
Management and Performance in the Public Sector: Evidence from German Municipalities

Florian Englmaier (LMU Munich)

Gerd Muehlheusser (University of Hamburg)

Andreas Roider (University of Regensburg)

Niklas Wallmeier (University of Hamburg)

Discussion Paper No. 397

May 14, 2023

Management and performance in the public sector: Evidence from German municipalities^{*}

Florian Englmaier[†] Gerd Muehlheusser[‡]
Andreas Roider[§] Niklas Wallmeier[¶]

May 14, 2023

Abstract

We study management practices and performance in a representative sample of German municipalities, which provide the bulk of direct administrative services for citizens and firms in Germany. Surveyed municipalities differ substantially in their use of structured management practices, and this heterogeneity is also pronounced within all federal states, regional types, and population size brackets. Moreover, we document a systematic positive relationship between the degree of structured management and a diverse set of performance measures capturing municipalities' attractiveness for citizens and firms. Topic modelling (LDA) of survey responses suggests that the predominant management style is to use relatively little structured management.

Keywords: management practices, public sector organizations, local government, municipal performance, state capacity, World Management Survey (WMS)

JEL Classification: D20, D73, H11, H73, R50

^{*}The project was preregistered at the Open Science Foundation (<https://osf.io/3k7sb/>). We gratefully acknowledge comments by Wouter Dessin, Ricard Gil, Anna Gumpert, Renata Lemos, Imran Rasul, John van Reenen, Rafaella Sadun, Daniela Scur, Uwe Sunde, Joachim Winter, and Mu-Jeung Yang. We also received very helpful suggestions from participants of the 2021 Meeting of the Committee for Organizational Economics of the German Economic Association, the 2022 SIOE conference in Toronto, the 2022 Firms and Labor Workshop in Holzhausen, the 2022 Empirical Management Conference in Washington D.C., as well as from seminar audiences at the Universities of Bonn, Hamburg, Regensburg and University College Dublin. Marte Anders, Anna Eisner, Lisa Eitinger, Pamela Fritzsche, Martina Kraus-Pietsch, Sebastian Kunz, and Leven Nendza provided excellent research assistance. We are grateful to the *Bertelsmann Stiftung* for giving access to their data, and to Dieter Daminger, Ulrich Hörning, and Siegfried Schloß for their input for our municipal survey. Further, support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged. Finally, we thank the *Kommunale Gemeinschaftsstelle für Verwaltungsmanagement (KGSt)*, and in particular Marc Groß, Laura-Elena Diehl, Sophia Eich, Anika Krellmann, and Volker Scharfen, for their support in conducting our survey.

[†]LMU Munich, Department of Economics, and CEPR, IZA, and CESifo; florian.englmaier@econ.lmu.de

[‡]University of Hamburg, Department of Economics, and IZA, and CESifo; gerd.muehlheusser@uni-hamburg.de

[§]University of Regensburg, Department of Economics, and CEPR, IZA, and CESifo; andreas.roider@ur.de

[¶]University of Hamburg, Department of Economics; niklas.wallmeier@uni-hamburg.de

1 Introduction

Motivation In advanced economies, the public sector is an important economic actor. For example, in the G7 countries, general government spending ranged between roughly 45% and well over 60% of GDP in 2021.¹ At the same time, recent topical events such as the COVID-19 pandemic have highlighted the importance of capable state institutions. However, for a long time, the question of state effectiveness has not been at the forefront of economic research. This might not be surprising given that the influential Public Choice school has generally held a negative view of public-sector organizations, for example by attributing a generally non-benevolent role to the state, and by seeing bureaucrats as agents who strive to expropriate citizens (see e.g., Tullock, 1967; Stigler, 1971; Peltzman, 1976). This line of reasoning has paved the way for an era of political leaders (such as Ronald Reagan and Margaret Thatcher) who viewed a smaller state as a better state. This perspective on state institutions might also explain why, for a long time, the question of what makes states effective has not been looming large in economic research. Instead, much of the literature has focussed on the (undoubtedly also important) optimal design of policies (e.g., optimal tax rates or optimal social insurance schemes), and not on how public services should be “produced” in an effective way.

More recently, state effectiveness has been put at centre stage by the literature on long-run growth and development, and this line of research has come to view the design of institutions as a key determinant of successful development (see e.g., Acemoglu and Robinson, 2012; Besley et al., 2021). Thereby, this literature picks up on a more positive view on states and public administrations, as for example brought forward in Max Weber’s (1922) seminal contribution. In his view, a formal, rule-driven operation comprising professional individuals with appropriate hierarchical delivery structures embodied a durable solution for implementing public policies. Weber (1922) also argues that running a state and running a firm, at least from a purely operational perspective, are not much different, and that the bureaucracies of (then) modern capitalist firms are good role models for public administrations. Of course, this raises the question of how to optimally structure such an organization.

For the private sector, substantial progress towards answering this question has been made as the empirical study of the management of private-sector firms has seen an unprecedented boom since the turn of the millennium. Pioneered by Ichniowski et al. (1997) and Bloom and Van Reenen (2007), there is by now a host of studies on *how* firms are run (see e.g. Bloom et al. 2012, 2015c and 2019, and the survey by Scur et al. 2021). In particular, these researchers have developed the *World Management Survey (WMS)*, a survey instrument where (production) managers at the establishment level are queried on the use of management practices in a broad set of domains (e.g., human resources and process management). Based on the survey responses, WMS studies typically construct a *management score* as a measure of “structured management”.² These studies

¹See e.g., <https://data.oecd.org/gga/general-government-spending.htm#indicator-chart>, accessed on February 13, 2023.

²For an overview of the many studies employing the WMS methodology, see also <https://worldmanagementsurvey.org>. While initially these studies were conducted as semi-structured phone surveys, later literature has also employed closed-form versions of the WMS (see e.g. Bloom et al., 2019). Our paper also adopts this approach. Moreover, we also

document that (i) how firms are managed can be meaningfully measured, (ii) there are persistent differences in the management practices employed, (iii) the degree of “structured management” is positively correlated with readily available performance measures such as profits, and (iv) there is a causal link from structured management to better firm outcomes (see e.g. Bloom et al., 2013).³

For the public sector, by employing the WMS approach, important insights have been gained on the role of management in hospitals, schools, universities, and in public projects in developing countries (for a survey, see again Scur et al., 2021). For example, a number of studies document that more structured management in hospitals is associated with lower mortality rates (Bloom et al., 2015b, 2020), in schools it goes hand in hand with higher SAT scores (Bloom et al., 2015a),⁴ and in universities it is associated with teaching and research performance (McCormack, Propper, and Smith, 2014). Moreover, Rasul and Rogger (2018) and Rasul, Rogger, and Williams (2021) find systematic variation of select management practices with task and project completion in the Ghanaian and Nigerian Civil Service, respectively.⁵

The private-sector and public-sector studies discussed so far have yielded very valuable insights. At the same time, they mainly focus on the role of management in achieving some well-defined goal in settings where informative performance measures are readily available. However, many public-sector organizations are legally obliged to pursue broader, more elusive goals such as the welfare of their respective stakeholders. This holds true for many regional bodies (such as states, counties, or municipalities). For the case of German municipalities (which provide a large fraction of public services in Germany and on which the present paper will focus), this is, for example, illustrated by Bavarian state law. It posits that municipalities shall “create and sustain (within the boundaries of economic feasibility) the public facilities that are necessary for the economic, social, and cultural well-being of its citizens”.⁶ Given such a broad, relatively elusive goal of improving the well-being of citizens, in comparison to other settings, it is also less clear how to measure performance relative to the goal. Thus, it would seem to make sense to simultaneously consider various dimensions of the performance of such public-sector institutions.

A similar argument applies to *comparing* performance *across* public-sector organizations. This is the case because, even if some public-sector organizations share the same goal (such as “maximizing well-being of citizens”), different organizations might pursue different strategies for achieving it. For example, while some might strive to attract new businesses, others might try to improve lo-

follow Bloom et al. (2019)’s definition of structured management practices as “those that are more specific, formal, frequent or explicit” (see their online appendix, p.2).

³Thereby, this literature contributes to explaining widely observed persistent productivity differences even in narrowly defined industries, which is, for example, documented in the overview article by Syverson (2011).

⁴See also Adelman et al. (2020) and Lemos et al. (2021).

⁵Outside the WMS approach, a number of papers have studied the efficacy of specific components of modern management in public sector organizations. For example, Bandiera, Prat, and Valletti (2009) explore the role of institutional characteristics (such as autonomy) for procurement efficiency in Italy. In the context of developing countries, recent studies have investigated the role of monitoring (Duflo, Hanna, and Ryan, 2012) and performance pay (Leaver et al., 2021) in schools, and career incentives in the health sector (Ashraf, Bandiera, and Lee, 2015). For an overview, see e.g. Besley et al. (2021). To support the build-up of effective bureaucracies in developing countries, the World Bank has established the *World Bank Bureaucracy Lab*, see <https://www.worldbank.org/en/research/dime/brief/Bureaucracy-Lab>.

⁶See Article 57 of the Municipal Code of the Federal State of Bavaria, <https://www.gesetze-bayern.de/Content/Document/BayGO>true> (own translation).

cal infrastructure, and yet another group might emphasize speedy decision-making. Hence, even if for some of these dimensions (e.g., the speed of decision-making) precise performance measures were available (which they are usually not), focussing on such selected performance measures might not allow for a meaningful comparison of performance across organizations given that some organizations might put more emphasis on other things to improve citizens' overall well-being. Thus, in absence of an objective aggregate measure of well-being, it seems paramount to measure performance of municipalities along a comprehensive set of separate metrics that pick up on these different directions that municipal activity might take.

Framework and results Against this backdrop, we believe that there is a need for a systematic empirical analysis of (i) the management practices employed in local administrations as one important class of public-sector organizations, which we conduct for the case of Germany, and (ii) the relationship between these management practices and performance across a broad set of relevant outcomes.

Our preregistered study makes an attempt in this direction and studies these two issues using a representative sample of German municipalities. Focussing on the municipal level is pertinent as municipalities play a core role for state capacity. For example, in Germany, they provide a vast number of public services such as registry services, issuing business licences, fostering economic activity, and child care. How well a municipality performs in this respect will affect its attractiveness for (new) citizens and enterprises and thus for economic prosperity.

One important contribution is to provide a comprehensive and systematic picture of the management practices employed in the administrations of German municipalities. Counseled by experts from our project partner and other practitioners, we developed a survey that is inspired by the various versions of the WMS. Thereby, we refined and trimmed the set of potential questions in order to fit the institutional setting in which German municipalities operate. This led to a total of 52 (sub-)questions from six categories: i) goal setting, (ii) monitoring of service quality, (iii) organizational learning, (iv) structured processes, (v) human resource management, and (vi) digitization, with the aim of measuring, in the spirit of the WMS, the degree of *structured management* in municipal administrations. We approached the top-level executives (i.e., mayors or heads of administration) of all 3,083 German municipalities with more than 5,000 inhabitants (covering roughly 80% of the German population) by email and invited them to participate in the survey. A response rate of 19% yielded 600 participating municipalities, covering a population of 16 million inhabitants and being representative of German municipalities with respect to a wide set of characteristics.

In a first set of results, we present an anatomy of the management practices employed in the administrations of German municipalities, where we document a large degree of heterogeneity with respect to employed practices. As a consequence, municipalities also differ widely with respect to the overall degree of structured management, which is measured by a *management score* as in Bloom et al. (2019). Interestingly, this heterogeneity is not driven by differences across states (which regulate municipal responsibilities and set the framework for administrations), by popu-

lation size, or by regional type.

We then analyze the empirical relationship between the management score and municipal performance. For this purpose, we use extensive municipality-level data provided by the Bertelsmann foundation (*Wegweiser Kommune*). The data set provides a set of indicators measuring municipalities' financial health and their attractiveness for citizens, businesses, and other stakeholders along a variety of dimensions. Taken together, we argue these indicators provide a sensible idea about how well municipalities perform with respect to their objectives. We find that municipalities differ widely with respect to these performance indicators. In a second set of results, we find a systematic positive and sizeable relationship between these municipal performance indicators and the management score. This finding is robust when controlling for municipal characteristics, and to a host of different specifications (including alternative measures of structured management such as the z-score). Hence, for the case of Germany, our paper clearly establishes a positive relationship between the degree of structured management and the performance of public administrations with respect to their objectives.

In a next step, we investigate in more detail the role of the six different management categories covered in our survey. Constructing a management subscore for each category, we find that it is mainly the use of practices related to the monitoring of service quality and human resource management that exhibit a systematic relationship with municipal performance. The answers to free-form questions in the survey corroborate the importance of human resource management for municipal performance.

Furthermore, we analyze *management styles*, i.e. which management practices are systematically used in combination. For example, some municipalities might focus on human resource management, while others might emphasize the monitoring of service quality instead. To investigate this, we resort to unsupervised machine learning and employ the *Latent Dirichlet Allocation* (LDA) introduced by Blei et al. (2003), which is more and more widely used in economics (see e.g. Hansen et al., 2018; Bandiera et al., 2020, Englmaier et al., 2022). We follow the approach of Bandiera et al. (2020), which posits two “pure” styles, and then characterizes every municipality as a convex combination of the two. It turns out that the LDA algorithm sorts municipalities along their management scores. In particular, on the one end of the spectrum are municipalities that score poorly on all six management categories, i.e. they pursue an “administered” style. By contrast, municipalities at the other end of the spectrum exhibit a high degree of structured management in all six categories, thereby pursuing an “actively managed” style. Hence, the analysis suggests that the municipalities in our sample do not specialize in different management categories to achieve their goals. Moreover, the “actively managed” style is associated with better municipal performance. However, “actively managed” municipalities are not too prevalent in our sample, and most municipalities rather closely follow the “administered” style.

The remainder of the paper is structured as follows: Section 2 provides some institutional background. Section 3 presents the details of the management survey (design, implementation, and responses) as well as the management score derived from it. Section 4 studies the relationship between the management score and municipal performance. Section 5 uses machine learning

techniques (LDA) to study municipal management styles. Section 6 discusses our findings and concludes. Appendix A contains a translation of the municipal survey (originally in German) and gives summary information on the responses. Appendix B contains additional tables and figures.

2 Institutional background

Germany is a federal state and broadly structured as follows: Below the federal level, there are 16 states (*Länder*). As of 2021, these are partitioned into 294 counties (*Landkreise*), which in turn are partitioned into a total of 10,799 municipalities (*Kommunen*).⁷

Of course, counties and municipalities are subject to federal and state law, but the German constitution grants them strong rights of self-government. For example, according to Article 28 of the constitution, “municipalities must be guaranteed the right to regulate all local affairs on their own responsibility within the limits prescribed by the laws”. As a result the vast majority of direct administrative services for citizens and businesses are provided at the local level, and the importance of municipal administrations is highlighted by the fact that they account for around 30 percent of total public service personnel in Germany (see e.g., Ruge and Ritgen, 2021, p. 124). In particular, municipal administrations provide services such as the maintenance and development of the local infrastructure (e.g., local roads, land-use plans, and development plans to promote the local economy), public facilities (e.g., parks, sports facilities), child care (e.g., public nurseries and kindergartens), and registry services (e.g., ID and civil status registration).⁸ For all services provided, municipalities have leeway how to organize and manage them.⁹

With respect to the decision-making process in municipalities, most states stipulate a dual structure, with a *representative body* of (elected) council members on the one side facing an *executive body* represented by the (elected) mayor on the other (see e.g., Ruge and Rittgen, 2021, pp.134.) The council holds the key management and control rights, while the mayor, in the role as *chief administrative officer*, is responsible for executing the decisions taken by the council and for managing all administrative matters.

To summarize, municipalities have a considerable degree of autonomy over the scope and the quality of the public services provided. Their performance in this respect contributes to a

⁷The different layers of public administration in Germany are described in more detail by Ruge and Ritgen (2021) and Schrappner (2021). Some states have an additional administrative layer formed by groups of counties (e.g., *Regierungsbezirke*).

⁸Note that in contrast to the U.S., German municipalities are, in general, not responsible for schools and the local police force, which are mainly administered at the state level (with some responsibilities at the county level). For a more detailed description, see Ruge and Ritgen (2021, p. 129). Counties are mainly responsible for services in areas such as social welfare and unemployment benefits, public health, garbage collection, immigration, motor vehicle registration, and building permits. At the time when our study was conducted, there were 107 (typically larger) municipalities in Germany that are not part of a county (*Kreisfreie Städte*). In these cases, all of the services usually provided by the relevant county are provided by the administration of the respective municipality as well.

⁹In terms of autonomy, there are three broad types of services that German municipalities deliver. Some services are mandatory and provided on behalf of the state or the federal level (e.g., registry services, support for federal or state elections). A second type of services is also mandatory, but there municipalities have more leeway to determine how to fulfil their obligations. For example, this is the case for the provision of public child care or fire brigades. Finally, there are services that municipalities can provide on a voluntary basis, and where they are free in how strongly to engage in such activities. Examples include sports and cultural facilities or projects to promote the local economy.

municipality's attractiveness for its citizens, businesses, and other stakeholders.

3 The municipal management survey

In this section, we discuss in detail our municipal management survey. Sections 3.1 and 3.2 explain the design and the implementation, respectively. Section 3.3 contains an overview over the questions and the main patterns in the survey responses. Finally, in Section 3.4, the *management score*, a measure for the degree of structured management, is constructed for municipalities in the spirit of Bloom et al. (2019). A translation of the questionnaire (originally in German) is provided in Appendix A, alongside with more detailed information about the answer categories and responses.

3.1 Survey design

To the best of our knowledge, our study is the first to provide a comprehensive and representative picture of the management practices employed in the local governments of a G7 country. For the case of German municipalities, we elicit the prevalence of various management practices in line with the literature building on the *World Management Survey*. This allows us to assess the degree of structured management employed in an important segment of public administrations.

While there is a large related literature that investigates this issue in the private sector, these studies do not provide an off-the-shelf template for our municipal survey. Rather, when devising the questionnaire for our public-sector setting, we needed to take into account the similarities and differences between private and public sector organizations. On the one hand, both when running a manufacturing firm and a municipal administration of equal size, one must set strategies, ensure a smooth operation of the day-to-day business, and recruit qualified employees as well as motivate them to do a decent job. On the other hand, municipalities have only a limited choice concerning which goods and services to provide, and their primary aim is certainly not profit maximization. Moreover, in Germany, as stipulated by public-sector labor law, municipalities have only limited scope for using standard human-resource management tools for incentivizing employees, such as incentive pay, the threat of dismissal, or perks such as company cars. Moreover, while there exist management surveys for (public) entities such as hospitals and schools, they are not readily applicable to our context of German municipal administrations, and hence, we have adapted established surveys as explained in the following.

For the design and the implementation of the survey, we have cooperated with the *Kommunale Gemeinschaftsstelle für Verwaltungsmanagement* (KGSt), a well-known non-profit organization, that specializes in consulting services for the administrations of its member municipalities.¹⁰ KGSt have extensive expertise when it comes to the institutional background of public administration in the federal structure of Germany, specific managerial aspects, and challenges in the public sector.¹¹

¹⁰Membership is voluntary and provides several benefits, but there is also a fee.

¹¹Importantly, KGSt was, however, not involved in the later data analysis or interpretation of our findings.

To design our (closed-form) survey, we started with a large set of potential questions (130+) from all versions of the *World Management Survey*, from which we selected questions and adapted them to fit our context.¹² Thereby, we consulted with KGSt as well as with several active and former mayors and other executives of municipal administrations to identify management practices that are potentially important for public-sector performance and where (given the high level of regulation) municipal administrations might differ.

In the resulting questionnaire we ask municipal executives (i.e. the mayor or, in large municipalities, the head of the municipal administration) about their management practices in six categories: (i) goal setting, (ii) monitoring of service quality, (iii) organizational learning, (iv) structured processes, (v) human resource management, and (vi) digitization. Specific questions and responses are discussed in more detail in Section 3.3.¹³

The answer categories for most of the questions are either binary (yes/no) or a 5-point Likert scale, but the questionnaire also contains some free-form questions. As explained in more detail in Section 3.4 below, we use the answers from a total of 52 closed-form questions and sub-questions to construct a measure for the degree of structured management.

3.2 Implementation of the survey and participation

The municipal survey was conducted online in April and May 2021. It was programmed in LimeSurvey, and all data was directly stored on a server of the University of Hamburg.

KGSt supported us logistically by sending out the invitation emails for the survey. Because municipal executives regularly receive invitations to participate in KGSt surveys, we were expecting a positive effect on the response rate from using this channel. In the communication with the municipalities regarding our survey, KGSt informed them that it was carried out together with us as part of a scientific project.¹⁴ In particular, invitations were sent to the mayors or heads of administration of all 3,083 municipalities with at least 5,000 inhabitants (in 2019). There were two main reasons for restricting the sample to these sufficiently large municipalities: First, one key aspect of our paper is to relate the management practices elicited in the survey to various measures of municipal performance. As detailed in Section 4 below, a rich set of performance measures is only available for municipalities with more than 5,000 inhabitants. Second, very small municipalities often do not have an administration on their own. Instead, they frequently join forces with other small neighboring municipalities to form an administrative entity (e.g. *Verbandsgemeinde*), and the administration for the whole entity is located in the main municipality. In the other municipalities of the entity, the local “administration” often consists solely of a mayor who holds this office as an honorary side job, possibly supported by a (part-time) clerk.

¹²In doing so, we proceed similarly to the *Management and Organizational Practices Survey* (MOPS), which is a closed form version the *World Management Survey* implemented by the U.S. Census Bureau, see e.g., Buffington et al. (2017).

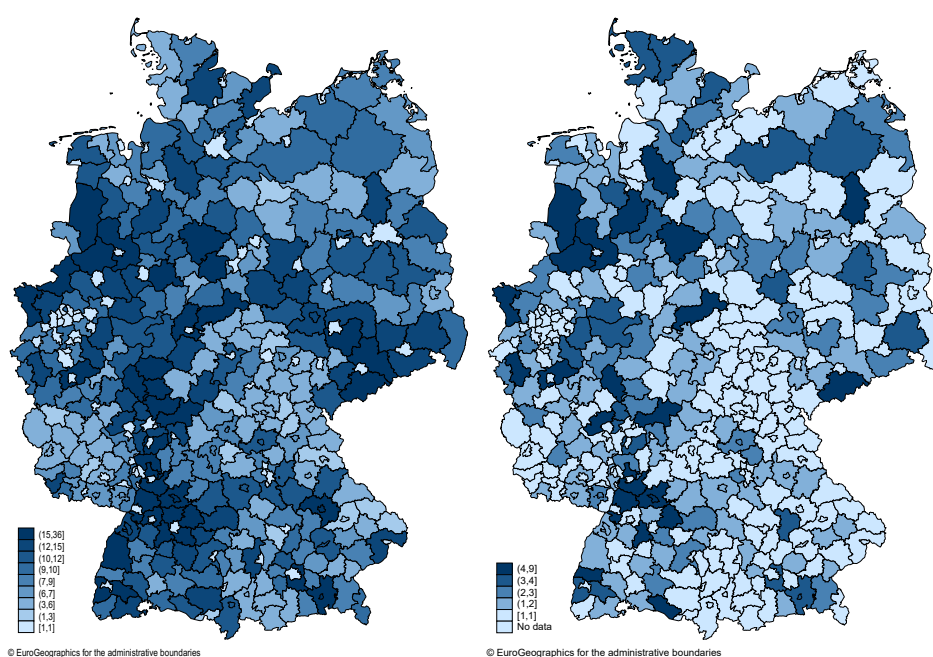
¹³With respect to (ii), as municipalities are legally required to record a variety of financial performance measures, in our survey we explicitly ask about the tracking of *service quality*, which is optional.

¹⁴To further incentivize participation, the invitation emails stated that 20 out of the responding municipalities would be randomly selected and invited to participate in an exclusive workshop “Challenges and Impulses for Public Management”, which we have conducted jointly with KGSt in February 2022.

Out of these 3,083 municipalities, 1,602 are KGSt members, and KGSt had the contact details of the mayor or other top executive in their database. In order to increase the number of observations, we also invited all 1,481 remaining municipalities with more than 5,000 inhabitants that are not KGSt member municipalities. To this end, we have hand-collected the respective contact details from municipalities' web sites, and sent personalized invitations to the mayor. To keep the channel of communication constant, also these invitations were sent out by KGSt.

In our analysis, we use those 600 out of the 3,083 municipalities contacted that have responded and completed at least 32 out of the 52 questions (including sub-questions) relevant for our management score.¹⁵ This corresponds to a response rate of 19.5 percent. We refer to these 600 municipalities as the “participating” municipalities, and they cover approximately 20 percent of the German population. Figure 1 provides an overview of the geographical distribution of the municipalities we approached (left panel) and the municipalities that participated (right panel). Both panels display broadly similar geographical patterns. As shown in Section 4, our sample is also representative with respect to municipal performance as well as population and regional type of municipalities.

Figure 1: Geographical distribution of invited municipalities (left panel) and participating municipalities (right panel) aggregated at the county level



Notes: Both panels show Germany partitioned into counties. In the left panel, the color indicates the number of municipalities with a population larger than 5,000 in the respective county (all of which were invited to the survey). In the right panel, the color indicates the number of municipalities in each county that have participated in the survey. To preserve anonymity, counties in which zero or one municipality participated are marked in the same color.

¹⁵In setting this threshold of 62 percent (32 out of 52 questions), we follow Bloom et al. (2019) who require at least 10 out of their 16 questions to be completed (63 percent). An additional 86 municipalities have participated in our survey, none of which has answered more than 27 questions, and 75% of these municipalities have answered at most 8 questions. These 86 municipalities were excluded from the analysis.

3.3 Overview of survey questions and responses

In the following, we provide a brief summary of the reported prevalence of the various management practices along the six main question categories described in Section 3.1 (see Appendix A for detailed information on the responses to each question).

First, in the set of questions on “goal setting”, we asked municipalities whether they have (easily accessible) “mission statements” directed either at citizens or at employees, whether these mission statements postulate goals, and the time horizon of these goals. Around 80 percent of municipalities do not have any mission statement directed at citizens, and around 20 percent of existing mission statements do not formulate any goals. A similar picture emerges with respect to mission statements directed at employees.

Second, with respect to the category “monitoring of service quality”, we inquired (i) whether applicable key performance indicators (such as processing times, waiting times, frequency of complaints) are tracked, (ii) whether customer satisfaction is elicited, and (iii) about the role such measures play for municipal management. It turns out that the use of such indicators is not at all widespread: In particular, more than 80 percent of municipalities do not use them at all to monitor service quality. Even when such indicators are used, about 85 percent of municipalities report that these indicators impact municipal management only to a small or intermediate degree.

Third, as for the category “organizational learning”, we focussed on (i) whether there are formal processes and incentives for employees to make suggestions, (ii) their impact on municipal management, and (iii) the sharing of management practices. It turns out that 20 percent of municipalities provide incentives for employee suggestions, but in 70 percent of these municipalities, the suggestions have at most an intermediate impact on management. In about half of the municipalities there is a regular exchange within the administration or with other municipalities about management practices.

Fourth, for the category “structured processes”, we had questions on whether tools such as project or process management are used. More than half of municipalities use project management and attach a high value to this tool, while process management is only employed by a minority of municipalities.

Fifth, as for the category “human resource management”, as in many other countries, there are legal restrictions on the use of (monetary) incentives in the German public sector. Nevertheless, municipalities have some leeway with respect to performance management and work organization, and our respective questions aimed to elicit to which degree this leeway is exploited. We find that the remuneration of employees reflects actual differences in performance only to a small degree. Non-monetary rewards for good performance are also scarce. Moreover, flexible working arrangements (in particular, the possibility to work from home) have been only rarely in place.¹⁶ To attract high-skilled labor (a major challenge in the public sector), only a small fraction of municipalities offers remuneration packages above the collective agreement level, while flexible work hours, and perks such as smartphones for office use are more common.

Finally, with respect to the category “digitization”, for a number of areas such as human re-

¹⁶Note that we explicitly asked about the prevalence of such arrangements *before* the COVID-19 pandemic.

source management, financial management, or procurement processes we asked (i) which fraction of data is stored digitally, and (ii) to which degree data-driven forecasting is employed. The idea is that digitized data is easily accessible, lends itself to straightforward analysis, and can serve as the basis for data-driven decision making. It turns out that the use of such tools is limited. Even for an area such as financial management a substantial fraction of municipalities reports that only part of the data is stored digitally. Moreover, very few municipalities use their data to make projections, e.g. to gauge the future demand for child care.

3.4 The management score of municipal administrations

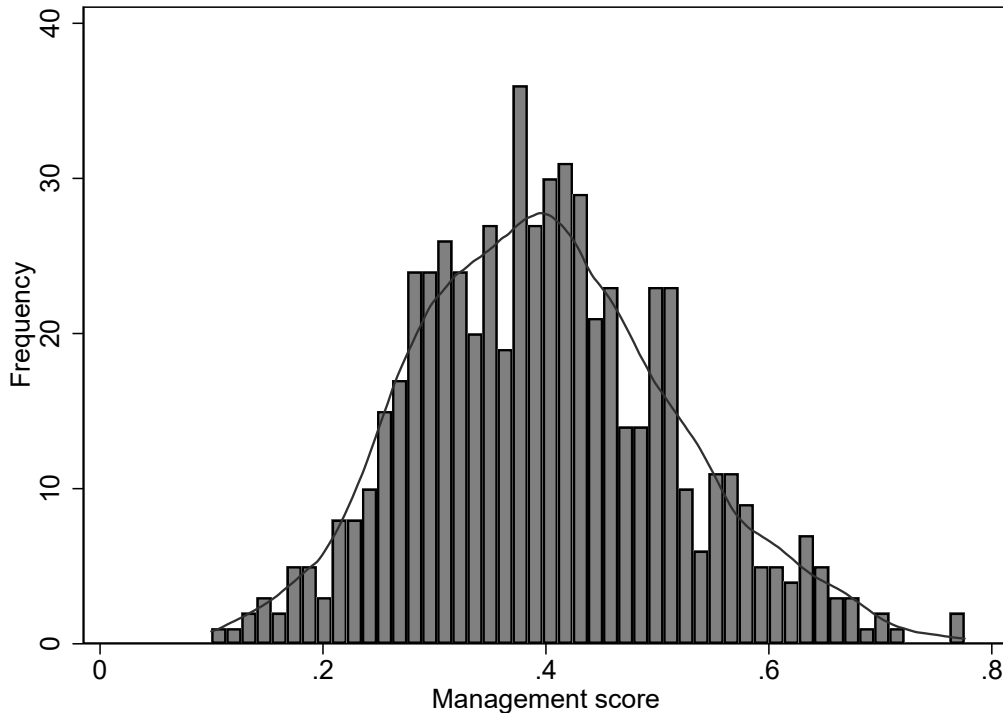
In a next step, we construct a *management score* (as, for example, in Bloom et al., 2019), thereby obtaining a measure for the degree to which a municipal administration relies on structured management. We proceed as follows.

The survey consists of a total of 49 questions, where 42 are directly linked to management practices.¹⁷ Out of these 42 questions, 39 do not contain sub-questions, while questions Q30, Q45, and Q46 contain 3, 5, and 5 sub-questions, respectively. This leads to a total of $39 + 3 + 5 + 5 = 52$ items entering the management score. In a first step, we normalize the responses to each of the 52 management questions on a 0-1 scale, where 0 (1) indicates the least (most) “structured” management practice. For questions with more than two possible answers, we assign equidistant values between 0 and 1. For example, for a question with five possible answers, we assign values of 0, 0.25, 0.5, 0.75 and 1. In Appendix A, we detail for each management question the number of answer options and which of them is assigned the lowest value of 0. In a second step, we calculate the *management score* as the unweighted average of the normalized responses to each of the management questions. By construction, the value of the score lies between zero and one, and higher values indicate a higher degree of adoption of structured management practices.

Figure 2 displays the frequency distribution of the management score across the participating municipalities. As already suggested by the discussion in Section 3.3, we find considerable heterogeneity. Given the relatively high degree of regulation in the public sector, this variation in the use of management practices seems remarkable. This documents that municipalities retain considerable leeway with regard to the employed management practices, and that these practices are not used uniformly across Germany. As municipalities are mainly regulated on the state level, one might suspect that the observed heterogeneity is mainly driven by variation *across states*. However, Appendix B.1 shows that there is considerable heterogeneity also within states as well as with respect to other observable characteristics of municipalities such as their regional types (metropolitan, urban, rural close to urban, and rural) or population size. Moreover, as discussed in Section 6, these variables explain only a small part of the variation in the management score.

¹⁷In the remaining seven questions we elicit (i) information regarding the relationship of the municipal administration with the local parliament (Q28), (ii) descriptive information for the respective municipal administration (Q29, Q42, Q43), and (iii) respondents’ opinions about future challenges for municipal management (free-form questions, Q47, Q48, Q49). See the discussions in Sections 4.3 and 6.

Figure 2: Frequency distribution of the management score



Notes: $N = 600$. Mean score: .40, standard deviation: .12, 25%-percentile: .31, 50%-percentile: .39, 75%-percentile: .48. The black line indicates the corresponding (Epanechnikov) kernel density.

4 The management score and municipal performance

In this section, we investigate the relationship between the management score and municipal performance. Obviously, this raises the question regarding the goals that municipalities are meant to pursue, and how to best measure municipal performance with respect to these goals.

In the private sector, firms are usually assumed to maximize their long-term value,¹⁸ and they are free to decide which goods and services to offer (and at which prices and quality levels) in order to best achieve this goal. Moreover, direct and informative performance measures such as annual profit, sales, market share, return on equity, or total factor productivity are often available.

This situation is different for many public sector organizations such as municipal administrations. First, for the vast majority of municipal services (including voluntary ones), the underlying goal is certainly not profit maximization. Rather, municipalities are supposed to follow a broader mission of providing an attractive environment for their stakeholders (in particular its citizens and the firms located in its constituency), taking into account the legal and institutional setting in which they operate. For example, according to Article 57 of the municipal code of the state of Bavaria, municipalities shall “create and sustain (within the boundaries of economic feasibility) the public facilities which are necessary for the economic, social and cultural well-being of its citizens”.¹⁹

Second, municipalities are constrained in pursuing their mission because they are less flexible

¹⁸For a critical discussion of the goals pursued by private firms, see e.g., Tirole (2001).

¹⁹See <https://www.gesetze-bayern.de/Content/Document/BayGO>true> (only available in German).

than private sector organizations in terms of their product portfolio and (personnel) policies. For example, municipalities are legally obliged to provide certain services (e.g. registry services) and hence cannot simply shut them down when considered undesirable with respect to their mission. Importantly, however, it is largely in the hands of municipalities to decide how exactly to manage the provision of both obligatory and voluntary services (e.g. how to organize citizen service centers or how intensely to engage in voluntary activities to attract new businesses), and some municipalities might turn out to be more apt in this than others. In this respect, and as already noted by Weber (1922), from a purely operational point of view, executives in municipal administration will have to cope with similar management challenges than their counterparts in the private sector.

Third, the broad mission of municipalities also renders the measurement of performance more challenging compared to private firms. For example, even if detailed information on some dimension of municipal performance were available (e.g. time to reach a decision on a given type of request), such a one-dimensional performance measure might not be suitable to compare performance across municipalities. The reason is that not all municipalities might optimize along this dimension to provide an attractive environment for their stakeholders. For example, while some municipalities might prioritize service quality, others might focus instead on improving infrastructure. As a consequence, we rely on a set of more comprehensive indicators of municipal performance, which we consider as (more) informative about the overall degree to which municipalities fulfil their mission as defined by law. These performance indicators (which are explained next) are obtained from a large German foundation, the *Bertelsmann Stiftung*.

4.1 Measures of municipal performance

The *Bertelsmann Stiftung* is one of the largest foundations in Germany, and they engage in a variety of projects in areas such as education, health, or the economy.²⁰ They also provide a database with rich annual statistical information on all German municipalities with a population size of at least 5,000, compiled from a variety of sources such as statistical offices, ministries, and other agencies (both at the state and federal level).²¹ Their database contains information at the municipal level from categories such as finance, demographics, education, infrastructure, and migration, and they kindly allowed us to use it for the present study.

In the light of municipalities' responsibilities and elusive goals as discussed above, and after counseling with KGSt and other expert practitioners, we consider a diverse set of eight measures from the Bertelsmann data that arguably are informative indicators of municipal performance. For all these indicators, higher values are associated with better performance.

First, with respect to whether a municipal constitutes an **attractive environment** for (new) citizens and (new) businesses, we consider the following indicators (all measured at the municipal level, monetary variables in Euro, variable names in italics):

1. *Net migration*: the net gain in population in a given year (per 1000 current inhabitants). We

²⁰See <https://www.bertelsmann-stiftung.de/en/home>.

²¹See their internet portal "Wegweiser Kommune" at <https://www.bertelsmann-stiftung.de/de/unsere-projekte/wegweiser-kommunede>.

use this as a measure for a municipality's attractiveness as a place to live for current and new citizens.

2. *Academic workforce*: the share of employees with an academic degree relative to the total employed workforce in the municipality. This variable measures the municipality's ability to attract high-skill jobs.
3. *Start-ups*: the number of newly founded firms per 1000 inhabitants. We use this variable as a measure of a municipality's attractiveness for founders.
4. *Broadband coverage*: the share of private households with an internet broadband connection of 50 Mbit/s or larger. We use this variable as a measure for the quality of the municipality's digital infrastructure.
5. *Business tax base*: per-capita tax base (i.e., loosely speaking, corporate profits) on which municipalities levy a business tax (*Gewerbesteuer*) to be paid by the establishments residing in the respective municipality. We interpret a high tax base as an indicator that a municipality provides an attractive environment for businesses to reside there and to enable them to generate profits. The business tax base is calculated by dividing a municipality's (per-capita) business tax revenue by the business tax rate (*Hebesatz*), which is set by municipalities.
6. *Income tax revenue*: the per-capita revenue from income tax paid by citizens and partnerships. In the same spirit as with the previous variable, we use this as an indicator that a municipality provides an attractive environment for citizens and partnerships to reside there and to generate income. Importantly, in contrast to the business tax, German municipalities have no influence on the income tax rate, which is set at the federal level. Instead, they receive as income tax revenue an amount that is by and large proportional to the income tax accruing in their constituency.²² Therefore, income tax *revenue* is arguably a good indicator for municipalities' income tax *base*.

German municipalities are granted a considerable degree of financial autonomy (see Article 28 of the German constitution). For example, they can (within legal and fiscal limits) decide on their expenditures (e.g. the size of the workforce in the administration or which infrastructure projects to finance) and the level of debt. As a result, also **fiscal health** seems to be an important measure of municipal performance, and we consider two basic indicators:

7. *Primary balance*: the difference (per capita) between income (e.g. from taxes) and expenditures, net of interest payments and repayment of debts. This is a prominent indicator for the overall financial health of a fiscal entity. A larger primary balance indicates that a municipality has more leeway, for example with respect to quality and/or scope of service provision or with respect to infrastructure investments.
8. *Financial balance*: the difference (per capita) between the financial income (e.g. dividends and interest) and financial expenditures (e.g. interest payments made to creditors). It mea-

²²For a detailed discussion of the distribution of income tax revenue from the federal to the municipal level, see https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Oeffentliche_Finzen/Foederale_Finanzbeziehungen/Kommunalfinzen/GemeindeanteilEst-2015.pdf (available in German only).

sures how strongly the municipal budget is constrained by obligations from outstanding debt.

It should be noted that in contrast to related studies in the private sector such as e.g. Bloom and Van Reenen (2007), Bloom, Sadun, and Van Reenen (2012, 2015c) or Bloom et al. (2019), any connection between the management practices employed in the municipal administration and these indicators will be rather indirect, as such practices cannot directly affect municipal performance. For example, when certain management practices in place ensure an effective provision of municipal services, this will not directly lead to more business tax revenue, but might be one factor in a firm’s decision to settle in the municipality and to pay its business tax there in the future. For this reason, our indicators will tend to be noisier measures of performance compared to those employed in studies of private sector firms. In this respect, our setting is somewhat stacking the deck against finding a (positive) statistically significant relationship between the management score and municipal performance. Nevertheless, our results reported below strongly support the view that such a relationship in fact exists.

Table 1: Indicators of municipal performance: Descriptive statistics for participating and non-participating municipalities

Indicator	Participate	N	Mean	SD	25%	50%	75%
Net migration	yes	583	7.43	7.3	3.6	6.8	9.7
	no	2,443	7.33	5.6	4.0	7.0	10.3
Academic workforce	yes	583	9.46	5.2	6.2	8.2	10.9
	no	2,444	9.72	5.5	6.4	8.5	11.4
Start-ups	yes	582	6.64	1.8	5.4	6.5	7.8
	no	2,443	6.56	2.2	5.2	6.5	7.7
Broadband coverage	yes	583	80.4	17.9	72.5	85.3	94.1
	no	2,444	79.5	20.0	71.3	86.4	94.3
Business tax base	yes	600	107.1	104.3	57.0	83.2	121.2
	no	2,483	116.5	187.5	57.1	85.3	125.9
Income tax	yes	600	462.2	134.8	368.7	460.8	544.6
	no	2,483	468.7	142.6	365.6	470.6	569.8
Primary balance	yes	600	228.1	191.9	119.0	200.0	312.6
	no	2,479	248.3	234.3	128.4	222.4	332.0
Financial balance	yes	600	-7.39	32.3	-21.3	-6.4	4.4
	no	2,479	-7.20	34.8	-21.4	-6.2	6.0

Notes: The table shows descriptive statistics for the eight indicators of municipal performance for both (i) the 600 participating municipalities for which a management score was calculated and (ii) the 2,483 non-participating municipalities with a population larger than 5,000 for which no management score could be calculated. The unit of observation for each indicator is the 5-year average from 2015 to 2019 (the five most recent years of availability) in the respective municipality. For some of the 600 participating municipalities, some indicators are not available during the five-year period. *Business tax base*, *Income tax*, *Primary balance* and *Financial balance* are measured per capita (in Euro), *Net migration* and *Start-ups* are measured by 1,000 inhabitants, *Academic workforce* and *Broadband coverage* are measured as shares.

The Bertelsmann data set covers the period from 2006 until 2019, which is the most recent wave. For our baseline analysis, we take 5-year averages over the period 2015-2019 to smooth out

fluctuations.²³ Table 1 provides descriptive statistics for our performance indicators. As can be seen, municipalities are widely dispersed with respect to all measures. Moreover, our sample of the 600 municipalities that have participated in the management survey seems to be representative in the sense that the differences between participating and non-participating municipalities are quite small. For example, as for mean comparisons, seven out of the eight corresponding t-tests are not statistically significant at the 10 percent level, and only for *Primary balance* there is a significant difference between the two means ($p = .05$). As shown in Table 2, also the pairwise correlations between our eight indicators are relatively small.

Table 2: Pairwise correlations of indicators of municipal performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Net migration	1.000							
(2) Academic workforce	0.112*	1.000						
(3) Start-ups	0.148*	0.206*	1.000					
(4) Broadband coverage	0.132*	0.241*	0.331*	1.000				
(5) Business tax base	0.078	0.397*	0.319*	0.186*	1.000			
(6) Income tax	0.053	0.351*	0.453*	0.333*	0.284*	1.000		
(7) Primary balance	0.170*	0.178*	0.348*	0.166*	0.698*	0.415*	1.000	
(8) Financial balance	0.009	0.155*	-0.041	0.031	0.196*	-0.105	-0.030	1.000

Notes: * indicates statistical significance at the 10 per cent level.

4.2 Main results

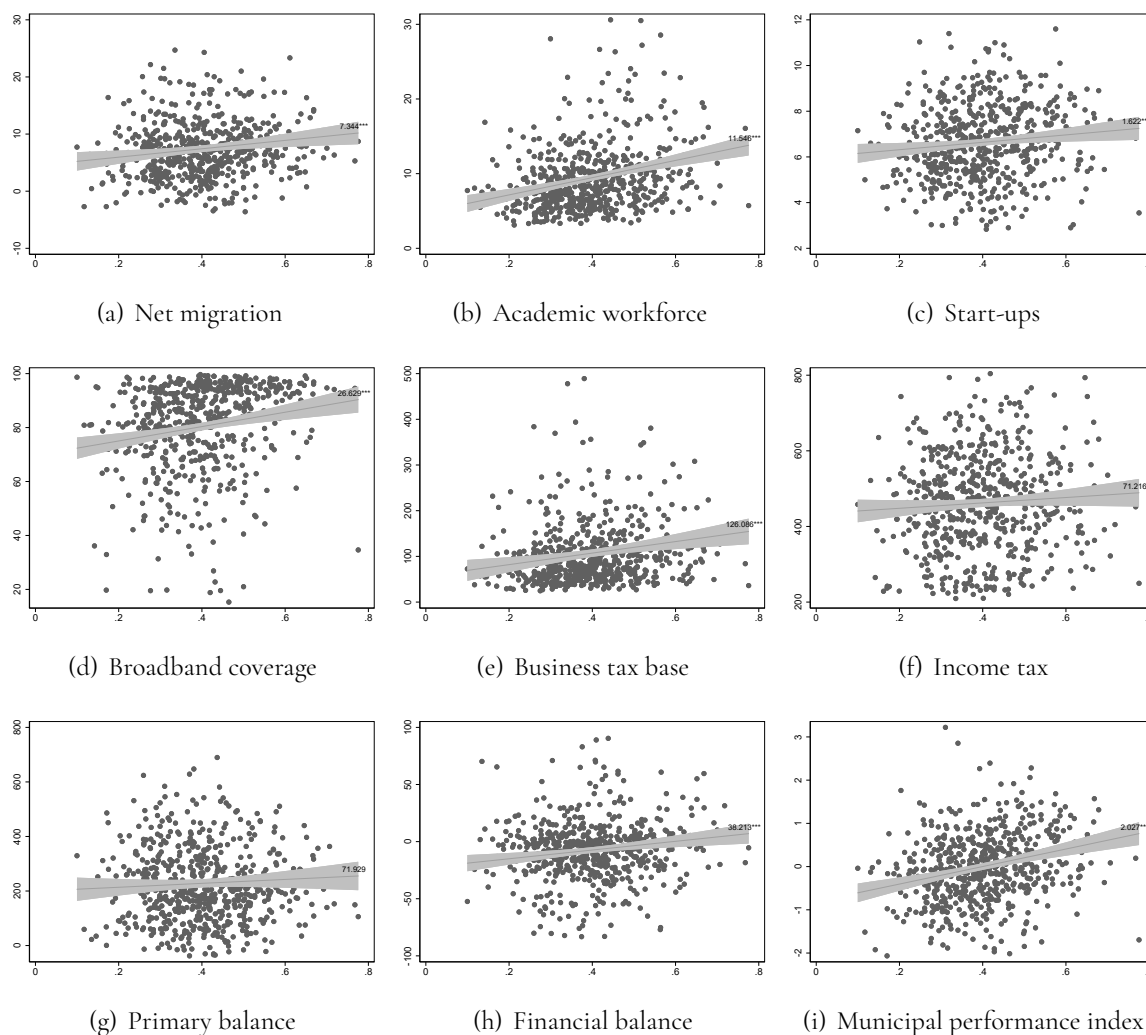
We now present our results on the empirical relationship between the management score and the eight municipal performance indicators. As a first step, we consider the raw data without any controls. Figure 3 displays scatter plots relating each indicator (on the vertical axis) to the management score (on the horizontal axis). Each panel of Figure 3 also includes a fitted regression line and the respective 95% confidence interval resulting from a simple univariate OLS regression without clustering of standard-errors. Note that for all eight indicators, municipal performance is positively related to the management score, and in six cases this relationship is statistically significant.

In a second step, we perform a more thorough regression analysis with the respective indicator as the dependent variable and the management score as the key independent variable. For each indicator of municipal performance, we estimate four regression models.²⁴ The first (baseline) model does not include additional controls. In the second model, we use state dummies to control for state-specific effects (e.g. legal, political, and economic factors) under which the respective municipalities located in a given state operate. In the third model, we additionally include dum-

²³In robustness checks, we verify that our main results uphold also when considering alternative time periods (see Appendix B.3).

²⁴In all of these regressions, standard errors are clustered at the state level. Participating municipalities are located in 13 out of the 16 German states. Given the relatively low number of 13 clusters, we employ a (wild cluster) bootstrap procedure.

Figure 3: Management score and municipal performance measures (raw data)



Notes: Each panel depicts the management score on the horizontal axis and the respective performance measures on the vertical axis (averages over the 5 most recent years of availability, 2015 to 2019). Each panel also show the fitted line and the 95% confidence interval resulting from a univariate OLS regression (without clustering of standard-errors) between the management score and the respective performance measures, where *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

mies that control for the type of region in which a municipality is located. Thereby, in line with the *RegioStaR* classification scheme provided by German Federal Statistical Office (see Destatis, 2021, p.16ff), we distinguish four regional types of municipalities: (i) metropolitan, (ii) urban, (iii) rural close to urban, and (iv) rural.²⁵ Finally, in the fourth model, in addition to state and regional type dummies, we control for the size of municipalities. While all of our performance indicators are defined on a per capita basis or as shares, it might still be the case that larger municipalities differ qualitatively from smaller ones. To address this issue, we introduce dummies for six population size brackets based on the GKBK7 categorization of the German Federal Statistical Office (see Destatis, 2021, p.44): (i) 2,001 - 5,000, (ii) 5,001 - 20,000, (iii) 20,001 - 50,000, (iv) 50,001

²⁵Note that most German states feature all four regional types, and that regional types are not constrained by state boundaries. That is, a small rural municipality that lies close to a regipolitan city in a neighboring state will be classified as “rural close to urban”.

- 100,000, (v) 100,001 - 500,000, and (vi) more than 500,000 inhabitants. While municipalities' exact population size varies over time (and might depend on how well a given municipality is managed), the population size bracket does not vary across the sample period 2015-2019 for 99% of the 600 participating municipalities.²⁶ Finally, Table 7 in Appendix B.2 displays the distributions of participating and non-participating municipalities across regional types and population brackets, respectively. Kolmogorov-Smirnov tests indicate that our sample is also representative along these dimensions as there are no statistically significant differences between these distributions.

In addition to the eight performance indicators, we also construct a composite measure, the *Municipal Performance Index* (MPI). The construction of the MPI follows the standard procedure for *z-scores*. In a first step, we determine the *z-score* for each of the eight indicators, i.e. we normalize it to exhibit a mean of zero and a standard deviation of one. In a second step, the average of these eight *z-scores* yields an index value for each municipality. Finally, normalizing the index values again to exhibit a mean of zero and a standard deviation of one then yields the MPI used in the analysis.

The results of all regressions are shown in Table 3, the structure of which requires some explanation: Each row corresponds to a given performance indicator as dependent variable, and shows the coefficient and *p-value* of the management score in each of the four regression specifications discussed above (while the coefficients of control variables are not displayed). For example, the entry 7.34 (5.92) in the upper left (right) cell is the coefficient of the management score for the effect on net migration when not adding any further controls (when controlling for state, regional type, and population bracket).

The regression analysis largely confirms the findings of Figure 3. In particular, all 36 coefficients for the management score indicate a systematic positive relationship with performance. Moreover, 25 of them are statistically significant, and in the other cases the *p-values* are generally not far off. For the MPI, the coefficient of the management score is significant in all four specifications. Our findings thus clearly point to a systematic positive relationship between the degree of structured management in municipal administrations (as measured by the management score) and municipal performance.²⁷

As shown in Appendix B.3, the main results as reported in Table 3 are qualitatively robust (i) when averaging the outcome measures over the three most recent years (instead of five years) and when looking at the most recent year only (see Tables 8 and 9, respectively), and (ii) when using the *z-score* of the survey answers instead of the management score (see Table 10), which is an alternative measure of structured management employed in the WMS literature (see e.g., Scur et al., 2021).²⁸

In order to gauge the economic significance of our findings, we perform an interquartile com-

²⁶Seven out of these 600 municipalities are assigned to a different population bracket in one out of the five years of the sample period. To these municipalities we assign the 2015 population bracket, which in each of these cases was also the modal value. The results of column (4) of Table 3 remain robust when dropping these seven municipalities from the sample.

²⁷With respect to column (4), note that there is a positive and significant correlation between population size and the management score ($r = .27$), which might pick up part of the effect of the latter.

²⁸The correlation between our management score and the *z-score* is 0.98, which is significant at the 1 percent level.

Table 3: Management score and municipal performance

	(1)	(2)	(3)	(4)
Net migration	7.34*** (0.00)	7.27*** (0.00)	6.68*** (0.00)	5.92*** (0.00)
Academic workforce	11.55*** (0.00)	11.46*** (0.00)	9.93*** (0.00)	4.16 (0.18)
Start-ups	1.62*** (0.00)	1.24*** (0.00)	0.96** (0.01)	0.49 (0.25)
Broadband coverage	26.63*** (0.00)	27.41*** (0.00)	24.48*** (0.00)	11.34 (0.13)
Business tax base	126.09*** (0.01)	144.66*** (0.01)	140.35** (0.03)	117.79 (0.14)
Income tax revenue	71.22 (0.17)	80.71*** (0.00)	37.32 (0.14)	52.24** (0.02)
Primary balance	71.93 (0.27)	126.36 (0.12)	124.98 (0.12)	127.79 (0.20)
Financial balance	38.21** (0.03)	36.50** (0.04)	36.12*** (0.00)	20.95 (0.14)
Municipal performance index (MPI)	2.03*** (0.00)	2.09*** (0.00)	1.85*** (0.00)	1.22** (0.04)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: Each row corresponds to a given performance indicator as dependent variable (5-year averages), and shows the coefficient and p-value of the management score in each of the four regression models discussed in the main text (while the coefficients of control variables are not displayed). For example, the entry 7.34 (5.92) in the upper left (right) cell is the coefficient of the management score for the effect on net migration when not adding any further controls (when controlling for state, regional type, and population bracket). In all regressions, standard errors are clustered at the state level (employing a wild cluster bootstrap), where *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The number of observations is 600, except for *Net migration*, *Academic workforce* and *Broadband coverage* where it is 583, and for *Start-ups* where it is 582 (see Table 1).

parison between municipalities with a management score at the 25%- and the 75%-percentile, where the respective score values are .31 and .48, leading to a difference of .17. In Table 4, for each performance measure we then evaluate this score difference of .17 with the respective coefficient of the management score in the richest regression specification (4) as reported in Table 3. That is, the first entry in column (4) of Table 4 denotes the absolute change of the respective performance measure, and is obtained by multiplying the respective coefficient in column (3) by

Table 4: Economic significance

	(1)	(2)	(3)	(4)
Indicator of municipal performance	Mean	SD	Coefficient of management score	Effect of interquartile score change of 0.17
Net migration	7.4	7.3	5.92	1.01 (14%, 14%)
Academic workforce	9.5	5.2	4.16	0.71 (7%, 14%)
Start-ups	6.6	1.8	.49	0.08 (1%, 5%)
Broadband coverage	80.4	17.9	11.34	1.93 (2%, 11%)
Business tax base	107.1	104.3	117.79	20.02 (19%, 19%)
Income tax	462.2	134.8	52.24	8.88 (2%, 7%)
Primary balance	228.1	191.9	127.79	21.72 (10%, 11%)
Financial balance	-7.4	32.3	20.95	3.56 (-48%, 11%)

Notes: Columns 1 and 2 display the mean and standard deviation of each of the eight indicators of performance of the participating municipalities (see also Table 1). Column 3 re-states the coefficient for the management score from the richest regression specification (4) in Table 3. Column 4 shows the absolute ceteris-paribus change in the respective performance measure for the interquartile difference in the management score (with values of .31 and .48 at the 25% and 75% percentile, respectively, leading to a difference of .17). That is, the entries in column 4 are obtained by multiplying the respective coefficient in column 3 by .17. The first (second) number in brackets denotes the absolute change as a fraction of the respective mean (standard deviation). For example, for the case of net migration, we have: $1.01 = .17 \cdot 5.92$ ($14\% = 1.01 \cdot 100/7.4$, $14\% = 1.01 \cdot 100/7.3$).

the score difference of .17. The two entries in brackets state this absolute change as a fraction of the mean and the standard deviation, respectively, for the performance indicator under consideration. For example, for the case of net migration, the absolute change is $1.01 = .17 \cdot 5.92$, which constitutes 14% of both the mean (7.4) and the standard deviation (7.3). Overall, the economic effects appear to be sizeable.

Keeping in mind that the link between the use of structured management in municipal administrations and the measures of municipal performance is arguably rather indirect, our findings seem all the more intriguing.

4.3 Management subscores and municipal performance

Recall from Section 3.1 that our survey contains questions from six different management categories (goal setting, monitoring of service quality, organizational learning, structured processes, human resource management, and digitization). In this section, we analyze the role of these categories in more detail.²⁹ In a first step, we construct a *management subscore* for each category from the respective survey questions.³⁰ We then regress municipal performance (measured by the composite Municipal Performance Index (MPI), as introduced in Section 4.2) on these subscores. The results are shown in Table 5. We consider the same four regression specifications as in main Table

²⁹This part of the analysis was not addressed in the preregistration. We have nevertheless included it as it turns out to be helpful for the subsequent (preregistered) analysis of municipal management styles.

³⁰The subscores are constructed in the same way as the overall management score used so far. See Appendix A for more information on which questions enter which subscore.

Table 5: Management subscores and municipal performance

	(1)	(2)	(3)	(4)
Subscore				
goal setting	-0.17 (0.16)	-0.03 (0.66)	-0.08 (0.23)	-0.22*** (0.00)
monitoring of service quality	0.57* (0.07)	0.69** (0.03)	0.63** (0.01)	0.43 (0.18)
organizational learning	0.27 (0.45)	0.41 (0.20)	0.42 (0.17)	0.46 (0.13)
structured processes	-0.17 (0.87)	0.23 (0.13)	0.10 (0.38)	-0.05 (0.73)
human resource management	1.95*** (0.00)	0.82* (0.06)	0.74** (0.01)	0.72*** (0.01)
digitization	0.12 (0.60)	0.06 (0.72)	0.13 (0.52)	0.09 (0.64)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The table reports regressions where the dependent variable is the municipal performance index (MPI), i.e., the z-score of the eight performance indicators. The key independent variables are the management subscores for the six categories of questions in the survey. Specification (1) corresponds to a baseline with no additional controls, while specifications (2) – (4) consecutively include dummies for state, regional type, and population size (as explained in Section 4.2). Each entry in the table shows the coefficient and the p-value for the respective management subscore. In all regressions, standard errors are clustered at the state level (employing a wild cluster bootstrap), where *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

3 above, but in Table 5 each column refers to just one regression, i.e. each coefficient indicates the effect of the respective subscore on the MPI. As can be seen, it is particularly the use of management practices related to the monitoring of service quality and human resource management that exhibit a statistically significant relationship with municipal performance.

The topicality of human resource management policies is corroborated by the answers to the three free-form questions included in the municipal survey (Q47-Q49). We asked respondents what they perceive as major challenges for implementing high-quality municipal services, and it turns out that topics related to human resource management, such as flexible remuneration, ensuring an attractive workplace, or attracting talent loom large in the answers. Taken together, this suggests that in discussions of changing management practices human resource management should receive particular attention. See Section 6 for a more detailed discussion.

The analysis of subscores provides some first indication on the relative importance of the

various management categories. In a next step, we analyze whether management practices are systematically used in combination – thereby giving rise to distinct municipal *management styles*– and whether there exists a systematic relationship between management styles and municipal performance. In the subsequent section, we resort to tools from unsupervised machine learning to study these issues in more detail.

5 Municipal management styles

In this section, we investigate in more detail potentially distinct patterns in the answers to the survey questions. Thereby, our aim is to gain a better understanding of different management styles employed in German municipalities.

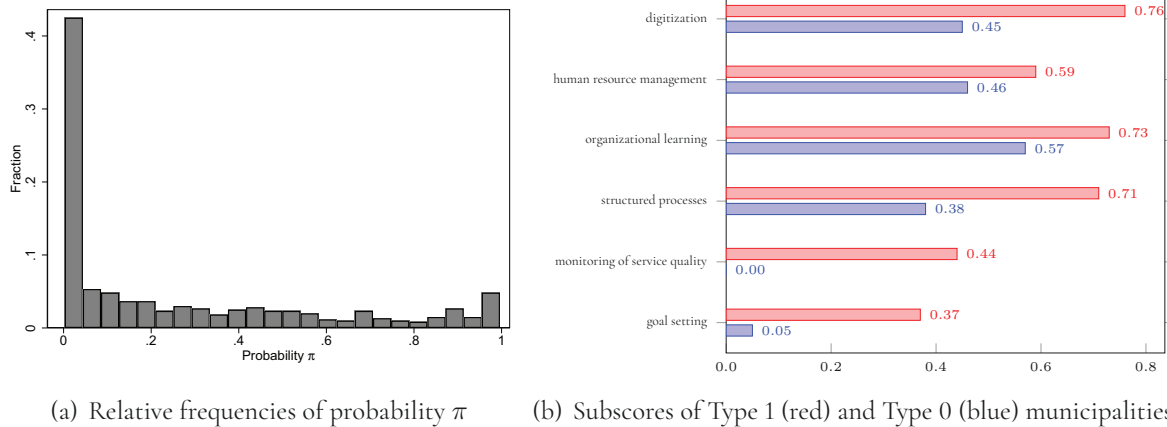
To analyze the high-dimensional survey data, we apply machine learning techniques for pattern detection in the survey responses in order to classify municipalities according to their management styles. In particular, we consider the *Latent Dirichlet Allocation* (LDA) (Blei et al., 2003), an unsupervised learning algorithm. Compared to other algorithms (such as principal component analysis or k-means analysis), the LDA allows to include a large number of variables in the classification and also to take into account correlations between combinations of the variables. The LDA is also a suitable tool for the analysis of survey data (see e.g., Airolidi et al., 2014). In the context of economics, it has recently been employed to analyze central bank protocols (Hansen et al., 2018), CEO management styles (Bandiera et al., 2020), and management practices in Spanish manufacturing firms (Englmaier et al., 2020).

Our approach follows Bandiera et al. (2020). That is, we postulate two “pure” management styles, which the LDA then extracts from the survey responses. In a next step, each municipality’s management style is then expressed as a convex combination of these two pure styles. In particular, to each municipality the LDA assigns probabilities π and $1 - \pi$ as the probability weights on either pure style. Panel (a) of Figure 4 shows the histogram of the probability π across municipalities. As can be seen, roughly 30 percent and 5 percent of municipalities closely resemble one of the pure styles, respectively. All other municipalities are (more or less uniformly) distributed between these extremes.

In a next step, we investigate key features of municipal administrations whose management style either exactly or very closely matches one of the two pure styles, where we use $\pi \leq 0.1$ and $\pi \geq 0.9$ as cutoffs, and refer to these municipalities as “Type 0” and “Type 1”, respectively. Panel (b) of Figure 4 shows the differences between the two styles for each of the six management subscores as introduced in Section 4.3. As can be seen, municipalities of Type 1 exhibit substantially higher scores in all six categories compared to their Type 0 counterparts. Hence, the analysis suggests that the municipalities in our sample do not specialize in different management categories.

The difference between the two types is particularly stark for the categories *monitoring of service quality* and *goal setting*, where municipalities of Type 0 are hardly active (if at all), as indicated by management subscores that are (close to) zero. In the light of these differences, one might refer to municipalities of Type 1 and 0 as “actively managed” and “administered”, respectively.

Figure 4: Management styles using the LDA



Strikingly, the latter constitute half (49.8%) of all municipalities in our sample. Moreover, 78% of municipalities are closer to the “administered” style than to the “actively managed” style (i.e. they are characterized by $\pi \leq 0.5$).

Using the LDA results, we have also investigated the relationship between each municipality’s management style and municipal performance. In particular, we use the probability weight π on the “actively managed” Type 1 as an alternative measure of structured management (instead of the management score used in main analysis). It turns out that these two measures are positively correlated (pairwise correlation: 0.70, significant at the 1 percent level). Moreover, Table 11 in Appendix B.3 replicates the analysis of main Table 3 using this alternative measure of structured management, and the results are qualitatively robust.

6 Discussion and conclusion

In this paper, we study management practices in public-sector organizations. In particular, we have conducted a survey among German municipalities eliciting the relevance of various management practices in their administrations. The survey gives us a sample of 600 participating municipalities that is representative along a variety of dimensions. As our first key finding, we document that municipalities differ widely in how intensively various management practices are employed.

We then focus on municipal performance. As municipalities arguably pursue more elusive goals (i.e., fostering the well-being of relevant stakeholders) than private firm or other public-sector organizations (such as schools or hospitals), this makes municipal performance harder to define and noisier to measure. We use data from the *Bertelsmann Stiftung*, which contains a diverse set of indicators measuring municipalities’ financial health and their attractiveness for citizens, businesses, and other stakeholders along a variety of dimensions.³¹ Our second key finding is

³¹In the preregistration, in addition to the performance indicators provided by the *Bertelsmann Stiftung*, we had also envisioned to consider direct measures of municipal performance in the form of customer satisfaction data. Since such data does not exist on the municipal level in Germany, we have designed and conducted two online surveys (one among Facebook users with a focus on private citizens and one among all firms registered in any of the 600 municipalities

the existence of a systematic positive and sizeable relationship between these municipal performance indicators and the degree of structured management, and this finding is robust to a host of different specifications. Thereby, practices related to human resource management and towards monitoring the service quality seem to play an important role.

Furthermore, the use of unsupervised machine learning (in particular, the LDA) allows to classify municipalities' survey responses into two polar management styles of running German municipal administrations, "administered" and "actively managed". The latter are characterized by a more intensive use of structured management practises across all six management categories captured in the survey, and actively managed also exhibit better municipal performance.

All in all, our paper points to a stable empirical relationship between the degree of structured management and the performance of municipal administrations. Thereby, comparable findings from other sectors are also corroborated for this important class of public-sector organizations, which pursue more elusive goals.

In future research, it would be interesting to investigate in more detail how the large heterogeneity in the use of structured management practices in municipal administrations comes about. Looking at the distribution of management scores suggests that the main drivers of heterogeneity are not municipalities' geographical location (state), regional type, or size (see Figures 5-7 in Appendix B.1). This observation is corroborated by a regression analysis (reported in Table 6 with the management score as the dependant variable and the controls for state, regional type, and population bracket as right-hand side variables (added consecutively)). As shown in columns (1)-(3) of Table 6, the adjusted R^2 -values indicate that these variables only explain a (very) small fraction of the variation in the management score. Out of these variables, the population bracket dummies seem to play the biggest role.

Table 6: Potential sources of the heterogeneity in the management score

	(1)	(2)	(3)	(4)
Adjusted R^2	0.01	0.01	0.13	0.13
State dummies	Yes	Yes	Yes	Yes
Regional type dummies	No	Yes	Yes	Yes
Population bracket dummies	No	No	Yes	Yes
Historic primary balance	No	No	No	Yes

Notes: The table reports results from OLS regressions with the management score as dependent variable. In specifications (1)-(4), state dummies, regional type dummies, population bracket dummies, and the variable primary balance (averaged over the years 2006 and 2007) are consecutively added.

In principle, heterogeneity in the management score could have emerged if municipalities that were relatively affluent in the past exhibit a higher degree of structured management today, e.g.

participating in our management survey). In both of these surveys, the number of respondents was unfortunately so low (less than 1 per cent, leading to very low numbers of observations per municipality) such that the obtained data was not amenable to sound statistical analysis. We have also elicited the Google reviews for the municipal offices (e.g. civil registry) of all 600 participating municipalities. Again, for the vast majority of municipalities, the number of reviews was very small and hence too noisy to be useful.

because they had sufficient financial resources to systematically evaluate, refine, or change their management policies.³² One possibility to investigate this issue is to approximate a municipality's historic prosperity by its (average) primary balance in the past (where the primary balance is a key measure of financial health as defined in Section 4.1 above). However, we are limited by the fact that our data go only back to 2006 and municipalities' current sets of practices might very well have emerged before that date. Therefore, as a preliminary exercise, we consider the average primary balance over the first two years contained in our data set (2006-2007), and we include this measure as an additional control in the regression analysis of Table 6. Column (4) reveals that this has no effect on the explained variation in the management score (again measured by the adjusted R^2).³³

Apart from structured management, a recent literature studies the importance of managers (and more generally, human capital) for the performance of public sector organizations (see e.g., Dwenger and Gumpert, 2023; Fenizia, 2022; Hjort et al., 2021; Decarolis et al., 2020, for causal evidence).³⁴ While we do not have information on the composition of the workforce of the municipal administrations in our sample, in our survey we elicited the size of the workforce, i.e. the number of employees in (core parts of) the respective administration (Q29, see Appendix A). Interestingly, our data reveal a positive relationship between the management score and the number of municipal employees per capita (of the population); a finding reminiscent of the positive relationship between the management score and firm size documented for private-sector (manufacturing) firms in the WMS literature. In principle, it could be that, independent of the degree of structured management, municipalities with a larger workforce score better with respect to our performance indicators. To investigate the robustness of our main analysis (see Table 3), we include the number of (per-capita) employees in the administration as an additional right-hand side variable. As shown in Table 12 in Appendix B.3, this does not qualitatively affect our findings on the role of the management score.

Taking together our results and the results of prior literature on private sector firms, schools, hospitals, and public administrations in developing countries, we would argue that there is by now compelling evidence for a positive link between the degree of structured management practices and the effectiveness of public service provision. Given the severe challenges for states' capacities that lie ahead, and the already widely felt strain on public administrations, we are convinced that any upcoming public service reforms should be inspired by these findings. Such an evidence-based approach should be employed not only at the planning stage, but also during implementation, for example through field experiments. This would allow us to learn more about the exact mechanisms through which structured management practices facilitate public service provision.

³²The financial resources needed to implement the management practices surveyed are of course open to debate.

³³In Germany, there is a long-standing debate about whether heavily indebted municipalities rely too much on checking account overdrafts (*Kassenkredite*). This suggests that checking account overdrafts could be an alternative measure for (historic) prosperity. However, using the 2006-2007 average of this variable (instead of primary balance) does not affect the explained variation of the management score relative to column (3).

³⁴Carreri and Payson (2021, 2023) survey roughly 300 US mayors and city managers to elicit their management skills and show that they are positively associated with municipal performance. See Carreri (2021) for a similar study on Italian mayors.

References

- ACEMOGLU, D. AND J. ROBINSON (2012): *Why Nations Fail: Origins of Power, Poverty and Prosperity*, Crown Publishing Group.
- ADELMAN, M., J. BARON, AND R. LEMOS (2020): “Managing shocks in education: Evidence from hurricane Matthew in Haiti,” *World Bank, mimeo*.
- AIROLDI, E. M., D. M. BLEI, E. A. EROSHEVA, AND S. E. FIENBERG (2014): “Introduction to mixed membership models and methods.” *Handbook of Mixed Membership Models and Their Applications*, 100, 3–14.
- ASHRAF, N., O. BANDIERA, AND S. LEE (2015): “Do-gooders and go-getters: career incentives, selection, and performance in public service delivery,” *Harvard Business School, mimeo*.
- BANDIERA, O., A. PRAT, S. HANSEN, AND R. SADUN (2020): “CEO behavior and firm performance,” *Journal of Political Economy*, 128, 1325–1369.
- BANDIERA, O., A. PRAT, AND T. VALLETTI (2009): “Active and passive waste in government spending: evidence from a policy experiment,” *American Economic Review*, 99, 1278–1308.
- BESLEY, T. J., R. BURGESS, A. KHAN, AND G. XU (2021): “Bureaucracy and development,” *NBER Working Paper No. 29163*.
- BLEI, D. M., A. Y. NG, AND M. I. JORDAN (2003): “Latent Dirichlet allocation,” *Journal of Machine Learning Research*, 3, 993–1022.
- BLOOM, N., E. BRYNJOLFSSON, L. FOSTER, R. JARMIN, M. PATNAIK, I. SAPORTA-EKSTEN, AND J. VAN REENEN (2019): “What drives differences in management practices?” *American Economic Review*, 109, 1648–83.
- BLOOM, N., B. EIFERT, A. MAHAJAN, D. MCKENZIE, AND J. ROBERTS (2013): “Does management matter? Evidence from India,” *Quarterly Journal of Economics*, 128, 1–51.
- BLOOM, N., R. LEMOS, R. SADUN, AND J. VAN REENEN (2015a): “Does management matter in schools?” *Economic Journal*, 125, 647–674.
- (2020): “Healthy business? Managerial education and management in health care,” *Review of Economics and Statistics*, 102, 506–517.
- BLOOM, N., C. PROPPER, S. SEILER, AND J. VAN REENEN (2015b): “The impact of competition on management quality: evidence from public hospitals,” *Review of Economic Studies*, 82, 457–489.
- BLOOM, N., R. SADUN, AND J. VAN REENEN (2012): “Americans do IT better: US multinationals and the productivity miracle,” *American Economic Review*, 102, 167–201.

- (2015c): “Do private equity owned firms have better management practices?” *American Economic Review*, 105, 442–46.
- BLOOM, N. AND J. VAN REENEN (2007): “Measuring and explaining management practices across firms and countries,” *Quarterly Journal of Economics*, 122, 1351–1408.
- BUFFINGTON, C., L. FOSTER, R. JARMIN, AND S. OHLMACHER (2017): “The management and organizational practices survey (MOPS): An overview,” *Journal of Economic and Social Measurement*, 42, 1–26.
- CARRERI, M. (2021): “Can good politicians compensate for bad institutions? Evidence from an original survey of Italian mayors,” *The Journal of Politics*, 83, 1229–1245.
- CARRERI, M. AND J. PAYSON (2021): “What makes a good local leader? Evidence from US mayors and city managers,” *Journal of Political Institutions and Political Economy*, 2, 199–225.
- (2023): “Local leaders and the pursuit of growth in US cities: The role of managerial skill,” *Political Science Research and Methods*, forthcoming.
- DECAROLIS, F., L. M. GIUFFRIDA, E. IOSSA, V. MOLLISI, AND G. SPAGNOLO (2020): “Bureaucratic competence and procurement outcomes,” *Journal of Law, Economics, and Organization*, 36, 537–597.
- DESTATIS (2021): *GV-ISys: Verzeichnis der Regional- und Gebietseinheiten Definitionen und Beschreibungen*, Available at: https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/beschreibung-gebietseinheiten.pdf?__blob=publicationFile.
- DUFLO, E., R. HANNA, AND S. RYAN (2012): “Incentives work: Getting teachers to come to school,” *American Economic Review*, 102, 1241–78.
- DWENGER, N. AND A. GUMPERT (2023): “Building fiscal capacity,” *mimeo*, University of Tübingen.
- ENGLMAIER, F., J. GALDON-SANCHEZ, R. GIL, M. KAISER, AND H. STRANDT (2020): “Management practices and firm performance during the Great Recession: Evidence from Spanish survey data,” *Mimeo*, LMU Munich.
- FENIZIA, A. (2022): “Managers and productivity in the public sector,” *Econometrica*, 90, 1063–1084.
- HANSEN, S., M. MCMAHON, AND A. PRAT (2018): “Transparency and deliberation within the FOMC: A computational linguistics approach,” *Quarterly Journal of Economics*, 133, 801–870.
- HJORT, J., D. MOREIRA, G. RAO, AND J. F. SANTINI (2021): “How research affects policy: Experimental evidence from 2,150 Brazilian municipalities,” *American Economic Review*, 111, 1442–80.
- ICHNIOWSKI, C., K. SHAW, AND G. PRENUSHI (1997): “The effects of human resource management practices on productivity,” *American Economic Review*, 87, 291–313.

- LEAVER, C., O. OZIER, P. SERNEELS, AND A. ZEITLIN (2021): "Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary schools," *American Economic Review*, 111, 2213–46.
- LEMOS, R., K. MURALIDHARAN, AND D. SCUR (2021): "Personnel management and school productivity: Evidence from India," *NBER Working Paper No. 28336*.
- MCCORMACK, J., C. PROPPER, AND S. SMITH (2014): "Herding cats? Management and university performance," *Economic Journal*, 124, F534–F564.
- PELTZMAN, S. (1976): "Toward a more general theory of regulation," *Journal of Law and Economics*, 19, 211–240.
- RASUL, I. AND D. ROGGER (2018): "Management of bureaucrats and public service delivery: Evidence from the nigerian civil service," *Economic Journal*, 128, 413–446.
- RASUL, I., D. ROGGER, AND M. J. WILLIAMS (2021): "Management, organizational performance, and task clarity: Evidence from ghana's civil service," *Journal of Public Administration Research and Theory*, 31, 259–277.
- RUGE, K. AND K. RITGEN (2021): "Local self-government and administration," in *Public Administration in Germany*, Palgrave Macmillan, Cham, 123–141.
- SCHRAPPER, L. (2021): "The Administration of the Länder," in *Public Administration in Germany*, Palgrave Macmillan, Cham, 105–121.
- SCUR, D., R. SADUN, J. VAN REENEN, R. LEMOS, AND N. BLOOM (2021): "The World Management Survey at 18: Lessons and the way forward," *Oxford Review of Economic Policy*, 37, 231–258.
- STIGLER, G. (1971): "The theory of economic regulation," *Bell Journal of Economics and Management Science*, 3–21.
- SYVERSON, C. (2011): "What determines productivity?" *Journal of Economic Literature*, 49, 326–65.
- TIROLE, J. (2001): "Corporate Governance," *Econometrica*, 69, 1–35.
- TULLOCK, G. (1967): "The welfare costs of tariffs, monopolies, and theft," *Economic Inquiry*, 5, 224–232.
- WEBER, M. (1922): *Economy and society*, Mohr Siebeck, Tübingen.

Appendix

A Questionnaire sent to municipalities

In this Appendix we provide a translation of the survey questions (originally in German). In the list below, each question (Q 1 through Q 49) is accompanied by a comment field in square bracket, which is structured as follows:

1. “Type and result” identifies the question type by spelling out the possible answer options (separated by “|”), followed by the respective response frequencies (in percent) in the order of the answer options. Some questions are conditional in the sense that they are only reached if some prior question is answered with “yes”, indicated in parentheses (see e.g., Q 2).
2. “Subscore” indicates whether the respective question is included in the management score, which is (i) not the case if the entry is “n.a.”, and (ii) otherwise it is included in the management score (and in the stated subscore). As the score aims to measure the degree of structured management, “worst outcome” indicates the answer option receiving a score of zero. Each subscore is based on the following questions: goal setting (Q1-Q8), monitoring of service quality (Q9-Q16), organizational learning (Q17-Q22), structured processes (Q24-Q27), human resource management (Q23, Q30(i)-Q30(iii), Q31-Q41, Q44), and digitization (Q45(i)-Q45(v) and Q46(i)-Q46(v)).
3. The questionnaire contains various question types:
 - (a) “yes | no”, “Likert” (5-point scale), and “Interval in days” indicates that participants can select one answer option only (and if the percentage results do not add up to 100 this implies that some participants did not answer this question, see e.g., Q 3).
 - (b) “multiple options” implies that participants are free to select any number of the available answer options (and as a consequence the percentage results generally do not add up to 100, see e.g., Q 4).
 - (c) “number” allows participants to enter an integer.
 - (d) “free text” solicits a free-form input from participants.
 - (e) Questions Q 30, Q 45, Q 46 contain several subquestions, which each individually enter the management score. Therefore, the comment field provides the results separately for each subquestion.

List of questions

Q 1: Does your municipality have a mission statement that is directed at citizens? [Type and result: *yes | no = 26 | 74. Subscore: goal setting (worst outcome = no)*]

Q 2: Is the mission statement that is directed at citizens easily accessible for them (e.g., on the municipal webpage or newsletter)? [Type and result: *yes / no (conditional on Q1 = yes) = 84 / 16*. Subscore: *goal setting (worst outcome = no)*]

Q 3: Does the mission statement that is directed at citizens entail goals? [Type and result: *yes / no (conditional on Q1 = yes) = 86 / 12*. Subscore: *goal setting (worst outcome = no)*]

Q 4: What is the time horizon envisioned for these goals? [Type and result: *multiple options (conditional on Q1 = yes and Q3 = yes): one year / two to four years / five to ten years / more than 10 years = 7 / 21 / 62 / 31*. Subscore: *goal setting (worst outcome = no option chosen)*]

Q 5: Does your municipality have a mission statement that is specifically directed at the employees of the municipal administration? [Type and result: *yes / no = 21 / 78*. Subscore: *goal setting (worst outcome = no)*]

Q 6: Is the mission statement that is directed at employees easily accessible for them (e.g., on the municipal intranet)? [Type and result: *yes / no (conditional on Q5 = yes) = 88 / 11*. Subscore: *goal setting (worst outcome = no)*]

Q 7: Does the mission statement that is directed at employees entail goals? [Type and result: *yes / no (conditional on Q5 = yes) = 77 / 23*. Subscore: *goal setting (worst outcome = no)*]

Q 8: What is the time horizon envisioned for these goals? [Type and result: *multiple options (conditional on Q5 = yes and Q7 = yes): one year / two to four years / five to ten years / more than 10 years = 18 / 47 / 34 / 27*. Subscore: *goal setting (worst outcome = no option chosen)*]

Q 9: Does your municipality employ quantitative measures to evaluate administrative performance or service quality (e.g., with regard to processing times)? [Type and result: *yes / no = 16 / 83*. Subscore: *monitoring of service quality (worst outcome = no)*]

Q 10: Which of the following indicators are employed to measure administrative performance or service quality in your municipality? [Type and result: *multiple options (conditional on Q9 = yes): time to complete a process / waiting times of customers / frequency of complaints / other = 79 / 56 / 54 / 20*. Subscore: *monitoring of service quality (worst outcome = no option chosen)*]

Q 11: How are the indicators of administrative performance or service quality taken into account in municipal management? [Type and result: *multiple options (conditional on Q9 = yes): discussed at the employee level / at the management level / in the municipal council = 63 / 93 / 21*. Subscore: *monitoring of service quality (worst outcome = no option chosen)*]

Q 12: Overall, how strongly do such indicators affect the management of administrative performance or service quality in your municipality? [Type and result: *Likert (conditional on Q9 = yes): very large / large / medium / small / very small = 0 / 14 / 53 / 29 / 3*. Subscore: *monitoring of service quality (worst outcome = very small)*]

Q 13: Does your municipality conduct surveys among citizens and other customers to measure administrative performance or service quality? [Type and result: *yes / no = 16 / 84*. Subscore: *monitoring of service quality (worst outcome = no)*]

Q 14: How often does your municipality conduct surveys among citizens and other customers to measure administrative performance or service quality? [Type and result: *Likert (conditional on Q13 = yes): every year | every 2 years | every 5 years | every 10 years | less than every 10 years = 14 | 26 | 44 | 8 | 8*. Subscore: *monitoring of service quality (worst outcome = less than every 10 years)*]

Q 15: How are the results of surveys among citizens and other customers on performance or service quality taken into account in municipal management? [Type and result: *multiple options (conditional on Q13 = yes): discussed at the employee level | at the management level | in the municipal council = 62 | 87 | 25*. Subscore: *monitoring of service quality (worst outcome = no option chosen)*]

Q 16: Overall, how strongly do surveys among citizens and other customers affect the management of administrative performance or service quality in your municipality? [Type and result: *Likert (conditional on Q13 = yes): very large | large | medium | small | very small = 2 | 30 | 40 | 26 | 2*. Subscore: *monitoring of service quality (worst outcome = very small)*]

Q 17: In your municipality, is there a institutionalized way for employees to suggest improvements? [Type and result: *yes | no = 73 | 26*. Subscore: *organizational learning (worst outcome = no)*]

Q 18: Is there a reward system (financial or otherwise) for employees' suggestions for improvement in your municipal administration? [Type and result: *yes | no = 26 | 74*. Subscore: *organizational learning (worst outcome = no)*]

Q 19: How strongly do employees' suggestions for improvement affect municipal management? [Type and result: *Likert: very large | large | medium | small | very small = 3 | 26 | 34 | 25 | 12*. Subscore: *organizational learning (worst outcome = very small)*]

Q 20: Do employees who submit a suggestion for improvement usually receive feedback even if their proposal is ultimately not implemented? [Type and result: *yes | no = 91 | 8*. Subscore: *organizational learning (worst outcome = no)*]

Q 21: Is there a regular exchange between different departments of your municipal administration with the aim of improving processes or services? [Type and result: *yes | no = 72 | 27*. Subscore: *organizational learning (worst outcome = no)*]

Q 22: Does your municipality regularly exchange ideas with other municipalities on how to improve processes or services? [Type and result: *yes | no = 57 | 43*. Subscore: *organizational learning (worst outcome = no)*]

Q 23: What would happen in the case of staff shortages in key departments of your municipal administration? [Type and result: *multiple options: overtime | temporary relocation of idle capacity | hiring = 91 | 60 | 40*. Subscore: *human resource management (worst outcome = no option chosen)*]

Q 24: Does your municipal administration employ the instrument of "project management"? [Type and result: *yes | no = 49 | 51*. Subscore: *structured processes (worst outcome = no)*]

Q 25: How large is the value added of “project management” for your municipality? [Type and result: *Likert (conditional on Q24 = yes): very large | large | medium | small | very small = 7 | 51 | 34 | 7 | 0.* Subscore: *structured processes (worst outcome = very small)*]

Q 26: What role does the instrument of “process management” play in your municipal administration (e.g., the structured organization of processes in different departments)? [Type and result: *Likert: not used | few processes defined | core processes defined | processes defined in many areas | processes defined in all areas = 30 | 37 | 24 | 8 | 1.* Subscore: *structured processes (worst outcome = not used)*]

Q 27: Does your municipality use the instrument of a “product-oriented budgeting”? [Type and result: *yes | no = 79 | 20.* Subscore: *structured processes (worst outcome = no)*]

Q 28: To what extent do you agree with the following statement: The municipal administration and the municipal council work together in a constructive fashion? [Type and result: *Likert: very strongly | strongly | medium | weakly | very weakly = 19 | 39 | 30 | 9 | 3.* Subscore: *n.a. (worst outcome = n.a.)*]

Q 29: How many people are employed in your municipal administration (in full-time equivalents)? [Type and result: *number = 357 (mean), 80 (median).* Subscore: *n.a. (worst outcome = n.a.)*]

Q 30: How frequently does your municipal administration use the following instruments to remunerate its employees? [Type and result: *Likert: very often | often | sometimes | rarely | never. Three separate questions: (i) compensation above collective agreement, (ii) other upwards deviations from collective agreement, (iii) fast-track promotions) = (i) 1 | 5 | 19 | 32 | 39 , (ii) 0 | 2 | 12 | 28 | 51, (iii) 0 | 10 | 40 | 32 | 16.* Subscore: *human resource management (worst outcome = never)*]

Q 31: Does your municipal administration employ the instrument of “performance-based remuneration”? [Type and result: *yes | no = 78 | 22.* Subscore: *human resource management (worst outcome = no)*]

Q 32: To what extent do you agree with the following statement: In the municipal administration, the award of “performance-based remuneration” does indeed reflect actual performance differences? [Type and result: *Likert (conditional on Q31=yes): very strongly | strongly | medium | weakly | very weakly = 5 | 22 | 28 | 24 | 21.* Subscore: *human resource management (worst outcome = very weakly)*]

Q 33: Apart from “performance-based remuneration”, what other instruments does your municipal administration use to reward high performance of employees? [Type and result: *multiple options: promotion | nonmonetary awards | rewards (e.g., vouchers) | work cell phone or tablet | training opportunities | other = 42 | 2 | 13 | 28 | 54 | 28.* Subscore: *human resource management (worst outcome = no option chosen)*]

Q 34: For promotion decisions in your municipal administration, how big a role does the seniority of eligible candidates play in practice? [Type and result: *Likert: very large | large | medium | small | very small = 0 | 8 | 31 | 36 | 25.* Subscore: *human resource management (worst outcome = very small)*]

Q 35: Before the onset of the COVID19 pandemic, how frequently did your municipal administration's employees "work from home"? [Type and result: *Likert: very often | often | sometimes | rarely | never = 2 | 6 | 20 | 54 | 18*. Subscore: *human resource management (worst outcome = never)*]

Q 36: Before the onset of the COVID19 pandemic, what challenges did your municipal administration encounter when aiming to implement "work from home"? [Type and result: *multiple options: insufficient technical equipment | lack of IT skills | data security concerns | hampered communication | lack of monitoring | unequal treatment of employees = 64 | 14 | 51 | 44 | 29 | 44*. Subscore: *human resource management (worst outcome = all options chosen)*]

Q 37: Before the onset of the COVID19 pandemic, which of the following events were usually organized in your municipal administration? [Type and result: *multiple options: christmas party | office party | honoring jubilees | other = 88 | 76 | 85 | 44*. Subscore: *human resource management (worst outcome = no option chosen)*]

Q 38: In your municipal administration, is there a collective agreement in place concerning individual "review meetings" with employees? [Type and result: *yes | no = 57 | 42*. Subscore: *human resource management (worst outcome = no)*]

Q 39: How frequently do individual "review meetings" with employees take place? [Type and result: *Likert: more than once per year | once per year | regularly but not every year | only when need arises | never = 11 | 58 | 10 | 18 | 3*. Subscore: *human resource management (worst outcome = never)*]

Q 40: In the individual "review meetings", do employees typically receive feedback on their performance? [Type and result: *yes | no (conditional on Q39 = yes) = 79 | 20*. Subscore: *human resource management (worst outcome = no)*]

Q 41: What would typically happen if employees in your municipal administration underperform over a prolonged period of time? [Type and result: *multiple options: individual review meeting | training | re-assignment | deadline for improvement | termination during probation period | other = 97 | 65 | 31 | 41 | 76 | 19*. Subscore: *human resource management (worst outcome = no option chosen)*]

Q 42: How many sick days do employees of the municipal administration approximately take (per employee and year, on average over the last few years)? [Type and result: *interval in days: 0 - 5 | 6 - 10 | 11 - 15 | 16 - 20 | 21 - 25 | 26 - 30 | 31 or more = 8 | 38 | 28 | 11 | 4 | 2 | 1*. Subscore: *n.a. (worst outcome = n.a.)*]

Q 43: When considering the annual average over the last few years: How many employees of the municipal administration have been on long-term sick leave such that a formal process of "re-integration" was triggered? [Type and result: *number = 17 (mean), 3 (median)*. Subscore: *n.a. (worst outcome = n.a.)*]

Q 44: In times of scarce talent it is increasingly important to attract qualified personnel. What measures does your municipal administration take to achieve this? [Type and result: *multiple options: compensation above collective agreement | flexible work schedule | work cell phone or tablet | providing accommodation | support in finding accommodation | participation in job fairs | internship program | other* = 15 | 80 | 41 | 2 | 29 | 29 | 31 | 39. Subscore: *human resource management* (worst outcome = *no option chosen*)]

Q 45: In your municipal administration, approximately what percentage of relevant data is stored in digital form in the following areas? [Type and result: *Likert: 0-20% | 21-40% | 41-60% | 61-80% | 81-100%*. Five separate questions: (i) *personnel data*, (ii) *financial data*, (iii) *customer feedback*, (iv) *procurement processes*, (v) *processing times and status of processes* = (i) 14 | 11 | 15 | 21 | 32, (ii) 2 | 4 | 8 | 21 | 57, (iii) 32 | 14 | 19 | 14 | 12, (iv) 9 | 16 | 23 | 22 | 23, (v) 43 | 17 | 14 | 10 | 8. Subscore: *digitization* (worst outcome = 0-20%)]

Q 46: How often does your municipal administration use data-based forecasting (e.g., statistical techniques that allow to make predictions) in the following areas? [Type and result: *Likert: very often | often | sometimes | rarely | never*. Five separate questions: (i) *customer services*, (ii) *financial planning*, (iii) *infrastructure planning*, (iv) *personnel planning*, (v) *childcare services* = (i) 2 | 8 | 20 | 32 | 31, (ii) 18 | 40 | 19 | 8 | 7, (iii) 6 | 26 | 29 | 20 | 11, (iv) 5 | 26 | 26 | 21 | 15, (v) 12 | 27 | 21 | 17 | 15. Subscore: *digitization* (worst outcome = *never*)]

Q 47: In your opinion, what are the biggest challenges for successful municipal management in the medium to long term? [Type and result: *free text*. Subscore: *n.a.* (worst outcome = *n.a.*)]

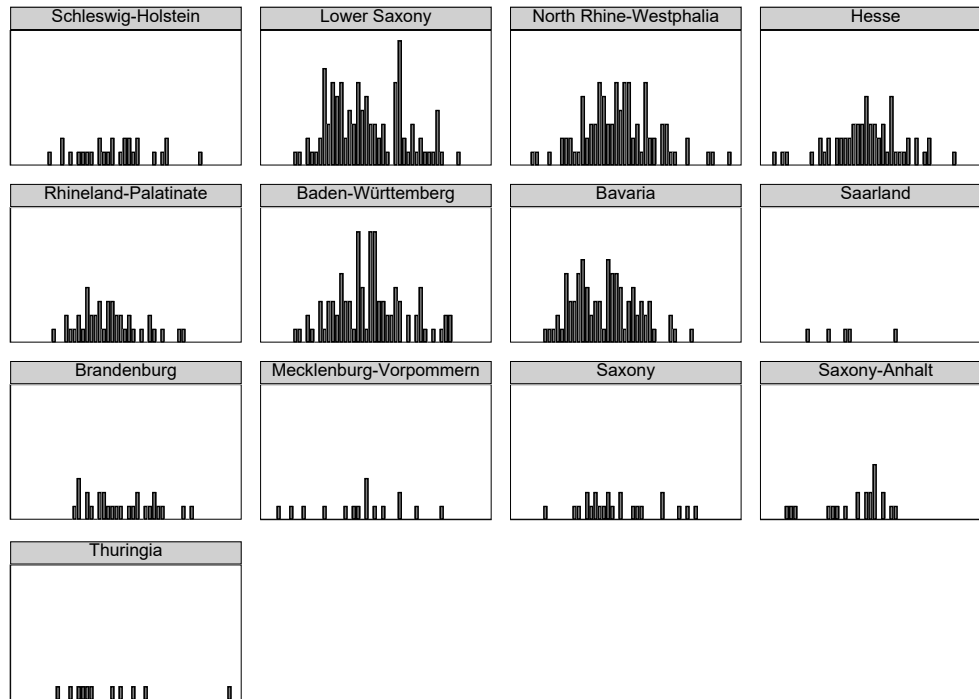
Q 48: What external conditions would have to change for municipal management to be better able to respond to pertinent challenges? [Type and result: *free text*. Subscore: *n.a.* (worst outcome = *n.a.*)]

Q 49: In your opinion, what hampers an effective and timely expansion of e-government services? [Type and result: *free text*. Subscore: *n.a.* (worst outcome = *n.a.*)]

B Additional tables and figures

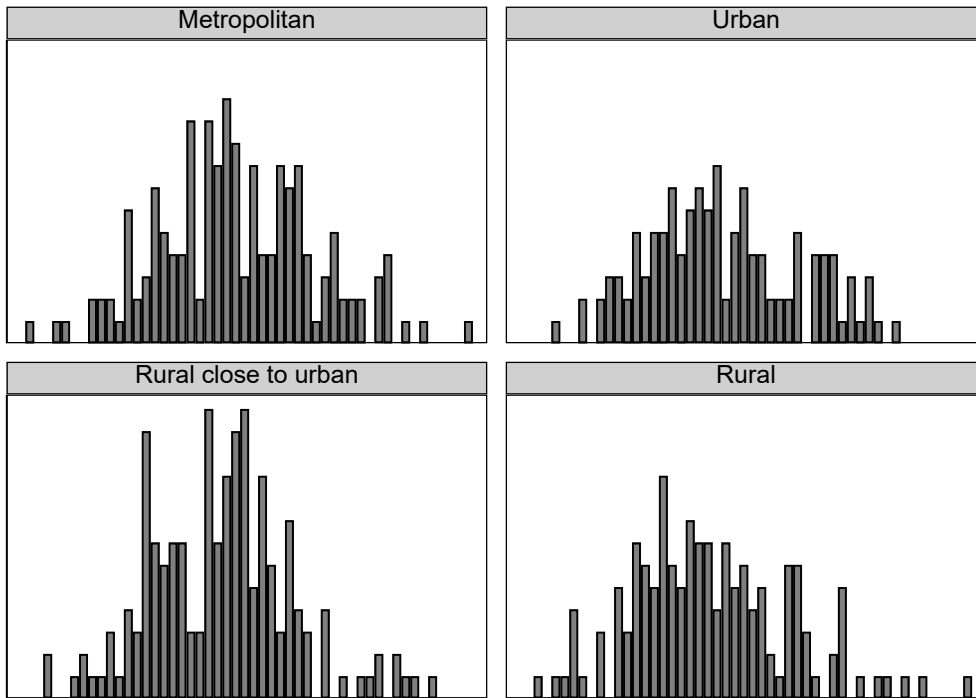
B.1 Heterogeneity of the management score in subsamples

Figure 5: Frequency distribution of the management score: By state



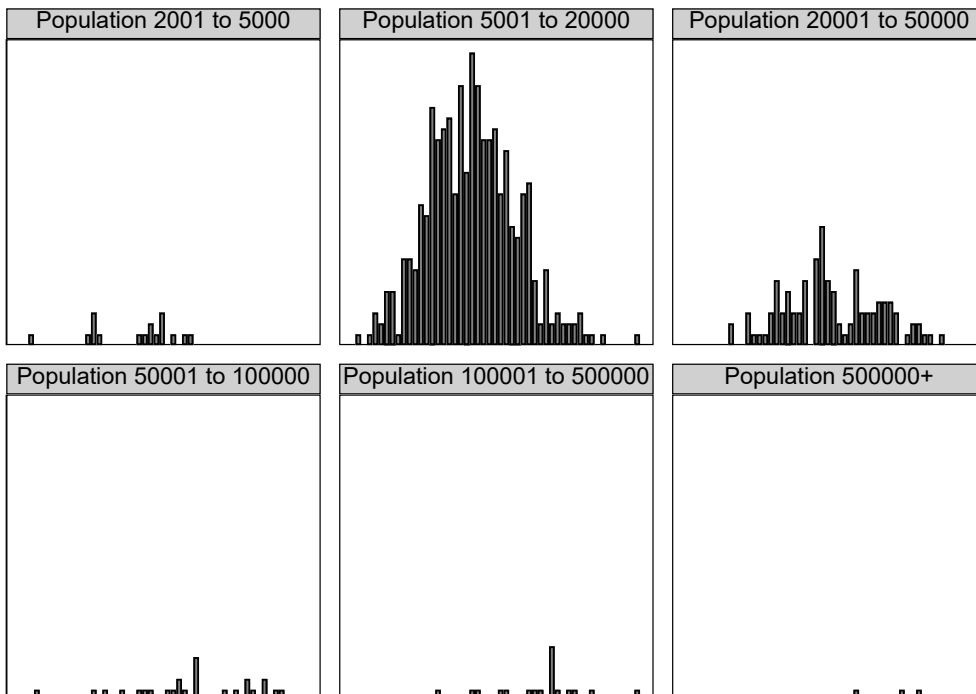
Notes: municipalities located in 13 out of the 16 German states participated in the survey. The three city states Berlin, Bremen, and Hamburg (each constituting one single municipality) did not participate.

Figure 6: Frequency distribution of the management score: By regional type



Notes: As discussed in Section 4, the four regional types were constructed from the *RegioStar* classification scheme of the German Statistical Office (see Destatis, 2021, p.16ff) .

Figure 7: Frequency distribution of the management score: By population bracket



Notes: The six population brackets follow the GKBIK7 classification of the German Statistical Office (see Destatis, 2021, p.44). As discussed in Section 4, the omitted seventh GKBIK7 bracket with population sizes smaller than 2,000 is not relevant for our analysis.

B.2 Regional types and population brackets: Distributions

Table 7: Regional and population types: Descriptive statistics for participating and non-participating municipalities

	Non-Participants	Participants	Total
Regional type			
Metropolitan	624 (27.05%)	162 (27.00%)	786 (27.04%)
Urban	465 (20.16%)	124 (20.67%)	589 (20.26%)
Rural close to urban	619 (26.83%)	171 (28.50%)	790 (27.18%)
Rural	599 (25.96%)	143 (32%)	742 (25.52%)
Total	2,307 (100.00%)	600 (100.00%)	2,907 (100.00%)
Population bracket			
2,001 to 5,000	40 (1.62%)	17 (2.83%)	57 (1.85%)
5,001 to 20,000	1844 (74.47%)	426 (71.00%)	2,270 (73.80%)
20,001 to 50,000	447 (18.05%)	111 (18.50%)	558 (18.14%)
50,001 to 100,000	85 (3.43%)	25 (4.17%)	110 (3.58%)
100,001 to 500,000	49 (1.98%)	18 (3.00%)	67 (2.18%)
more than 500,000	11 (.44%)	3 (.50%)	14 (.46%)
Total	2,476 (100.00%)	600 (100.00%)	3,076 (100.00%)

Notes: The table displays the distributions of participating and non-participating municipalities across regional types and population size brackets. The population size bracket to which a given municipality is assigned is not time-invariant across our sample period 2015-2019 for only seven out of the 600 participating municipalities. To these seven municipalities, we assign the 2015 population bracket.

B.3 Robustness: Management and performance

Table 8: Robustness: Management score and municipal performance (3-year- averages: 2017-2019)

	(1)	(2)	(3)	(4)
Net migration	5.37** (0.02)	5.58*** (0.00)	5.12*** (0.00)	4.56** (0.02)
Academic workforce	12.10*** (0.00)	12.00*** (0.00)	10.42*** (0.00)	4.61 (0.18)
Start-ups	1.61*** (0.00)	1.22*** (0.00)	0.94** (0.01)	0.49 (0.26)
Broadband coverage	27.41*** (0.00)	28.27*** (0.00)	25.31*** (0.00)	11.34 (0.16)
Business tax base	128.19*** (0.00)	150.05** (0.01)	146.59** (0.01)	125.88 (0.13)
Income tax revenue	70.36 (0.17)	80.11** (0.01)	36.19 (0.19)	57.25*** (0.01)
Primary balance	32.98 (0.54)	97.44 (0.34)	98.77 (0.32)	87.42 (0.44)
Financial balance	30.53*** (0.00)	28.30*** (0.00)	28.65*** (0.00)	12.78 (0.25)
Municipal performance index (MPI)	1.90*** (0.00)	1.97*** (0.00)	1.75*** (0.00)	1.10** (0.04)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The note below Table 3 applies with the only difference that three-year averages (2017-2019) of the indicators of municipal performance are considered.

Table 9: Robustness: Management score and municipal performance (2019 only)

	(1)	(2)	(3)	(4)
Net migration	3.49 (0.12)	3.78 (0.11)	3.21 (0.15)	2.68 (0.28)
Academic workforce	12.75*** (0.00)	12.65*** (0.00)	11.07*** (0.00)	4.87 (0.23)
Start-ups	1.85*** (0.00)	1.05 (0.12)	0.81 (0.24)	0.41 (0.70)
Broadband coverage	22.44*** (0.00)	23.18*** (0.00)	20.60*** (0.00)	10.07 (0.13)
Business tax base	122.65*** (0.00)	155.93** (0.03)	151.11** (0.04)	139.09* (0.10)
Income tax revenue	70.09 (0.18)	79.30** (0.02)	35.42 (0.20)	61.67** (0.02)
Primary balance	85.52 (0.48)	149.04 (0.32)	136.91 (0.34)	130.96 (0.46)
Financial balance	28.98*** (0.00)	25.66** (0.02)	25.87*** (0.01)	10.07 (0.46)
Municipal performance index (MPI)	1.78*** (0.00)	1.83*** (0.00)	1.59*** (0.00)	1.01* (0.08)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The note below Table 3 applies with the only difference that the most recent (i.e. 2019) values of the indicators of municipal performance are considered.

Table 10: Robustness: z-score and municipal performance

	(1)	(2)	(3)	(4)
Net migration	0.87*** (0.00)	0.84*** (0.00)	0.78*** (0.00)	0.69*** (0.00)
Academic workforce	1.35*** (0.00)	1.32*** (0.00)	1.15*** (0.00)	0.48 (0.20)
Start-ups	0.21*** (0.00)	0.16*** (0.01)	0.12*** (0.01)	0.07 (0.17)
Broadband coverage	3.08*** (0.00)	3.06*** (0.00)	2.72*** (0.00)	1.21 (0.13)
Business tax base	15.53*** (0.00)	17.34*** (0.00)	16.83** (0.01)	14.29 (0.11)
Income tax revenue	10.61** (0.05)	10.12*** (0.00)	5.06 (0.10)	6.83** (0.01)
Primary balance	11.46 (0.14)	16.19* (0.09)	16.03* (0.09)	16.47 (0.15)
Financial balance	4.65** (0.03)	4.54** (0.04)	4.52*** (0.00)	2.79 (0.12)
Municipal performance index (MPI)	0.25*** (0.00)	0.25*** (0.00)	0.22*** (0.00)	0.15** (0.04)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The note below Table 3 applies with the only difference that the z-score is used as the measure of structured management (instead of the management score). The construction of the z-score follows Bloom et al. (2019): The answers to each of the 52 questions entering the score are normalized to have a mean of 0 and a standard deviation of 1, and then averages are taken over all questions.

Table 11: Robustness: Management style and municipal performance

	(1)	(2)	(3)	(4)
Net migration	1.83** (0.03)	1.82*** (0.00)	1.58*** (0.01)	1.19* (0.06)
Academic workforce	4.31*** (0.00)	4.42*** (0.00)	3.73*** (0.00)	1.21 (0.15)
Start-ups	0.56*** (0.00)	0.49*** (0.00)	0.37* (0.05)	0.19 (0.31)
Broadband coverage	9.35*** (0.00)	9.41*** (0.00)	8.27*** (0.00)	2.51 (0.20)
Business tax base	45.83** (0.02)	51.95** (0.01)	50.02** (0.02)	41.15 (0.19)
Income tax revenue	13.94 (0.25)	24.66** (0.02)	4.90 (0.50)	10.18 (0.19)
Primary balance	33.53 (0.24)	55.30 (0.11)	55.25 (0.12)	59.31 (0.19)
Financial balance	14.76*** (0.00)	13.76*** (0.00)	14.06*** (0.00)	7.63 (0.26)
Municipal performance index (MPI)	0.71*** (0.00)	0.75*** (0.00)	0.65*** (0.00)	0.38* (0.08)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The note below Table 3 applies with the only difference that the probability for the Type 1 management style is used as the measure of structured management instead of the management score (see Section 5).

Table 12: Robustness: Management score and municipal performance (controlling for the number of employees (per capita) in the municipal administration)

	(1)	(2)	(3)	(4)
Net migration	6.59* (0.07)	5.98*** (0.00)	5.48*** (0.00)	4.86** (0.01)
Academic workforce	8.16* (0.08)	8.36* (0.08)	7.35* (0.08)	4.01 (0.25)
Start-ups	1.05** (0.02)	1.06** (0.02)	0.84** (0.02)	0.48 (0.26)
Broadband coverage	20.46*** (0.00)	22.91** (0.02)	20.75** (0.03)	13.28 (0.10)
Business tax base	107.97 (0.10)	128.28 (0.11)	125.99 (0.10)	117.12 (0.17)
Income tax revenue	36.48 (0.24)	73.45*** (0.01)	40.66 (0.14)	49.56** (0.03)
Primary balance	73.96* (0.07)	154.47* (0.10)	153.80* (0.06)	154.11 (0.17)
Financial balance	35.41* (0.08)	35.98* (0.05)	35.18** (0.05)	23.15 (0.14)
Municipal performance index (MPI)	1.64** (0.05)	1.87** (0.03)	1.69* (0.05)	1.26* (0.07)
State dummies	No	Yes	Yes	Yes
Regional type dummies	No	No	Yes	Yes
Population bracket dummies	No	No	No	Yes

Notes: The note below Table 3 applies with the only difference that the number of employees (per capita) in the municipal administration is included as an additional explanatory variable in all regressions.