Check for updates

RESEARCH ARTICLE

How does delegation affect the administrative capacity of the hollow state?

[version 1; peer review: 2 approved with reservations]

Thomas Balbach

Public Governance Institute, KU Leuven, Belgium

v1 **First published:** 04 Jul 2026, **6**:230
https://doi.org/10.12688/openreseurope.23326.1
Latest published: 04 Jul 2026, **6**:230
https://doi.org/10.12688/openreseurope.23326.1

Abstract

A hollow state results from extensive delegation via contracting out and decentralization. How does this affect its administrative capacity to steer contractors toward collective goals such as environmental sustainability? This study argues that delegation stymies state-level administrative capacity by diminishing 'demand power', i.e., the state's uniform and comprehensive power to formulate public demand. Contracting out fragments demand based on transaction characteristics, while decentralization fragments demand by buyer organization. This is tested by analyzing how transaction characteristics and political control over buyer organizations affect contracting for sustainability. Using data on all 566,442 public procurement procedures in Germany from 2021-2023, a logistic generalized linear mixed model finds that the specificity of assets needed to produce a good, service or work is the single strongest barrier to including sustainability criteria in a contract. While buyers with more political oversight are more likely to contract for sustainability, this effect is secondary. This is the first study to assess an entire hollow state, finding that demand fragmentation by product type is a more critical impediment to demand power than buyer fragmentation.

Keywords

hollow state, state of agents, state capacity, administrative capacity, demand power, transaction cost, agency drift, political steering, sustainable public procurement, green public procurement

Open Peer Review

Approval Status ? ?

	1	2
version 1 04 Jul 2026	? view	? view

1. **Ekaterina Domorenok**, University of Padua, Padua, Italy

2. **Karem Aboelazm**, United Arab Emirates University, Al Ain, United Arab Emirates

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Horizon Europe gateway.



This article is included in the Political Science gateway.

Corresponding author: Thomas Balbach (thomas.balbach@kuleuven.be)

Author roles: Balbach T: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the Horizon Europe Framework Programme under grant agreement No [101059288]
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Balbach T. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Balbach T. **How does delegation affect the administrative capacity of the hollow state? [version 1; peer review: 2 approved with reservations]** Open Research Europe 2026, **6**:230 https://doi.org/10.12688/openreseurope.23326.1

First published: 04 Jul 2026, **6**:230 https://doi.org/10.12688/openreseurope.23326.1

Introduction

Governments today often do not deliver public goods, services or works themselves. Instead, they delegate to a vast array of private firms, foundations, corporatized entities, and agencies, creating what is known as the hollow state[i]. This raises a critical question: How does this widespread delegation affect the hollow state's capacity to steer contractors toward collective goals such as environmental sustainability? While delegation is not inherently problematic, this study examines how two common forms of delegation – contracting out[ii] and decentralization[iii] to less politically controlled public sector organizations – can impact state capacity (Berwick & Christia, 2018).

The modern state is increasingly characterized by delegation. Functions once performed by core government bureaucracies are now routinely contracted out to firms, devolved to lower levels of government, or assigned to agencies, foundations and corporatized entities. The hollow state is a complex web of third-party governance where the lines of accountability and control are often blurred (Heinrich et al., 2010; Milward, 1996; Salamon & Lund, 1989). For decades, the debate surrounding this transformation has oscillated between concerns for democratic accountability and questions of performance, moving from the classic principal-agent problem between politicians and bureaucrats (Bendor et al., 2001; Weber, 2010 [1921]) to the problem of bureaucrats' responsibility to the public interest (Friedrich, 1940).

Despite this extensive body of research, a critical question remains largely unanswered: How does delegation affect the administrative capacity of the hollow state? While many studies have examined how delegated organizations deliver on collective goals, the aggregate impact of widespread delegation on *state-level* administrative capacity is poorly understood (e.g., Cheng et al., 2021; Raeymaeckers & Cools, 2021). Similarly, research on public procurement has identified many individual-level and organizational-level correlates of contracting for collective goals, such as awareness, prior experience, discretion, and trainings (Dimand & Neshkova, 2024; Sönnichsen & Clement, 2020; Testa et al., 2012; Zhu et al., 2013). Yet it is unclear how these correlates aggregate to the state level (Rosell, 2021; Shadrina et al., 2022) and impact the state-level administrative capacity to steer contractors.

This study addresses two overlapping research gaps. The first research gap regards the overall administrative capacity of hollow states. Most studies have focused on the performance of single organizations or on single forms of delegation, yielding mixed findings (Heinrich et al., 2010; Kelly LeRoux & Nathaniel S. Wright, 2021a, 2021b; Overman, 2016). Some studies suggest that contracting out can improve outcomes for simple, measurable services, while others find that contracting out complex services jeopardizes collective goals (Dubnick & Frederickson, 2010; Ford, 2021; Zhao & LeRoux, 2021). What is missing is the bigger picture that covers all public sector organizations and all goods, services, and works within a hollow state[iv]. The second research gap lies in identifying *which* forms of delegation, or combinations thereof, lead to lower administrative capacity. An untested proposition (Heinrich et al., 2010, p. i7) is that rather than longstanding forms of delegation, such as devolution and decentralization, it is newer forms of delegation such as contracting out that lead a diminished capacity to steer contractors toward collective goals.

This study refers to the overall administrative capacity of a state to steer firms toward collective goals as 'demand power'. Demand power is defined as the power to make public demands towards contractors with regards to how goods, services and works should be produced (compare Lundberg et al., 2016; Kapczynski & Michaels, 2024). Demand power can be enacted by introducing policy criteria in public procurement procedures, in grantmaking, or by having equity in contractors (Bozeman, 1987). Demand power is one dimension of state-level administrative capacity (cf. coordination capacity in Berwick & Christia, 2018, and administrative power in Kapczynski & Michaels, 2024). In turn, state-level administrative capacity is one aspect of state capacity, next to other aspects such as the capacity to wage war (Berwick & Christia, 2018; Williams, 2021). The concept of demand power highlights an understudied, yet frequently alluded-to dimension of administrative capacity in hollow states (Berwick & Christia, 2018; Williams, 2021). This study focuses on the state level while acknowledging that state-level capacities break down into heterogeneous organization-level capacities due to distinct performance cultures (Williams, 2021), relational contracting (Warsen et al., 2019), and local collaborative networks (Provan & Kenis, 2008) (see Future Research). Organization-level heterogeneity does not, however, imply that there are no general mechanisms that lead to state-level administrative capacity.

---

[i] In this article, the following terms are considered synonyms: 'hollow state', 'state of agents', 'third-party government'.

[ii] Contracting out is defined as the purchase of the delivery of public goods, services and works from firms, instead of producing them in-house (Overman, 2016).

[iii] Decentralization is defined as the delegation of authority from majoritarian organizations to non-majoritarian organizations. This includes phenomena such as agencification, corporatization and third-sector governance (see Appendix A in Balbach, 2026).

[iv] This research gap is surprising. After all, concepts such as the hollow state and the state of agents literally foreground the *state*, when compared to the related concept of governance (network governance, hybrid governance, collaborative governance) which is more suitably applied to single organizations or sets of organizations.

This study develops a theoretical mechanism that integrates transaction cost theory with agency drift theory. Demand power arises from the uniform and comprehensive pursuit of a collective goal across markets at the right level of stringency (Andrews, 1998; Halonen, 2021, p. 548; Lundberg & Marklund, 2013, pp. 14-15). Contracting out and decentralization diminish demand power by contributing respectively to the fragmentation of the demand side into thousands of transactions and thousands of buyers (compare Dubnick & Frederickson, 2010, p. i155; Langlois & Foss, 1999). The core argument is that extensive delegation systematically violates the necessary conditions for demand power to emerge, thereby stymieing one aspect of state-level administrative capacity.

These interrelated mechanisms are tested for the case of sustainable public procurement[v]. The scope of this case is limited in several ways. The hollow state's administrative capacity to steer contractors toward collective goals operates not only via public procurement and purchase contracts, but also via grantmaking and holding equity in contractors. Furthermore, sustainability is but one example of collective goals. Other examples of collective goals are public health, public safety, and justice. Collective goals have been variously characterized as goals that serve the public interest (Britton-Purdy et al., 2020), realize public values (K. LeRoux & N. S. Wright, 2021b; Brown et al., 2006), reflect basic human capabilities (Nussbaum, 2011) or ensure welfare (Hausman, 2024). For the purpose of this study, the relevant characteristic of collective goals is their breadth and transversality. Environmental sustainability is a collective goal because human societies need an environment that can sustain them, such as continued water supply, fertile soil, or clean air (Diamond, 2005). The public sector itself plays a major role in achieving environmental sustainability. The goods, services and works that governments procure create negative externalities, such as polluting the environment or exacerbating climate change. That is why governments seek to use government contracting as a tool to achieve environmental sustainability (Grandia & Meehan, 2017; Lerusse & Van de Walle, 2021; Swedish Competition Authority, 2012). The sheer size of public procurement makes it a powerful lever to achieve environmental sustainability. Public procurement is sized at about a third of government expenditure (Ortiz-Ospina & Roser, 2016), and roughly a tenth of the gross domestic product in low-, middle- and high-income countries (Bosio & Djankov, 2020; Organisation for Economic Co-operation and Development, 2021).

This study answers the following research question: When and how does delegation stymie the administrative capacity of the hollow state? Four pre-registered hypotheses (Balbach, 2025) express the theoretical framework. Building on transaction cost theory and agency drift theory, it posits that contracting for environmental sustainability is less likely in transactions with high asset specificity, elevated transaction frequency, and high measurability, and more prevalent in public organizations subject to greater political control. Demand-side fragmentation (Lindström et al., 2022; Lundberg et al., 2016; Saussier & Tirole, 2015, p. 7) is explained with the transaction properties of things procured (Williamson, 1985) and the political control over the buyer organization (Biela & Papadopoulos, 2014; Dahlström et al., 2021; Dahlström et al., 2011). Here, political control is viewed as a means of transmitting collective goals from elected officials to subordinate bureaucrats (Bogumil, 2007; Carpenter, 2001; Dahlström et al., 2011), while acknowledging that economists often view it as a source of harmful external uncertainty to a business relationship (Williamson, 1999).

This study uses a cross-sectional administrative data set covering all notified procurement procedures in Germany in 2021, 2022, and 2023, collected by the German Federal Statistical Office via continuous, survey-based reporting by public buyer organizations (N = 566,442). In a generalized linear mixed model, contracting for environmental sustainability is explained by recently vetted measures of transaction properties (Potoski et al., 2025) and a novel measure of political control, while controlling for competition and procedural discretion.

The findings show when and how delegation stymied demand power and the pursuit of environmental sustainability. Only 9.2% of contracts from 2021-2023 included environmental sustainability criteria, which were broadly defined. Contracting for environmental sustainability was mainly stymied by the specificity of the assets needed to produce goods, services and works. It is comparatively harder to apply a collective goal such as environmental sustainability to highly asset-specific production processes and to specify this within a purchase contract. It was also stymied by the practice of splitting contracts into multiple lots, which is a way to increase competition among bidders. Transaction frequency (framework agreements) and buyer discretion in picking the winning bid had the biggest positive association with contracting for environmental sustainability. As expected, the measurability of goods, services and works and political control of the buyer organization are associated with more contracting for environmental sustainability.

This study realizes the analytical promise of the hollow state literature to overcome the small-scale analytical focus of positive theories of delegation. Building on studies of single or multiple organizations affected by one or two forms of

[v]Contracting for environmental sustainability is often referred to as 'green public procurement' or 'sustainable public procurement'. In practice, this means writing down clauses and sentences in a procurement document that reflect broader environmental outcomes.

delegation (Dubnick & Frederickson, 2010; Heinrich et al., 2010; Zhao & LeRoux, 2021), this is the first study to analyze three forms of delegation across an entire hollow state (from higher to lower levels of government, from more to less control by elected officials, from public sector organizations to private firms). Thereby, it adds nuanced knowledge on the effect of delegation on demand power, which is an important dimension of administrative capacity in a hollow state. It is the fragmentation of demand into distinct goods, products, and services due to contracting out, rather than the fragmentation of the political control of buyers, that reduces demand power and stymies the pursuit of environmental sustainability in the hollow state.

The administrative capacity of the hollow state

Terms such as the hollow state, the state of agents, or third-party government, are part of a longstanding debate about the problem of delegation (Heinrich et al., 2010). The debate gradually moved down the chain of delegation, while oscillating between a focus on accountability or performance. It started with the classical idea of a clear split between elected officials who delegate to unelected bureaucrats (Weber, 2010 [1921]; Wilson, 1887). The main question was normative, namely, how to hold bureaucrats accountable to elected officials. The Friedrich-Finer debate opened up a new position, namely, that bureaucrats should be directly responsible towards the public interest (Friedrich, 1940). This position was ignored in formal positive theory, where delegation was reduced to games of control between a politician principal and a bureaucrat agent (Bendor et al., 2001; Britton-Purdy et al., 2020; Spence, 1997). By the 1990s, scholars studied a variety of delegation to agencies, corporations, third-sector organizations and firms (Levi-Faur, 2005; Milward, 1996; Overman, 2016; Salamon & Lund, 1989). This enabled scholars to view the state as a whole, which, due to multiple forms of delegation, had become a hollow state (Milward, 1996) or state of agents (Heinrich et al., 2010). Nevertheless, most studies examined how single organizations managed their public service delivery in hollowed out arrangements, rather than hollow states as a whole (e.g., Cheng et al., 2021; Raeymaeckers & Cools, 2021; Wright & Reames, 2021).

Empirical research offers mixed findings on how well the hollow state achieves collective goals such as sustainability, health or education. Dubnick and Frederickson (2010) find that combined devolution and contracting out did not contribute to selected organizations' achievement of collective goals, while they refrain from characterizing them as hindrances. They conclude that agents were held accountable for performance most effectively in "grant and contract management and oversight at the operational level" (p. i155). Ford (2021) finds that quasi-markets for charter schools do not sort out schools that deliver low performance, i.e., lower student outcomes on reading and math proficiency. Zhao and LeRoux (2021) find that contracting out simple and measurable public health services improves public health. Summing up, to date, hollow state scholarship has not yet examined the state-level administrative capacity to steer contractors toward collective goals (Bozeman, 1987; Berwick & Christia, 2018).

Next to academic debates, questions about the demand power of the hollow state frequently feature in public discourse, indicating that the stakes are high. Typically, the critics of contracting out do not overlap with the critics of delegation to less politically controlled entities. For example, some critics proffer that many non-government organizations receive public funding (decentralization type of delegation) but then pursue their own goals rather than collective goals (Gersdorf, 2025; Harris, 2024). Whereas other critics argue that governments contract out to firms which also pursue their own, rather than collective goals (Gammeltoft-Hansen, 2012). This study takes both sides of the public debate seriously. Societal and political cleavages about how the public sector should be organized may be explained by the hollow state's differential effect on specific collective goals. Generally, progress on collective goals does not proceed in lockstep. Progress toward the goal of environmental sustainability may be negatively related to progress on other collective goals (García-Rodríguez et al., 2025). The way that the public sector organizes to achieve collective goals may play an important role in this pattern. As part of a broader effort to bring evidence to this public debate, this study contributes evidence on which forms of delegation are problematic for collective sustainability goals. However, the scope of the present study is limited in two important ways. This study does not examine the cost-effectiveness of achieving collective goals (Lundberg & Marklund, 2013) nor does it estimate welfare, given that an effective pursuit the goal of sustainability may harm other goals, such as fair and increased competition (Halonen, 2021, p. 549; K. LeRoux & N. S. Wright, 2021; Saussier & Tirole, 2015, p. 7).

Theoretical framework

This study argues that the erosion of state-level demand power unfolds through three interrelated components (Figure 1). First, the voluntary and nonbinding nature of self-regulation for collective goals such as sustainability enables the fragmentation of the demand side (Andrews, 1998; Halonen, 2021, p. 548; Lindström et al., 2022). Second, the structural decentralization inherent in the hollow state creates buyer-side fragmentation. Drawing on agency drift theory (Heinrich et al., 2010), this study posits that public organizations with less direct political oversight and more formal autonomy are less inclined to contract for collective goals (cf. Dahlström et al., 2011; Windoffer, 2020). This variation in buyer inclination further diminishes the hollow state's demand power. Third, contracting out to private firms creates product

market fragmentation. Guided by transaction cost economizing principles and to avoid contracting hazards, public buyers rationally avoid including difficult-to-specify and hard-to-verify policy criteria (credence attributes) in high-stakes contracts (Balafoutas & Kerschbamer, 2020; Dulleck et al., 2011; Williamson, 1999). This leads to an inconsistent application of collective goals to contracts.

Delegation stymies demand power in the hollow state (compare 'administrative power' in Kapczynski & Michaels, 2024; cf. also Saussier & Tirole, 2015, p. 7). This is because demand power is given when three conditions are present (Lundberg et al., 2016). Collective goals are (1) uniformly and comprehensively applied across products (cf. also Saussier & Tirole, 2015, p. 7), (2) at the right level of stringency, so that (3) a growing share of products aligns with the collective goals. If firms encounter inconsistent policy criteria in procurement procedures of client public sector organizations, then there is less incentive to adapt to any of those criteria (Halonen, 2021, p. 548). Also, buyer organizations need to find the right level of stringency that would move their contractors jointly towards more environmentally friendly production technologies (Lundberg & Marklund, 2013, pp. 14-15).

In a hollow state, combinations of delegation forms, such as contracting out and decentralization, create demand-side fragmentation (Andrews, 1998; Halonen, 2021, p. 548), which violates the aforementioned conditions for demand power. In turn, reduced demand power leads to fewer contracts that pursue collective goals. For example, if a contract includes sustainability criteria, then they may or may not be implemented during contract execution. However, if a contract does not include sustainability criteria, it is quite certain that they will not be implemented. In that sense, demand power, as one dimension of state-level administrative capacity of a hollow state (Figure 1), determines the subsequent contract execution and eventual policy outcomes.

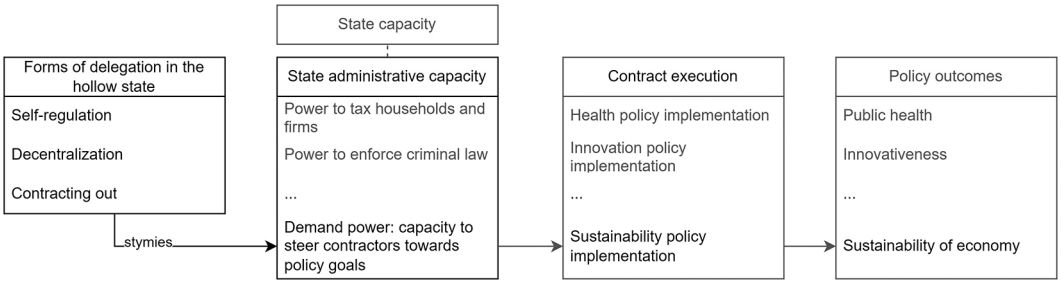

Figure 1. The effect of delegation on the state capacity of hollow states. The diagram illustrates how different forms of delegation in the hollow state (self-regulation, decentralization, contracting out) can stymie the development and exercise of state administrative capacity. State administrative capacity includes core coercive and extractive functions (e.g., taxation and law enforcement) and, critically, demand power (the capacity to steer contractors toward policy goals). Demand power, in turn, conditions contract execution (e.g., health, innovation, and sustainability policy implementation), which ultimately shapes downstream policy outcomes (e.g., public health, innovativeness, and sustainability of the economy). The dashed link indicates that state administrative capacity is a central component of broader state capacity.

Prior research called on scholars to conceptualize the exact mechanisms that affect state capacity as well as the mechanisms by which state capacity affects state performance (Williams, 2021). Three forms of delegation in a hollow state play a role in diminishing the administrative capacity of the hollow state, namely, self-regulation, decentralization, and contracting out (Figure 2). These forms of delegation affect the demand power of the hollow state via the general mechanisms of agency drift (Heinrich et al., 2010) and transaction cost economizing (Williamson, 1985, 1999). In other words, forms of delegation affect the extent to which collective goals are contracted for in public procurement procedures in a hollow state, i.e., the extent to which policy criteria are used in purchase contracts. This study focuses on policy criteria in purchase contracts, while there are other aspects of demand power, such as using policy criteria in grants and subsidies or holding equity in contractors (Bozeman, 1987).

In what follows, the mechanisms in Figure 2 are described in detail. Collective goals such as environmental sustainability, when pertaining to public procurement, are largely based on self-regulation, rather than command-and-control regulation. Buyer organizations draw on non-binding and voluntary self-regulation to describe sustainable qualities of goods, services, and works (Andrews, 1998; Halonen, 2021, p. 548; Lindström et al., 2022). A firm is entirely free to decide whether it will bid for a contract with environmental criteria. The impact on the environment that government contracting for environmental sustainability has, will crucially depend on the reaction of firms to competitive bidding for contracts with environmental criteria (Lundberg et al., 2016; Marron, 1997). In that sense, self-regulation enables the fragmentation of public demand by the following two forms of delegation.

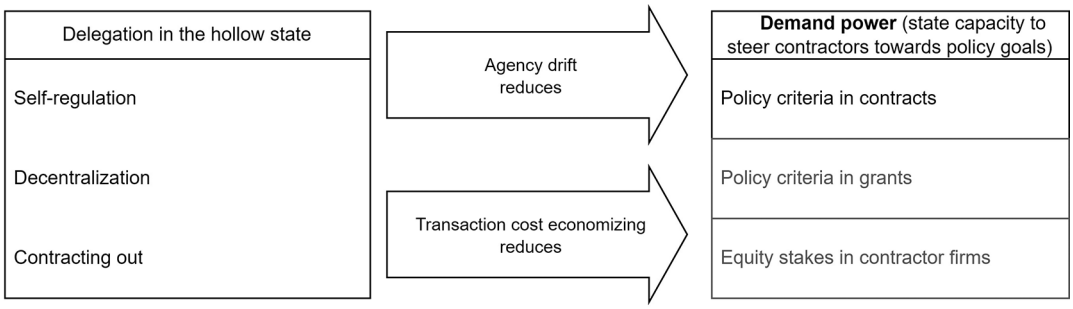

Figure 2. Mechanisms whereby delegation stymies demand power. The diagram theorizes how delegation in the hollow state - via self-regulation, decentralization, and contracting out - can weaken demand power, defined as the state's capacity to steer contractors toward collective goals. Two pathways are displayed: agency drift and transaction-cost economizing, each of which reduces the state's ability to embed collective goals in contracts, grants or other implementation instruments. The right-hand panel lists operational manifestations of demand power, including the inclusion of policy criteria in contracts (emphasized), as well as policy criteria in grants and equity stakes in contractor firms (shown as additional channels).

Structural decentralization to less politically controlled public sector organizations brings forth buyer-side fragmentation in hollow states (Bach & Jann, 2010) and comes with agency drift (Heinrich et al., 2010). Delegated public sector organizations are less likely to pursue policy goals in public procurement. This reduces the overall demand power of hollow states. The agency drift mechanism (Figure 2) is shaped by the political control structure of the public sector. More politically controlled buyer organizations are held accountable for collective goals by elected/appointed officials and are more likely to contract for them[vi].

Building upon the accountability regime framework by Biela and Papadopoulos (2014), the present study puts forward a political control framework (see operationalization in Appendix A in Balbach, 2026). It ranks political control regimes by the extent of the formal organizational autonomy of public buyer organizations and by the extent of their exposure to the influence of elected officials (cf. Dahlström et al., 2011; Windoffer, 2020). It captures public sector organizations' political control regime as their relative position in chains of delegation (cf. e.g., Bach et al., 2018; Bertelli, 2021; Maggetti & Verhoest, 2014, p. 249). More formal autonomy[vii] gives governments less opportunity to politically control the activities of administrative entities and coincides with fewer political appointments (Bach et al., 2018). More formal autonomy is linked to lower administrative accountability to elected government (Verschuere et al., 2006, pp. 6-7). Formal organizational autonomy is a proxy for de facto autonomy (Baekgaard et al., 2018). For example, bureaucrats at formally less autonomous organizations pay more attention to signals from government officials (Baekgaard et al., 2018).

Political control by elected government pertains to broad and transversal political preferences, not to detailed public policy. This fits with how elected government comes about. Citizens elect government based on their perceived representation of citizen preferences, not their views on specific public policies (Rubin, 2005). Elected officials' and unelected bureaucrats' care and concern about issues operate at different levels of granularity (Baekgaard et al., 2018; Schillemans & Busuioc, 2015). Carpenter (2001) argued that the relationship between elected government and bureaucracies cannot be conceived of as a contract with detailed prescriptions, but rather as rough guidelines (regarding Germany, cf. Bogumil, 2007). By contrast, bureaucrats are more committed to the granular public policies that are already in place (Wood, 1988, p. 231; cf. also Schillemans and Busuioc, 2015, p. 209).

Lastly, contracting for collective goals such as sustainability exacerbates the transaction cost of buying goods, services or works (Figure 2). Buyer organizations economize on transaction cost and the hazards of contracting for difficult-to-specify and hard-to-verify collective goals. Collective goals such as environmental sustainability are achieved during the production of a good, service or work. But to a buyer, this production process is unobservable or hard to observe. The buyer must believe that the production process contributed to the collective goal. For example, sitting on a chair, it is hard

[vi] A rival explanation of the role of political control on buyer organizations is proffered by transaction cost theory: elected officials interfere in the business relationship, making them a source of external uncertainty to the contract. Hence, the buyer organization will avoid contracting for collective goals, considering potential interference by politicians. The pre-registration of this study elaborates more on this rival hypothesis (Balbach, 2025).

[vii] The present study only examines formal autonomy, not perceived autonomy. It is unclear how tightly the two are linked, with large differences by country of study (van Thiel & Yesilkagit, 2014, p. 321).

to know how it was produced, and even harder to know how to describe the environmentally friendly aspects of its production. If an environmental sustainability certificate is attached to the chair, the buyer only knows that the company producing it has been certified as environmentally friendly (Balafoutas & Kerschbamer, 2020; Balineau & Dufeu, 2010). This buyer-side uncertainty of credence goods can refer to the difficulty of expressing what environmental sustainability means for a concrete good, work or service as well as the difficulty of verifying it once it has been delivered (compare Balineau & Dufeu, 2010; Dorée et al., 2011; Rietbergen et al., 2015).

Since public buyers have discretion in deciding when to contract for environmental sustainability, they are expected to choose which contracts they assign environmental criteria to. Based on transaction cost theory, this study expects public buyers to assign difficult-to-specify and hard-to-verify criteria to less hazardous transactions. For less hazardous transactions, liability and verifiability are more likely given. Liability and verifiability generally increase the trade of credence goods and the welfare that these transactions create (Dulleck et al., 2011). Low-hazard transactions feature low asset specificity, high measurability, low external uncertainty, and non-recurring purchases. Asset specificity is the ease by which an asset can be redeployed to a second-best use. Measurability is the ease by which the contract object can be specified and measured. External uncertainty is the extent to which future states of the world that affect the transaction are not known. Frequency is the number of times that a transaction will take place (Geyskens et al., 2006, p. 532).

The following pre-registered hypotheses (Balbach, 2025) apply the above theoretical mechanisms (Figure 2) to the pursuit of environmental sustainability in public procurement. Hypotheses 1 to 3 concern how contracting out stymies demand power, while Hypothesis 4 concerns how decentralization stymies demand power.

Hypothesis 1: Higher asset specificity is associated with less contracting for environmental sustainability.

Hypothesis 2: Higher measurability is associated with more contracting for environmental sustainability.

Hypothesis 3: Higher purchasing frequency is associated with less contracting for environmental sustainability.

Hypothesis 4: More political control is associated with more contracting for environmental sustainability.

Empirical setting and data

This study applies the theoretical framework to the case of contracting for environmental sustainability in public procurement procedures in Germany (also called 'green public procurement'). It is hard to gauge the baseline uptake of environmental criteria in contracts[viii] because there are no unified definitions and measurements. For example, baselines varied from 30% to 80% between Scandinavian countries (2003-2005), and around 20% in Belgium (2011-2016) (Grandia & Kruyen, 2020; Kozuch et al., 2022). Broad definitions of environmental criteria, such as in this study, find baseline uptake rates of around 70% in Italy (Testa et al., 2016) and 60% in Russia (Shadrina et al., 2022).

European Union citizens are broadly supportive of government contracting for environmental sustainability (Keulemans & Van de Walle, 2017). The German population has a strong preference for environmental sustainability but does not endorse specific sets of effective public policies (Blesse et al., 2024). The preference for environmental sustainability is ubiquitous in the German population (cf. yearly representative surveys, such as European Investment Bank, 2022; Federal Environment Agency, 2020); this also applies across all governing political parties (cf. also Eckersley et al., 2023).

Based on a law on procurement reporting (German Federal Government, 2016), the German Federal Statistics Office collects self-reported data from contracting authorities. Contracting authorities are legally obliged to report every procurement procedure. This study is based on three cross-sections of administrative data that cover all public procurement procedures in Germany in 2021 (n1=182,033), 2022 (n2=188,916), and 2023 (n3=195,493). Each row represents a public procurement procedure. The data is nested as follows: location (federal state) / buyer organization (anonymized) / procurement procedure.

All variables are manifest variables. The binary outcome variable is whether environmental sustainability (broadly defined) was contracted for in the requirements, contract execution specifications, selection criteria, or award criteria. The relevant reporting law provides a non-exhaustive list of examples of environmental criteria, such as the "requirement for

[viii] In the present study, 'contract' is used broadly to refer to binding promises made in a public procurement procedure, including, amongst others, product specifications, award criteria, and contract execution specifications.

fulfilling the prerequisites of an ISO 14024 Type I environmental label (e.g., Blue Angel, Nordic Swan, Austrian Ecolabel) or equivalent criteria" or the "requirement for compliance with the environmental management and audit scheme (EMAS) as set out by Regulation (EC) No. 1221/2009" (see Appendix 9 in German Federal Government, 2016).

The independent variables include asset specificity, measurability, transaction frequency, and political control. Asset specificity and measurability were calculated by taking the three leading digits of the Common Procurement Vocabulary code and matching the asset specificity scores from Potoski et al. (2025) to them. Asset specificity scores range from 1 to 5 with a mean score of 2.64 and a standard deviation of 0.45. Measurability was measured as the ease with which the contract object can be specified and measured. Measurability scores range from 1 to 5 with a mean score of 2.41 and a standard deviation of 0.49. For example, non-asset specific and easy-to-measure products are things such as coffee and tea (asset specificity score 2.0; measurability score 2.15). Highly asset specific and hard to measure products are things such as management training and continuing professional development (asset specificity score 3.58; measurability score 2.84) or IT consultancy and program development (asset specificity 3.43; measurability 2.98). Transaction frequency was operationalized as whether the contract object is procured a single time or more often than that, i.e., whether it was a framework contract (Casady et al., 2023; Petersen et al., 2022). Political control was operationalized as a rank-ordered variable that reflects a buyer organization's exposure to influence by elected government. It contains 6 regimes of public sector organizations in Germany, ranked by political control. Rank 6 contains organizations such as national or state ministries and municipalities, while rank 1 contains private law organizations under the dominant influence of public authorities such as Deutsche Bahn or defense firms where the federal state owns shares. The German legally defined terms for the buyer organizations within each regime are listed in Appendix A (Balbach, 2026).

The analysis controlled for buyer discretion (Domingos et al., 2024; Casady et al., 2023), i.e., the extent of procedural discretion in the award decision. Discretion is highest for awards based on cost and quality, and lowest for awards based solely on price. It also controlled for various kinds of competition (Petersen et al., 2022), such as the total number of bids received, the number of bids from the European Union, and whether lots were used in the procedure. The core model presented below included random intercepts per buyer location (in a federal state), exploiting the nested data structure. The year of the cross-section was controlled for. There is no intercorrelation among the control variables; they measure separate things (Appendix F in Balbach, 2026).

Research design and analytical strategy

The research design centers on iterative and transparent hypothesis testing. The 2021 cross-section on all procurement procedures in Germany was used to refine hypotheses, which were pre-registered in March 2025 (Balbach, 2025), and tested with the 2022 and 2023 cross-sections which were only received after preregistration. Because the sample size is very large, the estimation approach centers on variable importance rather than statistical significance testing. This means that the study presents p-values as well as confidence intervals and reflects on the model fit and the size of the average marginal effects.

The data analysis methods are observational and explanatory. Causal inference methods were not possible due to limitations of cross-sectional contract data with anonymized buyer organizations. Nevertheless, much is to be gained from the analysis because, logically, the focal explanatory variables precede the outcome variable. First, public buyers are told what needs to be bought, only afterwards do they think about how to buy it and whether to contract for environmental sustainability (Lundberg & Marklund, 2018).

The analytical strategy relies on the logit model, where the predictors are linearly related to the logit. The results below are based on a logistic generalized linear mixed model where the probability that environmental sustainability is contracted for is explained by variables $X_{1ij\text{-}pij}$ with each federal state having its own baseline level of the outcome probability (random intercept) b_{0j}, next to a cross-group fixed intercept β_0, where i is an observation in federal state j (Heinrich & Hill, 2010).

$$\operatorname{logit}(P_{ij}) = \ln\left(\frac{P_{ij}}{1-P_{ij}}\right) = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \cdots + \beta_p X_{pij} + b_{0j}$$

Adding the random intercept yields a lower value for the Akaike Information Criterion. Grouping by location (federal state) yields more internally similar groups than organization type, level of government or product group, based on the intra-class correlation coefficients, which was also reflected in smaller confidence intervals of the average marginal effects.

The data were cleaned by recoding impossible values. Contracts above the 95th percentile of contract value were excluded. Based on a visual analysis and tests of missingness, the missing at random data (MAR) were imputed with

multivariate imputation by chained equations, generating 50 imputed datasets in 5 iterations over which values stabilized and mingled. Robustness checks included pooled marginal effects to be robust against imputation-related uncertainty; a complete case analysis; separate estimation by years and by federal state in Germany; estimation with contract value covariate, with contract object type covariate; validating the statistical importance of the focal variables in the regression-based models with the mean decrease accuracy and mean decrease Gini for a random forest-based model and validating the contribution of each focal variable to the outcome variable using a partial dependence plot (see appendices in Balbach, 2026).

Results

The descriptive statistics of the data (2021-2023) before imputation are presented in Table 1. Only 9.2% of contracts included environmental sustainability criteria, even though the reporting law defines environmental sustainability broadly. This is surprising in light of the baseline rates in prior studies (cf. Empirical Setting & Data). Both asset specificity and measurability are unevenly distributed. Only 7.8% of contracts are framework agreements (transaction frequency). A plurality of buyer organizations is subject to very high political control (45.6%), however, the second-largest group of buyers is subjected to very low political control (20.0%).

Most award procedures featured low discretion (84.4%). Globally speaking, competition is very low: the median procurement procedure received two offers. However, the range of competition is broad; the total number of bids ranges from 1 to 2601 in Table 2. Competition from abroad is almost non-existent. 8.5% of contracts induce additional competition by splitting the contract into lots. A similar number of contracts is reported for each year, indicating that the quality of the 2021 data (the first ever reporting period) is sufficient for inclusion in the model.

91.1% of purchases are below the European Union threshold above which they must be published on Tenders Electronic Daily. Only 8.4% are above the threshold. This indicates that the results in this study are a more accurate reflection of government contracting in general than the results from studies using Tenders Electronic Daily data. After excluding the 95th percentile, contract value ranges from €1001 to €1,328,556, with a median of €79,069 and an exponential decay-shaped distribution. Only 1.5% of contracts are awarded to firms outside of Germany, however, 32% of contracts do not declare the headquarters of the contractor (see Appendix C in Balbach, 2026).

49.3% are construction contracts, 28.1% service contracts, and 22.5% supply contracts, with a remainder of 0.2% for social and specific services. This predominance of construction work is reflected in the most common product group names: 1. Works for complete or part construction and civil engineering work, 2. Tradesman services, 3. Engineering and architect consultancy, 4. Other construction work, 5. Site preparation work, 6. Motor vehicles, 7. IT consultancy and program development, 8. Building and construction material, 9. Other Laboratory, optical and precision equipment (excl. glasses) 10. Computers and tablets. Furthermore, buyer organizations always assign common procurement vocabulary (CPV) code 1, whereas 85.5% and 93.2% of CPV codes 2 and 3 are missing, respectively.

Table 1. Focal variables.

Variable	Valid (%)	Missing (%)	Moments	Frequencies (% of Valid)
Environmental sustainability	525,676 (97.7%)	12,443 (2.3%)	Min 0 Mean 0.1 Max 1	No: 477,565 (90.8%) Yes: 48111 (9.2%)
Asset specificity	509,347 (94.7%)	28,772 (5.3%)	Min 1.7 Median 2.4 Max 3	160 distinct numeric values
Measurability	509,347 (94.7%)	28,772 (5.3%)	Min 1.7 Median 2.8 Max 3.6	160 distinct numeric values
Transaction frequency (framework agreement)	376,124 (69.9%)	161,995 (30.1%)	Min 1.7 Mean 0.1 Max 3.6	No: 346,968 (92.2%) Yes: 29,156 (7.8%)
Political control	538,119 (100%)	0 (0%)	Min 1 Median 5 Max 6	1: 107,410 (20%) 2: 6847 (1.3%) 3: 25,634 (4.8%) 4: 65,204 (12.1%) 5: 87,380 (16.2%) 6: 245,644 (45.6%)

Table 2. Control variables.

Variable	Valid (%)	Missing (%)	Moments	Frequencies (% of Valid)
Discretion	387,629 (72%)	150,490 (28%)	Min 1 Median 1 Max 4	1: 327,233 (84.4%) 2: 2,086 (0.5%) 3: 53,140 (13.7%) 4: 5,170 (1.3%)
Total number of bids	405,886 (75.4%)	132,233 (24.6%)	Min 1 Median 2 Max 2601	153 distinct values
Bids from European Union	288,239 (53.6%)	249,880 (46.4%)	Min 0 Median 0 Max 186	60 distinct values
Contract split into lots	448,195 (83.3%)	89,924 (16.7%)	Min 0 Mean 0.1 Max 1	No: 410,111 (91.5%) Yes: 38,084 (8.5%)
Year	538,119 (100%)	0 (0%)	/	2021: 174,123 (32.4%) 2022: 179,030 (33.3%) 2023: 184,966 (34.4%)

Of the share of purchases, the municipal level represents 51.7%, the state level 29.6%, the federal level 10.8%, and other levels 7.9%. The share of purchases also reflects the population size of states. The most common procedure is the public tender (33.8%), followed by the negotiated award without participation competition (27.2%) and the restricted tender without participation competition (17.9%). More detail can be found in Appendix C in Balbach (2026).

The results in Table 3, on the logistic generalized linear mixed model, support the theoretical mechanisms (cf. Figure 2). The assumptions underlying a logistic Generalized Linear Mixed Model were met (Appendix F in Balbach, 2026). The results are robust to specification differences, data subgroups and outliers, imputation-related uncertainty, and different modeling approaches (Appendices G-N in Balbach, 2026).

By far, asset specificity plays the most important role in how contracting out stymies the pursuit of environmental sustainability (cf. z score in Table 3). Both measurability and political control are associated with increased use of environmental criteria, however, their effects are smaller than the effect of asset specificity. Contracting for environmental sustainability was positively associated with both the use of framework agreements (indicating a future elevated transaction frequency due to repeat procurement) and the presence of buyer discretion in picking a winning bid. The former result is at odds with the transaction cost theoretic hypothesis on transaction frequency. It seems that the

Table 3. Average marginal effects on contracting for environmental sustainability.

Variable	AME	Std Error	z	Pr(>|z|)	S	CI 2.5%	CI 97.5%
Asset specificity	-0.09573	0.0140027	-6.84	<0.001	36.8	-0.12318	-0.06829
Measurability	0.005135	0.0013927	3.69	<0.001	12.1	0.002405	0.007865
Framework agreement	0.014685	0.002589	5.67	<0.001	26.1	0.009611	0.01976
Political control	0.006123	0.0009127	6.71	<0.001	35.6	0.004334	0.007912
Discretion	0.015097	0.0022377	6.75	<0.001	35.9	0.010711	0.019482
Total bids	0.000326	0.0000717	4.54	<0.001	17.5	0.000185	0.000466
EU bids	-0.00521	0.0008987	-5.8	<0.001	27.1	-0.00697	-0.00345
Contract lots[ix]	-0.0101	0.0020102	-5.02	<0.001	20.9	-0.01404	-0.00616
Year 2022	-0.00313	0.0010408	-3.01	0.00264	8.6	-0.00517	-0.00109
Year 2023	0.006794	0.0013625	4.99	<0.001	20.6	0.004123	0.009464

[ix]The variable 'contract lots' refers to whether the contract was split into several lots ("subcontracts") for which firms can bid separately. This is a means to increase competition.

prospectively higher transaction frequency in framework contracts does not raise the transaction hazards of using environmental criteria. Splitting contracts into several lots that firms can bid for as well as the participation of bidders from outside Germany, seem to inhibit government contracting for environmental sustainability. This implies that more competition may sometimes be incompatible with the pursuit of environmental sustainability in government contracting.

The highest rate of contracting for environmental sustainability (e.g., 30%) is achieved when asset specificity is low and both political control and buyer discretion are high (cf. bottom-right panel in Figure 3). For example, a public buyer at a national ministry buys chairs, based on cost and quality (rather than a fixed-price contact). Generally, the asset specificity of the purchased goods, services and works matters most. Once asset specificity is medium high (2.5 on a scale from 1-very low to 5-very high), even the most politically controlled buyer organizations contract for environmental sustainability at the same rate (~10%) as the least politically controlled organizations when asset specificity is low (~1.6). At the higher end of the spectrum of asset specificity, other variables such as political control or buyer discretion matter very little.

A buyer organization in Berlin, Saxony or Mecklenburg-Western Pomerania is far more likely to contract for environmental criteria than a buyer organization in the Saarland, Bremen or the Rhineland-Palatinate (Figure 4). The low propensity of buyer organizations in Bremen and the Saarland may be due the general statistical pattern that small territories have more extreme statistics since they do not average out over a larger sample size. This reading is supported by their relatively large confidence intervals. Nevertheless, an interesting pattern emerges whereby the biggest absolute values of intercepts are located in states in the East of Germany. The size of the effect of asset specificity varies by the location of the buyer organization (Appendix G in Balbach, 2026). This indicates that regional economies may moderate the extent to which asset specificity inhibits the use of environmental criteria in government contracts.

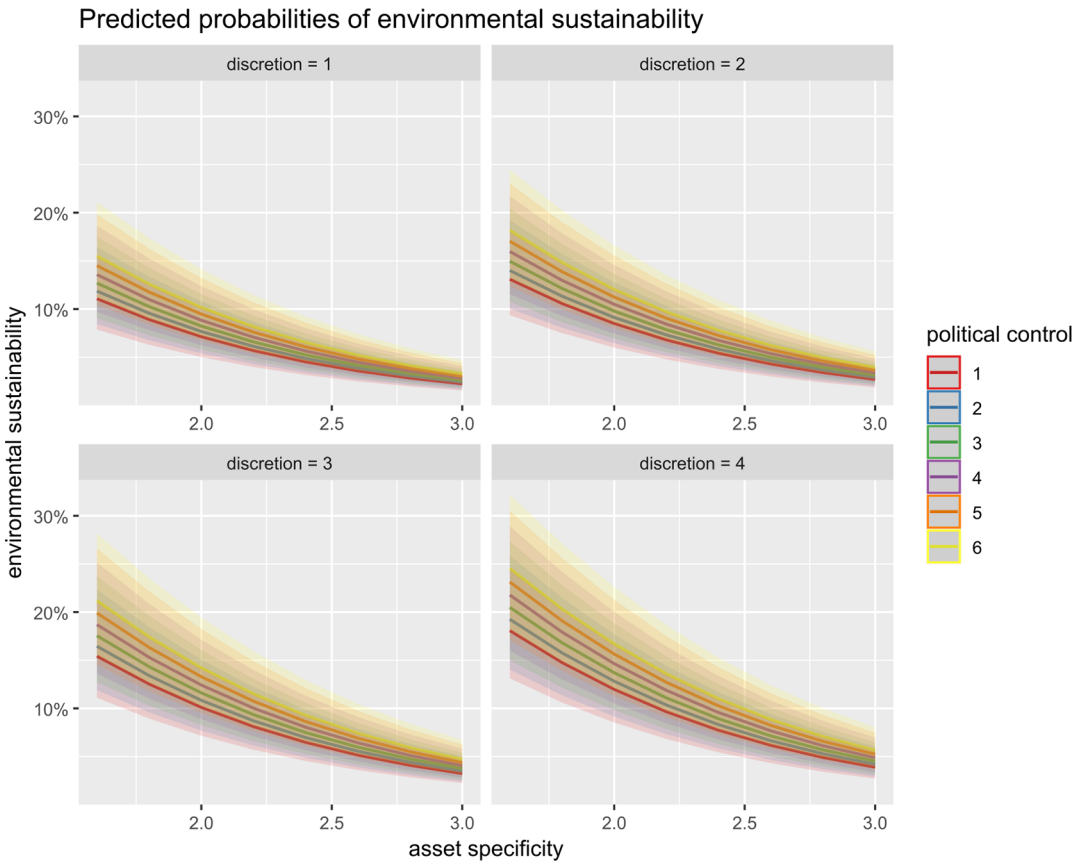

Figure 3. The interplay of asset specificity, political control and buyer discretion. Panels show model-based predicted probabilities of the environmental sustainability outcome as a function of asset specificity (x-axis) for four levels of discretion (1 to 4). Colored lines correspond to levels of political control (1 to 6). Shaded bands represent 95% Wald-type confidence intervals around the predictions. Across all discretion levels, predicted probabilities decline as asset specificity increases, while higher political control is associated with higher predicted probabilities at a given level of asset specificity.

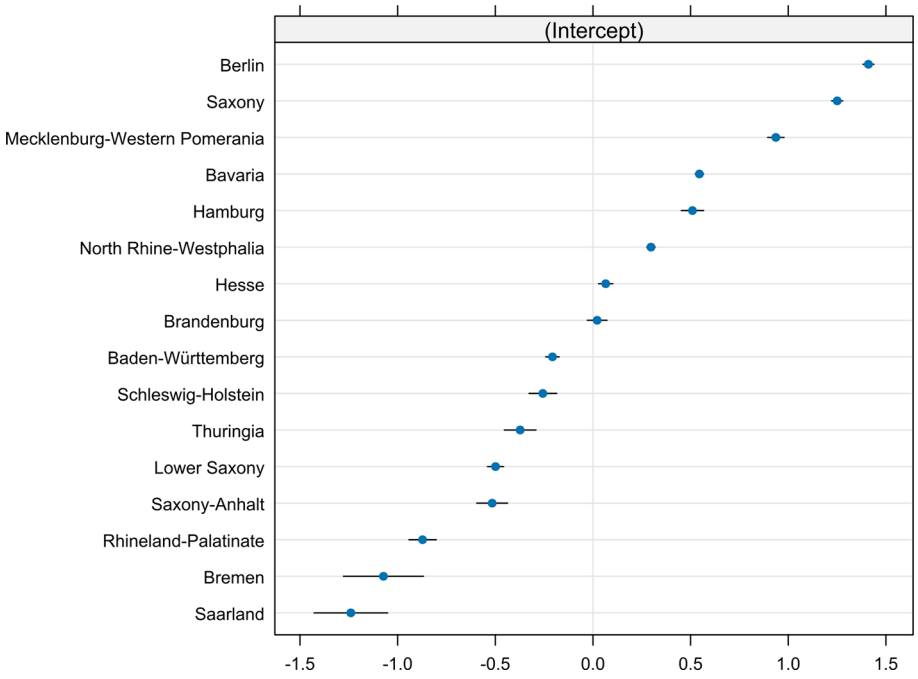

Figure 4. Random intercept by location (Federal State). The figure displays estimated state-specific random intercepts from the multilevel model for the German federal states (in German: "Bundesländer"), ordered by magnitude. Points indicate the estimated deviation of each state's intercept from the overall (grand-mean) intercept (0 on the x-axis); horizontal whiskers denote 95% Wald-type confidence intervals. Positive values indicate higher baseline levels of the outcome relative to the overall mean, and negative values indicate lower baseline levels.

Discussion

Generally, the results support the theoretical framework on when and how delegation stymies the demand power of the hollow state (cf. Table 3 and Figure 2). Hypotheses 1 to 3 concern how contracting out stymies demand power. Higher asset specificity was associated with less contracting for environmental sustainability, while higher measurability was associated with more contracting for environmental sustainability. However, higher purchasing frequency (framework agreement) was unexpectedly associated with more contracting for environmental sustainability. Hypothesis 4 describes how decentralization stymies demand power. As expected, more political control was associated with more contracting for environmental sustainability. The by far strongest focal variable predicting whether buyers contract for environmental sustainability was asset specificity, followed by measurability and political control. However, transaction frequency did not perform as expected, having a positive association instead of negative one. This echoes the broader transaction cost theoretic literature where transaction frequency is often not included in analyses and does not consistently perform as hypothesized (Cuypers et al., 2021; Mayer, 2009; Rindfleisch & Heide, 1997).

When and how does delegation stymie the administrative capacity of the hollow state to steer contractors toward collective goals? Generally, the descriptive findings, particularly the low baseline rate of contracting for environmental sustainability (9.2% from 2021-2023), underscore the first form of delegation (Figure 2), namely, that environmental sustainability in government contracting is regulated in a voluntary and nonbinding self-regulatory way. This baseline is lower than the baselines reported for Scandinavia, Italy and Belgium in the preceding decade (cf. Empirical Setting and Data; Grandia & Kruyen, 2020; Kozuch et al., 2022; Testa et al., 2016). Given the very high variation of the baseline rates between countries, this finding on Germany is not surprising. What does surprise is that self-reported uptake rates (here: by the buyer organization to the federal statistical office) normally tend to be higher than 'objective' uptake rates (Zhang et al., 2023). Since many of the estimates of uptake in other countries are objective, or at least, machine-counted (Grandia & Kruyen, 2020), the findings on Germany are surprisingly low. A potential explanation may be found in the other studies having worked with subsamples of the total population of government contracts.

The findings provide mixed evidence for the effects of contracting out (Figure 2). Public buyers assign environmental criteria to generally less hazardous transactions with lower asset specificity and higher measurability. What distinguishes the variables of transaction frequency and political control from asset specificity and measurability is their object of analysis. The former are about the context wherein a good, service or work is transacted, while the latter are about intrinsic

qualities of the good, service or work itself. In that sense, the mechanistic (Elster, 2007) core of transaction cost theory that describes properties of things procured rather than their context, has been confirmed in a novel empirical setting, concerning the collective goal of environmental sustainability (cf. Cuypers et al., 2021; Mayer, 2009; Rindfleisch & Heide, 1997). It remains unclear how exactly asset specificity, and to a lesser extent, measurement difficulty, decrease contracting for environmental sustainability – whether environmental criteria are simply too hard to formulate for purchase contracts of complicated goods, services and works, or whether buyers avoid using available criteria because they cannot be verified upon delivery of the good, service or work (compare Balineau & Dufeu, 2010; Dorée et al., 2011; Rietbergen et al., 2015).

The findings support the theoretical framework with regards to decentralization (Figure 2), whereby less politically controlled public sector organizations are less inclined to contract for collective goals. However, it must be noted that the political control variable, spanning all public sector organizations in Germany, is a broad and global variable. At this level of generality, political control over a buyer organization appears to influence contracting for environmental sustainability via political steering (Dahlström et al., 2011; Ebinger et al., 2019; Hartlapp, 2020). Furthermore, a robustness check using random forest models (Appendix N in Balbach, 2026) indicated that the effect of political control may have been mostly driven by the political control regime of rank 5. This regime includes upper, middle and lower authorities at the federal and state levels that are headed by unelected officials who are always overseen by an entity with an elected government head (cf. also Appendix A in Balbach, 2026). Public buyers at these organizations are far more likely to contract for environmental sustainability. This implies that each political control regime may have its own unique mix of formal organizational autonomy and political influence by elected officials (cf. Dahlström et al., 2011; Windoffer, 2020) or that there are other unobserved organizational characteristics that differ between them.

The findings fill the first identified research gap concerning overall state capacity. Covering all levels of government, all kinds of organizations obliged to conduct public procurement, and all sectors including defense, this study offers highly general or generalizable insights across extremely heterogeneous contexts, spanning from a rural municipality to a memorial foundation to a public water company. Delegation in the hollow state does not necessarily stymie the pursuit of collective goals at the organizational level (Heinrich et al., 2010; Kelly LeRoux & Nathaniel S. Wright, 2021a). In distinction to prior research (Bendor et al., 2001; Heinrich et al., 2010; Kelly LeRoux & Nathaniel S. Wright, 2021a), this study theorizes that delegation in the hollow state stymies the achievement of collective goals when it diminishes state-level demand power, i.e., a hollow state's power to make demands on how goods, services and works are produced so that they contribute to collective goals (compare 'administrative power' in Kapczynski & Michaels, 2024). Demand power arises from the uniform and comprehensive pursuit of a collective goal across markets (Lundberg et al., 2016; cf. also Saussier & Tirole, 2015, p. 7). In contrast, the observed pursuit of environmental sustainability was noncomprehensive, only comprising 9.2% of contracts. Furthermore, the pursuit was not uniform but fragmented between product groups. Public buyers barely contracted for the environmental sustainability of highly asset-specific and hard-to-measure goods, services and works. However, this study did not account for the level of stringency of the environmental criteria, nor for the extant share of environmentally friendly products per product market. The latter two conditions for demand power may yield a more complete picture of demand power.

The findings also fill the second identified research gap on when and how which forms of delegation, or combinations thereof, lead to low state capacity. The findings imply that it is one specific type of delegation, namely, contracting out, that diminished demand power, while decentralization to less politically controlled entities played a lesser role (Heinrich et al., 2010, p. i7). This may imply that the effect of delegation on hollow state capacity is not mainly contingent on the degree of political control by elected officials, as Weber, Wilson and Finer would have it (Weber, 2010 [1921]; Wilson, 1887). Rather, the effect of delegation may be contingent on whether the implementing entity is directly responsible to the public interest (Friedrich, 1940). Government contracts delegate public service delivery to contractors who are not directly responsible to the public interest, but towards the contract.

While this study shows that political control of the buyer organization increased contracting for environmental sustainability (cf. Bryngemark et al., 2023; Smith et al., 2016), prior studies showed that less independent public sector organizations are most at risk of favoritism in response to donations (Fazekas et al., 2023) or lobbying (Dusso et al., 2019) and most likely to have noncompetitive procedures and outcomes (Dahlström et al., 2021). This indicates that there may be a non-complementarity between the tools of government used to pursue the collective goal of environmental sustainability versus the tools used to pursue the collective goal of fair competition. This shows the importance of not conflating welfare with the objective-effective pursuit of a single collective goal. Environmental sustainability is but one collective goal, next to prosperity, personal liberty, social justice, and others (compare Britton-Purdy et al., 2020; Hausman, 2024; K. LeRoux & N. S. Wright, 2021). That is why one cannot take this study to have shown that contracting out in the hollow state stymies welfare. Welfare consists of achieving multiple collective goals, next to individual goals.

This implies that normative lessons must be drawn with caution from this study, while there is, indeed, initial evidence that some forms of delegation, notably, contracting out, do harm the pursuit of the collective goal of environmental sustainability within the entirety of a hollow state.

Furthermore, the effect of political control on contracting for environmental sustainability was secondary to the strong inhibiting effect of the asset specificity of contracted-for goods, services and works. Hence, the findings of this study raise doubts about the promise of political control in strategic government contracting to enhance welfare (Grandia & Meehan, 2017; Hartlapp, 2020). However, for fully different reasons than the more common doubts about the trustworthiness of elected officials (Lundberg et al., 2016; Spagnolo, 2012) which seek to benefit their clientele with government contracts (Dávid-Barrett & Fazekas, 2020; Essien & Etim, 2021; Sundell & Lapuente, 2012). Instead, contracts may not be able to express collective goals for highly and very highly asset-specific goods, services and works (cf. Figure 3) because the inhibiting effect of asset specificity overpowers the effect of political control.

Limitations

Several limitations apply. First, while offering causal mechanisms, this study relies on observational-explanatory methods. This is due to the anonymization of the buyer organization in the data set provided to the researchers. Anonymization obstructed more direct ways of accounting for unobserved heterogeneity to perform causal inference. Hence, this study cannot fully rule out that unobserved factors, such as organizational capacity, path dependence or budget pressures, confound the estimated effects. As a remedy, the study relied on an open and iterative research design. Starting from a theoretical framework that solely relied on transaction cost theory, the preliminary analysis with the 2021 dataset provided several learnings that prompted a reconceptualization of the role of political control. This enabled the final pre-registered study to test the refined framework using the 2022 and 2023 cross-sections as hold-out datasets.

Second, several limitations concern the measurement of demand power. (a) Demand power was measured as using environmental criteria in any section of the contract. This may be viewed as greenwashing since the sustainability criteria are not distinguished by their stringency. However, the low baseline rate of contracting for environmental sustainability (9.2%) may convince more skeptical observers that there was no greenwashing. (b) A related question concerns a specific case, namely, whether buyer organizations are truly expressing their demand power in those cases where sustainability criteria are mandated by law. For the purpose of this study, namely, to examine demand power as a dimension of *state-level* administrative capacity, this specific case can also be regarded as an expression of state-level demand power. Since it is the state, via binding regulation in this specific case, that formulates public demand. (c) Some collective goals, such as favoring local firms, may not be mentioned in a survey or visible in procurement language but still be pursued. Thus, another limitation of this study is its focus on socially acceptable and publicly disclosed collective goals. (d) Another measurement limitation concerning demand power is that policy criteria in public procurement procedures are but one instance of demand power. Other forms of steering contractors toward collective goals are grants and equity in contractors (Bozeman, 1987). (d) Unlike performance, which is fully captured in a cross-sectional measurement, demand power is a latent concept. Future research may refine the measurement model for demand power by constructing an indicator that builds on multiple exercises of demand power. (e) This study covered the supplier selection and bid selection phases of the public procurement procedure, but not the contract execution phase (Figure 1). Explicit mentions of sustainability criteria may or may not be followed by their implementation during contract execution. However, if sustainability criteria are not mentioned in the procurement procedure, there is no reason to expect them to be implemented.

Third, several limitations apply to this study due to its focus on the state level. Future inquiries into organization-level demand power will need to adapt the model. The focus of this study lies on *state-level* administrative capacity. More granular political control variables would be appropriate for organization-level models, potentially detecting stronger political control effects for specific organizations (Williams, 2021). For the purpose of this study, the state-level effects have high external validity because of the extensive dataset of procurement procedures across highly heterogeneous contexts. Furthermore, this study possesses a high ecological validity, owing to the analysis of reported, granular contractual design elements rather than latent attitudes or general perceptions of sustainable public procurement at an organizational level (e.g., Testa et al., 2012). Furthermore, the logistic generalized linear mixed model, coupled with a series of robustness checks (Balbach, 2026), helped circumvent many erroneous but common approaches to statistical inference from imperfect government data with missing values (van Ginkel et al., 2020). Covering all organizations that are obliged to conduct public procurement at all levels of government and all sectors, provided for a heterogeneous and large sample.

Fourth, due to the federal administrative system of the country studied here, it seems plausible that the results are most likely to replicate in other historically decentralized states, such as the United States of America or the Republic of India (Bach & Jann, 2010, pp. 446-451; Heinrich et al., 2010, pp. i7-i9; Jugl, 2025). However, other countries may be more

heterogeneous with regards to the degree of accountability and discretion of public sector organizations (Kim et al., 2026), or value collective goals such as sustainability differently.

Future research

To provide a more precise image of the hollow state, future research is advised to sharpen the analysis of the effect of delegation on the pursuit of collective goals. The first four suggestions are about improving state-level models of demand power, while the fifth is about extending the framework of this study to the organizational level and the sixth suggestion is about modeling buyer choices at the individual level.

First, future research may explore contracting for collective goals in requirements, selection criteria, award criteria, and contract execution specifications, respectively. This may allow to account for the stringency by which collective goals are applied to government contracts, which is an important condition for demand power (Lundberg et al., 2016). Second, future research may test whether the mechanisms observed in this study hold for other collective goals, such as social equity or innovation.

Third, if the data permits, future research could control for characteristics of buyer organizations, such as their organizational culture, budget constraints, network density, and network trust (Lindström et al., 2022; Lundberg & Bergman, 2017; Agranoff & McGuire, 1998). Thereby, one could refine the mechanisms (Figure 2) with organizational characteristics that mediate or moderate demand power, and perform causal inference.

Fourth, using the exact geographical location of the buyer organization, in conjunction with the product group, would allow to estimate the third condition for demand power, namely, the extent to which markets comply with collective goals, e.g., the share of environmentally friendly products.

Fifth, some scholars may be more interested in analyzing organization-level demand power (rather than the state-level analysis in this study). A future organization-level study may include random slopes per buyer organization, since the organizational perception of transaction properties may differ (compare static scores in Potoski et al., 2025). The model in this study did not include such random slopes because they did not improve model selection criteria such as the Akaike Information Criterion (Appendix F in Balbach, 2026). Further, future research may seek to more granularly characterize types of public sector organizations that strongly contract for collective goals, such as the organizations in the rank 5 political control regime in this study (Appendices A and N in Balbach, 2026).

Sixth, future qualitative and experimental research could explicitly model the heuristics used by public buyers when dealing with both elected officials and the qualities of the good, service or work they need to buy. For instance, this study showed that asset specificity overwhelms political control (Figure 3) and that contract splitting has an adverse effect on contracting for sustainability. At higher levels of specificity, politically controlled organizations fare no better than less controlled ones at contracting for sustainable goods, services and works. An explicit heuristics approach would translate theoretical framework to decision trees at the individual level of analysis (Williams, 2021), modeling how agency drift and transaction cost economizing interact with each other within individual public buyers.

Conclusion

Drawing on 566,442 public procurement procedures in Germany from 2021-2023, the logistic generalized linear mixed model showed that asset specificity and measurement difficulty are associated with less contracting for environmental sustainability, while the transaction frequency (framework agreements) and the political control of buyer organizations is associated with more contracting for environmental sustainability. Asset specificity, as an intrinsic transaction property of things procured, was most strongly negatively associated.

The broader theoretical implications of these findings relate to the limits of strategic government contracting given the fragmentation of the demand side in the hollow state. Properties such as asset specificity cannot be modified by buyers. For example, a quantum computer has high asset specificity, which no public buyer can modify. Hence, contracting for collective goals may be less suitable or feasible for complex and complicated goods, services and works. For some collective goals at least, other tools of government, such as emission pricing programs and command-and-control regulation, may be more adequate (Kapczynski & Michaels, 2024; Salamon & Lund, 1989).

Buyer organizations that were subject to weaker political control regimes contracted less for environmental sustainability. This is worrying, given the tendency of the last 50 years to decenter classical public administration by delegating to other more autonomous entities. Elected officials may be well-advised to consider opening a dialogue about environmentally sustainable goods, services or works with more formally autonomous organizations in their realm of influence, such as

foundations or agencies. Dialogues enable elected officials to communicate their preferences while acknowledging the formal limits to their influence.

To date, most studies on the hollow state studied subsets of public sector organizations or contracts. By examining the entirety of a hollow state and multiple forms of delegation, this study finds that certain forms of delegation, most notably contracting out, moderately reduce the power of the hollow state to work toward collective goals.

Ethics and consent

Ethical approval and consent were not required.

Data availability

The preregistration of this study is available at OSF: https://doi.org/10.17605/OSF.IO/EDFYS

The appendices A to N are available at Zenodo: https://doi.org/10.5281/zenodo.18860675

The original data was obtained from a third party and restrictions apply to the availability of the data. The data was obtained from the German Federal Statistical Office ("Destatis"). The institution of the researcher and the German Federal Statistical Office entered into a contract regarding the use of and access to the data. The institution of the researcher is not permitted to share the original data with other institutions or individuals at other institutions. For a reader or reviewer to access the data by the same means as the author, the reader or reviewer must submit a formal request to the German Federal Statistical Office for access to the research datasets based on the Regulation on Statistics on the Award of Public Contracts and Concessions ("Vergabestatistikverordnung – VergStatVO" 2016).

References

Agranoff R, McGuire M: **Multinetwork Management: Collaboration and the Hollow State in Local Economic Policy.** *J. Public Adm. Res. Theory.* 1998; **8**(1): 67–91.
Publisher Full Text

Andrews RNL: **Environmental Regulation and Business 'Self-Regulation'.** *Policy. Sci.* 1998; **31**(3): 177–197.
Publisher Full Text

Bach T, Hammerschmid G, Löffler L: **More delegation, more political control? Politicization of senior-level appointments in 18 European countries.** *Pub. Pol. Adm.* 2018; **35**(1): 3–23.
Publisher Full Text

Bach T, Jann W: **Animals in the administrative zoo: organizational change and agency autonomy in Germany.** *Int. Rev. Adm. Sci.* 2010; **76**(3): 443–468.
Publisher Full Text

Baekgaard M, Mortensen PB, Bech Seeberg H: **The bureaucracy and the policy agenda.** *J. Public Adm. Res. Theory.* 2018; **28**(2): 239–253.
Publisher Full Text

Balafoutas L, Kerschbamer R: **Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions.** *J. Behav. Exp. Financ.* 2020; **26**: 100285.
Publisher Full Text

Balbach T: **Green Public Procurement in Germany [Registration].** 2025, March 10.
Publisher Full Text

Balbach T: **How Does Delegation Affect the Administrative Capacity of the Hollow State: Appendix. [Extended Data].** 2026, March 4.
Publisher Full Text

Balineau G, Dufeu I: **Are fair trade goods credence goods? A new proposal, with French illustrations.** *J. Bus. Ethics.* 2010; **92**(Suppl 2): 331–345.
Publisher Full Text

Bendor J, Glazer A, Hammond T: **Theories of delegation.** *Annu. Rev. Polit. Sci.* 2001; **4**: 235–269.
Publisher Full Text

Bertelli AM: *Democracy administered: how public administration shapes representative government.* Cambridge University Press; 2021.

Berwick E, Christia F: **State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate.** *Annu. Rev. Polit. Sci.* 2018; **21**: 71–91.
Publisher Full Text

Biela J, Papadopoulos Y: **The empirical assessment of agency accountability: a regime approach and an application to the German Bundesnetzagentur.** *Int. Rev. Adm. Sci.* 2014; **80**(2): 362–381.
Publisher Full Text

Blesse S, Dietrich H, Necker S, *et al.*: **Wollen die Deutschen beim Klimaschutz Vorreiter sein und wenn ja, wie? Maßnahmen aus Bevölkerungsperspektive.** *ifo Schnelldienst.* 2024; **77**(1): 39–43.

Bogumil J: **Regierung und Verwaltung.** *Politische Bildung.* 2007; **4**.

Bosio E, Djankov S: *How large is public procurement?.* World Bank Blogs; 2020, February 5, 2020. Retrieved August 9, 2024.
Reference Source

Bozeman B: *All organizations are public: bridging public and private organizational theories.* Jossey-Bass; 1987.

Britton-Purdy J, Grewal DS, Kapczynski A, *et al.*: **Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis.** *Yale Law J.* 2020; **129**(6): 1784–1835.

Brown TL, Potoski M, Van Slyke DM: **Managing Public Service Contracts: Aligning Values, Institutions, and Markets.** *Public Adm. Rev.* 2006; **66**(3): 323–331.
Reference Source

Bryngemark E, Söderholm P, Thörn M: **The adoption of green public procurement practices: Analytical challenges and empirical illustration on Swedish municipalities.** *Ecol. Econ.* 2023; **204**: 107655.
Publisher Full Text

Carpenter D: **The Political Foundations of Bureaucratic Autonomy: A Response to Kernell.** *Studies in American Political Development.* 2001; **15**(1): 113–122.
Publisher Full Text

Casady CB, Petersen OH, Brogaard L: **Public procurement failure: The role of transaction costs and government capacity in procurement cancellations.** *Public Manag. Rev.* 2023; **1-28**.
Publisher Full Text

Cheng YD, Shi Y, Andrew SA: **Bridging charitable support and public service performance: a preliminary analysis of large US city park systems.** LeRoux K, Wright NS, editors. *Performance and Public Value in the 'Hollow State'.* Elgar; 2021.

Cuypers I, Hennart J-F, Silverman BS, *et al.*: **Transaction Cost Theory: Past Progress, Current Challenges, and Suggestions for the Future.** *Acad. Manag. Ann.* 2021; **15**(1): 111–150.
Publisher Full Text

Dahlström C, Fazekas M, Lewis DE: **Partisan Procurement: Contracting with the United States Federal Government, 2003-2015.** *Am. J. Polit. Sci.* 2021; **65**(3): 652–669.
Publisher Full Text

Dahlström C, Peters BG, Pierre J: **Steering from the Centre Strengthening Political Control in Western Democracies.** Dahlström C, Peters BG, Pierre JON, editors. *Steering from the Centre.* University of Toronto Press; 2011; pp. 3–24.
Reference Source

Dávid-Barrett E, Fazekas M: **Grand corruption and government change: an analysis of partisan favoritism in public procurement.** *Eur. J. Crim. Policy Res.* 2020; **26**(4): 411–430.
Publisher Full Text

Diamond J: *Collapse: how societies choose to fail or succeed.* Allen Lane; 2005.

Dimand A-M, Neshkova MI: **Buying green in U.S. local government: Internal commitment and responsiveness to external pressures.** *Public Adm.* 2024; **102**(2): 644–667.
Publisher Full Text

Domingos FD, Heinrich CJ, Saussier S, *et al.*: **Navigating contract renegotiations with sustainability at the helm: Societal benefits and transaction costs.** *J. Strateg. Contract. Negot.* 2024; 20555636231222171.
Publisher Full Text

Dorée A, Wal G, Boes H: **Client leadership in sustainability: How the Dutch railway agency created CO2 awareness in the industry.** *Association of Researchers in Construction Management, ARCOM 2011 - Proceedings of the 27th Annual Conference.* 2011; **2**: 685–694.

Dubnick MJ, Frederickson HG: **Accountable Agents: Federal Performance Measurement and Third-Party Government.** *J. Public Adm. Res. Theory.* 2010; **20**: i143–i159.
Reference Source

Dulleck U, Kerschbamer R, Sutter M: **The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition.** *Am. Econ. Rev.* 2011; **101**(2): 526–555.
Publisher Full Text

Dusso A, Holyoke TT, Schatzinger H: **The Influence of Corporate Lobbying on Federal Contracting.** *Soc. Sci. Q.* 2019; **100**(5): 1793–1809.
Publisher Full Text

Ebinger F, Veit S, Fromm N: **The partisan–professional dichotomy revisited: Politicization and decision-making of senior civil servants.** *Public Adm.* 2019; **97**(4): 861–876.
Publisher Full Text

Eckersley P, Flynn A, Ferry L, *et al.*: **Austerity, political control and supplier selection in English local government: implications for autonomy in multi-level systems.** *Public Manag. Rev.* 2023; **25**: 1–21.
Publisher Full Text

Elster J: *Explaining Social Behaviour: More Nuts and Bolts for the Social Sciences.* Cambridge University Press; 2007.

Essien EE, Etim GS: **Policies versus politics: Which is the more important determinant of public sector supplier selection decisions in Nigeria?** *Int. J. Bus. Emerg. Mark.* 2021; **13**(1): 74–88.
Publisher Full Text

European Investment Bank: **The EIB Climate Survey 2021-2022.** 2022.
Reference Source

Fazekas M, Ferrali R, Wachs J: **Agency Independence, Campaign Contributions, and Favoritism in US Federal Government Contracting.** *J. Public Adm. Res. Theory.* 2023; **33**(2): 262–278.
Publisher Full Text

Federal Environment Agency: **Environmental Awareness in Germany 2018.** 2020.
Reference Source

Ford MR: **The challenge of nonprofit accountability and quality control in the urban hollow state: the case of public education.** LeRoux K, Wright NS, editors. *Performance and Public Value in the 'Hollow State'.* Elgar; 2021; pp. 25–42.

Friedrich CJ: **Public Policy and the Nature of Administrative Responsibility.** Friedrich CJ, Mason ES, editors. *Public Policy. A Yearbook of the Graduate School of Public Administration, Harvard University.* Harvard University Press; 1940; 3–24.
Reference Source

Gammeltoft-Hansen T: **Can Privatization Kill?.** *New York Times.* 2012.
Reference Source

García-Rodríguez A, Núñez M, Pérez MR, *et al.*: **Sustainable visions: unsupervised machine learning insights on global development goals.** *PLoS One.* 2025; **20**(3): e0317412.
Publisher Full Text

Gersdorf H: **Warum die bisherige NGO-Förderpraxis verfassungswidrig ist [Why the current practice of funding NGOs is unconstitutional].** *Frankfurter Allgemeine Zeitung.* 2025.
Reference Source

Geyskens I, Steenkamp J-BEM, Kumar N: **Make, Buy, or Ally: A Transaction Cost Theory Meta-Analysis.** *Acad. Manag. J.* 2006; **49**(3): 519–543.
Publisher Full Text

Grandia J, Kruyen PM: **Assessing the implementation of sustainable public procurement using quantitative text-analysis tools: A large-scale analysis of Belgian public procurement notices.** *J. Purch. Supply Manag.* 2020; **26**(4): 100627.
Publisher Full Text

Grandia J, Meehan J: **Public procurement as a policy tool: using procurement to reach desired outcomes in society.** *Int. J. Public Sect. Manag.* 2017; **30**(4): 302–309.
Publisher Full Text

Halonen K-M: **Is public procurement fit for reaching sustainability goals? A law and economics approach to green public procurement.** *Maastricht J. Eur. & Comp. L.* 2021; **28**(4): 535–555.
Publisher Full Text

Harris AJ: **Groups That Run N.Y.C. Shelters Are Riddled With Problems, Report Finds.** *New York Times.* 2024.
Reference Source

Hartlapp M: **Measuring and Comparing the Regulatory Welfare State. Social Objectives in Public Procurement.** *Ann. Am. Acad. Pol. Soc. Sci.* 2020; **691**: 68–83.
Reference Source

Hausman DM; **Philosophy of Economics.** Zalta EN, Nodelman U, editors. *The Stanford Encyclopedia of Philosophy.* Fall 2024 ed. 2024.

Heinrich CJ, Hill CJ: **Multilevel Methods in the Study of Bureaucracy.** *The Oxford Handbook of American Bureaucracy.* Oxford University Press; 2010; pp. 497–517.
Publisher Full Text

Heinrich CJ, Lynn LE Jr, Milward HB: **A State of Agents? Sharpening the Debate and Evidence over the Extent and Impact of the Transformation of Governance.** *J. Public Adm. Res. Theory.* 2010; **20**(suppl_1): i3–i19.
Publisher Full Text

Jugl M: **Patterns of bureaucracy: Conceptualizing administrative traditions.** *J. Public Adm. Res. Theory.* 2025; muaf012.
Publisher Full Text

Kapczynski A, Michaels J: **Administering a Democratic Industrial Policy.** *Harvard Law & Policy Review.* 2024; **18**(2): 279.

Keulemans S, Van de Walle S: **Cost-effectiveness, domestic favouritism and sustainability in public procurement: A comparative study of public preferences.** *Int. J. Public Sect. Manag.* 2017; **30**(4): 328–341.
Publisher Full Text

Kim G, Kumar T, Ramalho R, *et al.*: **Institutional Capacity for Policy Implementation: An Analytical Framework** *Policy Research Working Paper Series 11279.* The World Bank; 2026.

Kozuch AC, von Deimling C, Eßig M: **Implementing green public procurement: A replication study.** *J. Clean. Prod.* 2022; **377**: 134424.
Publisher Full Text

Langlois RN, Foss NJ: **Capabilities and Governance: The Rebirth of Production in the Theory of Economic Organization.** *Kyklos.* 1999; **52**(2): 201–218.
Publisher Full Text

LeRoux K, Wright NS: **Assessing performance and public value in the hollow state.** *Performance and Public Value in the 'Hollow State'.* Elgar; 2021a; pp. 1–24.

LeRoux K, Wright NS: **Reconceptualizing performance: Filling the hollow state with public value.** *Performance and Public Value in the 'Hollow State.* Elgar; 2021b; pp. 125–139.

Lerusse A, Van de Walle S: **Public managers' valuation of secondary policy objectives in public procurement – results from a discrete choice experiment.** *J. Behav. Public Adm.* 2021; **4**(1).
Publisher Full Text

Levi-Faur D: **The Global Diffusion of Regulatory Capitalism.** *Ann. Am. Acad. Pol. Soc. Sci.* 2005; **598**: 12–32.
Reference Source

Lindström H, Lundberg S, Marklund P-O: **Green public procurement: An empirical analysis of the uptake of organic food policy.** *J. Purch. Supply Manag.* 2022; **28**(3): 100752.
Publisher Full Text

Lundberg S, Bergman MA: **Tendering design when price and quality is uncertain.** *Int. J. Public Sect. Manag.* 2017; **30**(4): 310–327.
Publisher Full Text

Lundberg S, Marklund P-O: **Green public procurement as an environmental policy instrument: cost effectiveness.** *Environ. Econ.* 2013; **4**(4).
Reference Source

Lundberg S, Marklund P-O: **Green public procurement and multiple environmental objectives.** *Economia e Politica Industriale.* 2018; **45**(1): 37–53.
Publisher Full Text

Lundberg S, Marklund P-O, Strömbäck E: **Is Environmental Policy by Public Procurement Effective?.** *Public Finance Review.* 2016; **44**(4): 478–499.
Publisher Full Text

Maggetti M, Verhoest K: **Unexplored aspects of bureaucratic autonomy: a state of the field and ways forward.** *Int. Rev. Adm. Sci.* 2014; **80**(2): 239–256.
Publisher Full Text

Marron DB: **Buying Green: Government Procurement as an Instrument of Environmental Policy.** *Public Finance Review.* 1997; **25**(3): 285–305.
Publisher Full Text

Mayer KJ: **Construct validity and other empirical issues in transaction cost economics research.** Bergh DD, Ketchen DJ, editors. *Research Methodology in Strategy and Management.* Emerald Group Publishing Limited; 2009; Vol. **5**: pp. 213–236.
Publisher Full Text

Milward HB: **Symposium on the Hollow State: Capacity, Control, and Performance in Interorganizational Settings.** *J. Public Adm. Res. Theory.* 1996; **6**(2): 193–196.
Publisher Full Text

Nussbaum MC: *Creating capabilities: The human development approach.* Harvard University Press; 2011.

Organisation for Economic Co-operation and Development: **Government at a Glance.** 2021.
Publisher Full Text

Ortiz-Ospina E, Roser M: **Government Spending.** 2016. August 9, 2024.
Reference Source

Overman S: **Great Expectations of Public Service Delegation: A systematic review.** *Public Manag. Rev.* 2016; **18**(8): 1238–1262.
Publisher Full Text

Petersen OH, Hansen JR, Houlberg K: **The administrative burden of doing business with the government: Learning and compliance costs in Business-Government interactions.** *Public Adm.* 2022.
Publisher Full Text

Provan KG, Kenis P: **Modes of Network Governance: Structure, Management, and Effectiveness.** *J. Public Adm. Res. Theory.* 2008; **18**(2): 229–252.
Publisher Full Text

Potoski M, Lund-Sørensen B, Petersen OH: **Measuring Transaction Costs in Public Sector Contracting through Machine Learning and Contract Text.** *Public Adm. Rev.* 2025.

Raeymaeckers P, Cools P: **Negotiating performance: the strategic responses of associations where people in poverty raise their voice.** LeRoux K, Wright NS, editors. *Performance and Public Value in the 'Hollow State'.* Elgar; 2021; pp. 102–124.

Rietbergen MG, van Rheede A, Blok K: **The target-setting process in the CO2 Performance Ladder: does it lead to ambitious goals for carbon dioxide emission reduction?** *J. Clean. Prod.* 2015; **103**: 549–561.
Publisher Full Text

Rindfleisch A, Heide JB: **Transaction cost analysis: Past, present, and future applications.** *J. Mark.* 1997; **61**(4): 30–54.
Publisher Full Text

Rosell J: **Getting the green light on green public procurement: Macro and meso determinants.** *J. Clean. Prod.* 2021; **279**: 123710.
Publisher Full Text

Rubin E: **The Myth of Accountability and the Anti-Administrative Impulse.** *Mich. Law Rev.* 2005; **103**(8): 2073–2136.

Salamon LM, Lund MS: *Beyond privatization: the tools of government action.* Urban Institute; 1989.

Saussier S, Tirole J: **Strengthening the Efficiency of Public Procurement (Conseil d'Analyse Economique, n°22).** 2015.
Reference Source

Schillemans T, Busuioc M: **Predicting Public Sector Accountability: From Agency Drift to Forum Drift.** *J. Public Adm. Res. Theory.* 2015; **25**(1): 191–215.
Reference Source

Shadrina EV, Vinogradov DV, Kashin DV: **Implicit incentives in green public procurement: Good intentions versus rigid regulations.** *Ecol. Econ.* 2022; **198**: 107458.
Publisher Full Text

Smith J, Andersson G, Gourlay R, *et al.*: **Balancing competing policy demands: the case of sustainable public sector food procurement.** *J. Clean. Prod.* 2016; **112**: 249–256.
Publisher Full Text

Sönnichsen SD, Clement J: **Review of green and sustainable public procurement: Towards circular public procurement.** *J. Clean. Prod.* 2020; **245**: 118901.
Publisher Full Text

Spagnolo G: **Public Procurement as a Policy Tool.** Swedish Competition Authority, editor. *The Cost of Different Goals of Public Procurement.* 2012; pp. 24–43.

Spence DB: **Agency Policy Making and Political Control: Modeling Away the Delegation Problem.** *J. Public Adm. Res. Theory.* 1997; **7**(2): 199–219.
Publisher Full Text

Sundell A, Lapuente V: **Adam Smith or Machiavelli? Political incentives for contracting out local public services.** *Public Choice.* 2012; **153**(3/4): 469–485.
Publisher Full Text

Swedish Competition Authority: **The Cost of Different Goals of Public Procurement.** 2012.
Reference Source

Testa F, Grappio P, Gusmerotti NM, *et al.*: **Examining green public procurement using content analysis: existing difficulties for procurers and useful recommendations.** *Environ. Dev. Sustain.* 2016; **18**(1): 197–219.
Publisher Full Text

Testa F, Iraldo F, Frey M, *et al.*: **What factors influence the uptake of GPP (green public procurement) practices? New evidence from an Italian survey.** *Ecol. Econ.* 2012; **82**: 88–96.
Publisher Full Text

van Ginkel JR, Linting M, Rippe RCA, *et al.*: **Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data.** *J. Pers. Assess.* 2020; **102**(3): 297–308.
Publisher Full Text

van Thiel S, Yesilkagit K: **Does task matter? The effect of task on the establishment, autonomy and control of semiautonomous agencies.** *Int. Rev. Adm. Sci.* 2014; **80**(2): 318–340.
Publisher Full Text

Verordnung zur Statistik über die Vergabe öffentlicher Aufträge und Konzessionen (Vergabestatistikverordnung - VergStatVO): 2016.
Reference Source

Verschuere B, Verhoest K, Meyers F, *et al.*: **Accountability and Accountability Arrangements in Public Agencies.** Christensen T, Laegreid P, editors. *Autonomy and Regulation.* Elgar: 2006; pp. 268–298.
Publisher Full Text

Warsen R, Klijn EH, Koppenjan J: **Mix and Match: How Contractual and Relational Conditions Are Combined in Successful Public–Private Partnerships.** *J. Public Adm. Res. Theory.* 2019; **29**(3): 375–393.
Publisher Full Text

Williams MJ: **Beyond state capacity: bureaucratic performance, policy implementation and reform.** *J. Inst. Econ.* 2021; **17**(2): 339–357.
Publisher Full Text

Weber M: *Politik als Beruf.* Duncker & Humblot GmbH; 2010 [1921].

Williamson OE: *The economic institutions of capitalism: firms, markets, relational contracting.* Free Press; 1985.

Williamson OE: *The mechanisms of governance.* Oxford University; 1999.

Wilson W: **The Study of Administration.** *Political Science Quarterly.* 1887; **2**(2): 197–222.
Publisher Full Text

Windoffer A: **Verwaltungskontrolle.** *Verwaltungswissenschaft - Eine Einführung.* Universitätsverlag Potsdam; 2020.
Publisher Full Text

Wood BD: **Principals, Bureaucrats, and Responsiveness in Clean Air Enforcements.** *Am. Polit. Sci. Rev.* 1988; **82**(1): 213–234.
Publisher Full Text

Wright NS, Reames TG: **The role of community action agencies in facilitating successful sustainable development in American cities.** LeRoux K, Wright NS, editors. *Performance and Public Value in the 'Hollow State'.* Elgar; 2021; pp. 64–84.

Zhang L, Ding T, Fu Y: **Do measurement methods matter? A meta-analysis of antecedents and outcomes of contractual governance in interorganisational relationships.** *Prod. Plan. Control.* 2023; 1–17.
Publisher Full Text

Zhao T, LeRoux K: **Are outsourced public health services linked to better health outcomes? Findings from a national study of service contracting by local health departments.** LeRoux K, Wright NS, editors. *Performance and Public Value in the 'Hollow State'.* Elgar; 2021; pp. 43–63.

Zhu Q, Geng Y, Sarkis J: **Motivating green public procurement in China: An individual level perspective.** *J. Environ. Manag.* 2013; **126**: 85–95.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 18 August 2026

https://doi.org/10.21956/openreseurope.25237.r77504

© **2026 Aboelazm K.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Karem Aboelazm
United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates

The manuscript examines how delegation affects the administrative capacity of what the author conceptualizes as the “hollow state.” It develops the concept of **demand power**, defined as the state's capacity to formulate public demand toward contractors in pursuit of collective goals, and argues that contracting out and decentralization fragment this capacity. The empirical application concerns environmental sustainability in German public procurement and tests four preregistered hypotheses concerning asset specificity, measurability, transaction frequency, and political control.

The empirical scope is impressive. The study draws on administrative data covering public procurement procedures in Germany in 2021–2023 and uses a logistic generalized linear mixed model, supplemented by multiple imputation and extensive robustness analyses. The manuscript is theoretically ambitious, engages extensively with relevant literature, is preregistered, and provides unusually broad procurement coverage. Nevertheless, some conceptual and statistical claims should be moderated or clarified.

1. Is the work clearly and accurately presented and does it engage with the current literature? — **Yes**

The manuscript is generally well written, logically organized, and extensively situated within literature on the hollow state, delegation, transaction-cost theory, agency drift, administrative capacity, and sustainable public procurement. The theoretical contribution—the concept of “demand power”—is clearly introduced and connected to existing concepts of coordination capacity and administrative power.

The literature engagement is particularly strong because the author combines classic foundational works with recent research through 2025–2026. The manuscript also identifies two explicit gaps: the lack of state-level analysis of administrative capacity in hollow states and uncertainty regarding which forms of delegation are most consequential.

One improvement would be to reduce repetition between the Introduction, theoretical framework, Discussion, and Conclusion. The claims that contracting out fragments demand and that asset specificity is more consequential than political control are repeated frequently. A more concise presentation would sharpen the manuscript's central contribution.

2. Is the study design appropriate and is the work technically sound? — **Partly**

The observational design is appropriate for examining associations across a very large administrative dataset, and the preregistered hold-out strategy is a notable strength. The author appropriately explains that the 2021 data were used to refine the hypotheses and that the preregistered framework was subsequently tested using the 2022–2023 data.
However, the manuscript occasionally moves from observed associations to language implying that delegation *causes* or "stymies" administrative capacity. The author explicitly acknowledges that causal inference was not possible because the data are observational and important unobserved factors may remain. The causal terminology should therefore be moderated consistently in the Abstract, Discussion, and Conclusion.
A second concern is the conceptual identification of **contracting out**. All observations appear to be procurement procedures and therefore already represent contracting. Variation in asset specificity among procured products is not itself variation in whether government functions were contracted out. Consequently, the finding that asset specificity predicts sustainability criteria does not directly demonstrate that *contracting out relative to in-house provision* reduces demand power.
This distinction is important. The study convincingly demonstrates heterogeneity **within public contracting**, but a stronger causal claim about the consequences of contracting out as a form of delegation would ideally require comparison with equivalent activities performed in-house or another research design identifying variation in the degree of outsourcing. The theoretical interpretation should be revised accordingly.
3. Are sufficient details of methods and analysis provided to allow replication by others? — **Yes**
The methodological reporting is generally strong. The manuscript describes the outcome, focal predictors, political-control measure, controls, multilevel structure, missing-data treatment, exclusion of high-value observations, and robustness checks.
The preregistration and extensive Appendices A–N further strengthen reproducibility, and the appendices are deposited on Zenodo.
Nevertheless, two details should be clarified. First, the manuscript should provide a stronger substantive justification for excluding contracts above the **95th percentile of contract value**, particularly because the largest contracts may be theoretically important cases of asset specificity and administrative capacity. Second, the rationale for modeling federal state as the random-intercept grouping should be distinguished from possible clustering at the buyer-organization level, especially given that procurement procedures are nested within buyer organizations as well as federal states.
4. Are all the source data and materials underlying the results available? — **Partly**
The preregistration is publicly available through OSF, and the extensive analytical appendices are available through Zenodo. This represents a high level of research transparency.
However, the original administrative dataset cannot be openly shared because it was obtained from the German Federal Statistical Office under contractual restrictions. The manuscript appropriately explains these restrictions and describes how other researchers can formally request access to the underlying data.
I therefore select "Partly," but this is **not a criticism of the author** where legal/data-provider restrictions prohibit redistribution. If permitted by the data agreement, providing complete analysis code and a synthetic or structurally equivalent example dataset would further facilitate reproducibility.
5. Is the statistical analysis and its interpretation appropriate? — **Partly**
The logistic generalized linear mixed-model approach is broadly appropriate for a binary outcome, and the use of average marginal effects makes the substantive magnitude of the estimates easier to interpret. The manuscript also reports confidence intervals, conducts multiple imputation, performs complete-case and subgroup analyses, and uses alternative random-forest analyses as

robustness checks.
The principal results are clearly reported. For example, asset specificity has an AME of approximately −0.096, substantially larger in absolute magnitude than the estimated effects of measurability and political control.
However, several issues need attention.
First, variable importance should not be inferred primarily from z-scores. The manuscript states that asset specificity plays the most important role "cf. z score." A z-statistic depends on both estimated effect and standard error and is therefore not a direct measure of substantive variable importance. The comparison should instead emphasize appropriately standardized effects, changes in predicted probabilities, AMEs over substantively meaningful contrasts, or another clearly justified importance measure.
Second, the treatment of political control as a rank-ordered linear predictor deserves stronger justification. The manuscript itself reports that a robustness analysis suggests the effect may be disproportionately driven by political-control regime 5. This raises the possibility that the relationship is not linear across ranks 1–6. I recommend presenting the political-control categories as categorical predictors in the main or supplementary analysis and testing whether the assumed monotonic relationship is empirically justified.
Third, the missingness is substantial for several controls. Transaction frequency is missing for approximately 30%, discretion for 28%, total bids for 24.6%, and EU bids for 46.4% of observations. Multiple imputation is reasonable, but the plausibility of the Missing At Random assumption should be discussed more substantively given these levels of missingness.
6. Are the conclusions drawn adequately supported by the results? — **Partly**
The narrower conclusions concerning associations are well supported. The study provides strong evidence that asset specificity is negatively associated with the inclusion of environmental sustainability criteria and that measurability and political control are positively associated with such criteria. It also transparently reports the unexpected positive association between framework agreements and sustainability criteria.
The broader claim that **contracting out diminishes the administrative capacity of the hollow state**, however, is stronger than the empirical design directly demonstrates. Because the study analyzes procurement procedures rather than comparing contracted-out provision with in-house provision, it demonstrates variation in sustainability contracting across transaction characteristics rather than directly estimating the effect of contracting out itself.
Likewise, the outcome measures whether sustainability criteria are *included in procurement documents*, not whether these criteria are stringent, implemented during contract execution, produce environmentally superior procurement, or ultimately improve environmental outcomes.
The author appropriately recognizes these limitations. These qualifications should be brought more prominently into the Abstract, Discussion, and Conclusion.
Points that should be addressed for scientific soundness
The most important revision is to align the theoretical claims with what the empirical design identifies. The manuscript should distinguish clearly between evidence that **transaction characteristics within contracted procurement are associated with the state's use of sustainability criteria** and evidence that **contracting out itself reduces state administrative capacity**. The latter interpretation is currently not directly identified because there is no comparison with in-house provision.
The author should also reconsider using z-statistics to establish the relative importance of predictors, examine political control without imposing an untested linear rank effect, provide additional justification/sensitivity analysis for excluding the top 5% of contract values, and discuss the MAR assumption more thoroughly given substantial missingness in several variables.

Finally, claims concerning generalizability should be moderated. The dataset provides exceptional coverage **within Germany**, but population coverage within one national system does not by itself establish external validity across other hollow states. The manuscript itself appropriately acknowledges that institutional arrangements and the valuation of sustainability may differ across countries.
Overall assessment
This is a theoretically ambitious, data-rich, preregistered, and methodologically thoughtful study. The unusually comprehensive German procurement dataset and the effort to integrate transaction-cost theory with agency-drift theory make the manuscript potentially valuable to scholarship on public administration, procurement, and state capacity.
My main concern is not the quality of the dataset or the general analytical approach, but the **distance between the empirical estimand and the strongest theoretical claim**. The data strongly support conclusions about variation in the use of environmental criteria *within government contracting*, but they provide less direct evidence that contracting out itself causes a loss of administrative capacity relative to non-delegated/in-house provision.
I therefore recommend **revision**, with particular attention to moderating causal language, sharpening the interpretation of contracting out, strengthening the treatment of variable importance and political control, and aligning the Abstract and Conclusion more precisely with what the observational evidence establishes.

Is the work clearly and accurately presented and does it engage with the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

Are all the source data and materials underlying the results available?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Public Law, Administrative Law, Administrative Judiciary, Public Policy, and constitutional Law

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 17 August 2026

https://doi.org/10.21956/openreseurope.25237.r77505

© 2026 Domorenok E. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Ekaterina Domorenok
University of Padua, Padua, Italy

This is a promising paper that raises important questions. However, in my view, it would benefit from some reworking to address the following points and thereby strengthen both the argument and the analytical quality of the study.
First, if the analysis relies on the concept of administrative capacity, it should clearly define this concept and explain how it is operationalized in the environmental domain. How does administrative capacity manifest in the context of a “hollow state,” and why does it matter? Which specific capacities are relevant in this regard? Moreover, how do these capacities relate to the three types of delegation and to the expected policy outcomes?
Rather than offering a broad overview of studies on state capacity, the paper would benefit from a more focused and nuanced discussion of administrative capacity. This would help to better introduce the key analytical issues. Figure 1 should be revised accordingly to reflect this more precise conceptual framing.
Relatedly, the theoretical framework needs to be strengthened, particularly with respect to administrative capacity and delegation. This should be supported by engagement with more recent literature, especially in the field of green public procurement. Doing so would help clarify the conceptual and analytical contours of the study. Alternatively, if the author intends to use the broader concept of “state capacity,” this should be made explicit to avoid creating expectations regarding administrative capacity.
The hypotheses should be derived from a coherent and clearly articulated theoretical and analytical framework. In their current form, it is not entirely clear how they relate to the broader debate the paper seeks to address. The dependent and independent variables, along with their operationalization, should be clearly specified in the research design section, together with the corresponding hypotheses. At present, the research design section focuses more on the method than on the design itself.
The discussion section could also be improved and further refined analytically by directly addressing the general and specific research questions outlined at the beginning of the paper. In particular, what do the findings reveal about the administrative capacity of a hollow state in the environmental domain?

Finally, I suggest incorporating the limitations of the study—streamlined to leave more space for analysis—and outlining avenues for future research in the conclusion. This section should also include clear and concise final remarks that summarize the main findings of the study

Is the work clearly and accurately presented and does it engage with the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.