Journal of Economic Behavior and Organization 227 (2024) 106716

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

What are the priorities of bureaucrats? Evidence from conjoint experiments with procurement officials☆

Janne Tukiainen [d,*], Sebastian Blesse [e,f,h], Albrecht Bohne [f], Leonardo M. Giuffrida [g,c], Jan Jääskeläinen [a], Ari Luukinen [b], Antti Sieppi [b]

[a] *Aalto University, Finland*
[b] *FCCA, Finland*
[c] *MaCCI, Germany*
[d] *University of Turku, Rehtorinpellonkatu 3, 20014, Finland*
[e] *University of Leipzig, Institute for Public Finance and Public Management, Germany*
[f] *ZEW Mannheim, Germany*
[g] *ZEW Mannheim and CESifo, Germany*
[h] *Ifo Fürth and CESifo, Germany*

ARTICLE INFO

JEL classification:
D73
D90
H11
H57
H83
K41
M54

Keywords:
Bureaucrats
Public procurement
Conjoint experiments
Preferences

ABSTRACT

While effective bureaucracy is crucial for state capacity, its decision-making remains a black box. We elicit preferences of 900+ real-world public procurement officials in Finland and Germany. This is an important pursuit as they report having sizeable discretion and minimal extrinsic incentives. Through conjoint experiments, we identify the relative importance of multiple features of procurement outcomes. Officials prioritize avoiding unexpectedly high prices over seeking low prices. Avoiding winners with prior bad performance is the most important feature. Officials avoid very low competition, while litigation risks and regional favoritism matter less. Preferences and office interests appear well-aligned among bureaucrats.

☆ We thank Eduard Brüll, Alessandro De Chiara, Adriano De Leverano, Guido Friebel, Nicolas Fugger, Angelika Ganserer, Jonas Gathen, Dominik Hangartner, Friedrich Heinemann, David Hilpert, Ari Hyytinen, Elisabetta Iossa, Max Jahnsson, Philipp Lergetporer, Mika Maliranta, Felix Rösel, Emmi Silvo, Michal Soltés, Vítezslav Titl, and various participants at the Annual EconPol Meeting, the University of Frankfurt, ZEW Mannheim, the Walter Eucken Institute, the Helsinki GSE, IIPF 2021, EEA 2021, at VATT as well as the Finnish Competition and Consumer Authority (FCCA) for helpful comments and suggestions. We thank for the fruitful collaboration with the Deutsche Vergabenetzwerk (DVNW), especially with Jan Buchholz, as well as with DBInformation S.P.A. (Telemat Division). Moreover, our thanks go to Cobi King for proofreading the paper. Blesse is thankful for funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project ID 139943784—SFB 884. Blesse, Bohne and Giuffrida are thankful to the Leibniz Association for its support through the project Market Design by Public Authorities. Jääskeläinen and Tukiainen are thankful for funding by the Yrjö Jahnsson foundation. Tukiainen is thankful to the Academy of Finland project funding 340045 and ERC-2021-COG (INTRAPOL-101045239). All authors are thankful for EconPol funding. The project is registered pre-trial in the AEA RCT Registry under number AEARCTR-0006401 (link: https://doi.org/10.1257/rct.6401-1.0).

* Corresponding author at: University of Turku, Rehtorinpellonkatu 3, 20014, Finland.
E-mail address: janne.tukiainen@utu.fi (J. Tukiainen).

https://doi.org/10.1016/j.jebo.2024.106716
Received 10 December 2023; Received in revised form 26 August 2024; Accepted 28 August 2024
Available online 5 October 2024

1. Introduction

The functioning of public bureaucracies is considered a principal driver of state capacity, an observation going back to Weber (1921). Arguably, an important input into bureaucracies are individual bureaucrats and their behavior. However, this insight has received little attention in economics as a stand-alone aspect of government effectiveness (Besley et al., 2022). Contrary to conventional wisdom, public servants are generally able to exercise a sizeable degree of autonomy in their day-to-day work. Recent evidence seems to confirm this, pointing to the benefits of granting these workers more discretion in countries with high public capacity (Bosio et al., 2022; Rasul et al., 2020; Bandiera et al., 2021; Rasul and Rogger, 2018; Duflo et al., 2018).[1]

Yet we know little about how bureaucrats make choices. In a principal–agent setup, bureaucrats may work as agents for the general public (Alesina and Tabellini, 2008). While in the private sector such a setting would be governed by strong extrinsic incentives, in the public sector such incentives are usually weak or draw on non-standard mechanisms (Burgess and Ratto, 2003; Prendergast, 2007; Bertrand et al., 2019; Khan et al., 2019). Existing evidence rather points to substantial intrinsic motivation of public sector employees (Friebel et al., 2019; Ash and MacLeod, 2015). The underlying preferences of bureaucrats are therefore a key element driving their decision-making process.

Recent research considers management practices as input to drive outcomes in private and public organizations (e.g. Bloom et al., 2015). In contrast, our paper focuses on the *preferences* of front-line employees. We focus on public procurement, an ideal laboratory to study bureaucratic decision-making more generally due to its cross-sectoral and cross-institutional nature and its substantial relevance.[2] Specifically, we elicit in surveys *how* procurement officials (POs) think that *desirable tender outcomes* can be achieved in their daily decision-making processes and which *tender attributes* they find beneficial.

POs face complex tradeoffs when making choices, with multiple objectives and constraints imposed by the institutional environment (e.g., the regulation regarding the contract and auction design), as well as idiosyncratic variation in knowledge, resources, and behavioral biases. Moreover, POs' decisions do affect the outcome of a procurement, but typically only indirectly. Observational data typically only show aggregate realizations of choices among different actors and mask POs' true preferences. In particular, tender outcomes are influenced by the choices of both the POs and the bidders.[3] Studying the "anatomy" of bureaucratic decision-making in procurement is therefore hard through observational data. Surveys and survey experiments may help researchers to uncover these otherwise intangible information directly (Stantcheva, 2023).

To provide causal insights into POs' preferences uninfluenced by real-world constraints, we design and field a so-called conjoint experiment (e.g., Stantcheva, 2023; Hainmueller et al., 2014; Bansak et al., 2019), which is essentially a type of stated choice experiment (see more details on the method below). Specifically, we make real-life POs repeatedly choose between pairs of hypothetical tender outcome scenarios that randomly differ across six attributes: the price in the selected bid, the quality of the good or service (as promised in the selected bid), the degree of competition (the number of placed bids), the reputation of the winning firm by its past performance, the presence of judicial litigation, and whether the chosen firm is from the same region as the public buyer.

These attributes of procurement projects have the following useful features for our research endeavor. First, they are more relevant for the general public than direct choices about, e.g., tender procedures and auction mechanisms. Second, these attributes are deemed the most relevant to procurement outcomes from practice and research. Third, they include both predictors (e.g., bid price) and realizations (e.g., number of bids) of the awarding process, both dimensions relevant to the outcomes. Considering both dimensions is indeed one of the benefits of our conjoint approach (Stantcheva, 2023). Fourth, the choice framing considering the end of the tender stage—that is, just after the contract has been awarded to the selected firm but before the goods or services have been provided—allows us to analyze critical criteria of the entire tender process (i.e., number of bids received and judicial complaints) while also considering key aspects of the winning bid (i.e., price and quality) and the selected winner (i.e., past performance and proximity) for heterogeneous procurements (i.e., with and without an execution stage). Moreover, this timing is very realistic and familiar for the respondents, as the end of the tender process is a natural stage to reflect upon the performance in their daily work. While the respondents' daily work may not involve a direct choice over some attributes (e.g., the extent of competition), they typically make choices that can influence these particular attributes (e.g., by being engaged in marketing & contract planning).

Conjoint experiments originated in marketing, were formalized in political science (Hainmueller et al., 2014), and recently adopted in economics.[4] For reviews on conjoints with their pros and cons, please see Stantcheva (2023) and Bansak et al. (2019).

[1] There are also studies that point to the negative effects of discretion (e.g., Boland and Godsell, 2021). More generally, the trade-off between rules and discretion for bureaucrats can be seen as a principal retaining more or less control over the agent's decisions depending on the congruence of objectives, i.e., via delegation of authority, within organizations (Aghion and Tirole, 1997).

[2] Procurement is amongst the largest sectors of government activity and with some of the most frequent and high-stakes decisions made in the public sector. For 2018, the expenditures for public procurement of the median OECD country amount to 13% of GDP and 41% of total government spending (OECD, 2019).

[3] Observable outcomes such as bid prices, number of bidders, or product characteristics are the joint result of choices by POs and bidders. First, POs choose, for example, the auction rules out of the set that regulation allows, and what is purchased. Second, bidders decide how to bid given their competitive and strategic environments. This toy example implies that procurement outcomes do not only reflect bureaucratic preferences as POs, despite substantial influence, cannot entirely choose them by themselves. For example, award prices may be high not because POs prefer high prices but because competition is low, there is bid rigging among suppliers, POs are constrained from choosing optimal auction rules for a specific tender, or make mistakes in specifying the object or the rules. While structural models of observable auction data allow the estimation of primitives such as bidder's production or entry costs (Gentry et al., 2018), they would not help analyze preferences over tender outcomes.

[4] Among others, conjoints study preferences about migration (Bansak et al., 2016; Hainmueller and Hopkins, 2015) and public policies (Bechtel et al., 2014; Bechtel and Scheve, 2013; Beetsma et al., 2022). Stantcheva (2023) provides a lengthy list of papers in economics which adopted conjoints and vignette designs. Bansak et al. (2019) provide a similar list for conjoint adaptations in political science.

Conjoints typically let respondents evaluate multiple hypothetical scenarios which are randomized across various attributes of interest. Compared to text-based vignette studies, they are often presented in tabular forms (Stantcheva, 2023). In our given context, the strength of the conjoint design (when compared to a typical experiment with a binary treatment) is that we can study the causal effects of multiple tender attributes on the tender evaluations of public buyers, a feature which is well-suited for the complex decisions involved in setting up and awarding public tenders. Moreover, and in contrast to a revealed preferences approach, we can decompose and quantify the relative importance of single tender attributes for POs in their multi-faceted choice framework.[5] The results, in turn, allow for the clean study of preferences underlying bureaucratic choices without agency or supply-side frictions. Conjoints also have the merit of being validated and shown to predict real behavior (Hainmueller et al., 2015), to alleviate concerns about experimenter demand and increase realism because subjects simultaneously trade-off various relevant factors (of the tendering process, in our case). As is typical for conjoints, we do not provide monetary incentives as we do not want to encourage any particular choices over others. Also, POs in our context do not have distinct pecuniary incentives for specific tender outcomes in their daily work (on incentives, see Section 4.3). However, we motivate serious and truthful responses by encouraging POs that their participation may make their "voice" visible in public discussion.

Conjoint experiments are very close to traditional stated preference experiments (e.g. Rakotonarivo et al., 2016; Ameriks et al., 2020; Fuster and Zafar, 2021). However, they differ in two major dimensions. First, they do not impose any specific modeling assumptions regarding the choice structure or the underlying utility function of the respondents which is a key advantage (Bansak et al., 2019). Being agnostic about these assumptions is important since we provide first-time evidence on the priorities of real-life public buyers. Second, conjoints are very tractable, easy to implement, and provide a clearly interpretable estimate measuring the marginal impact of each individual attribute in driving stated choices. This estimate is the average marginal component effect (AMCE), first formalized in the potential outcomes framework by Hainmueller et al. (2014), and later refined (Leeper et al., 2020; Bansak et al., 2021, 2018). Intuitively, this approach estimates how a given level of an attribute drives observed choices compared with an omitted base category of the same attribute—e.g., comparing a higher-than-expected price to the expected benchmark price of the PO. The key for identification is a full randomization of all attribute levels across choice scenarios and respondents.

We field our conjoint experiment on POs in Finland and Germany allowing for cross-country comparisons between two settings with high public sector capacity yet with similar (i.e., EU) procurement rules.[6] Investigating bureaucratic priorities is very relevant in our context since POs in both countries report low extrinsic incentives and sizeable discretion, but view rigid regulation as a major obstacle to desirable work outcomes.

We contacted about 8,500 real-life POs and received a total of 933 completed answers from both Finland and Germany. Notably, the Finnish sample is broadly representative based on detailed administrative data. While we cannot analyze whether the German sample is representative or not, the respective results from it are interesting as Germany has the largest economy and public procurement market in the EU. We also elicited socio-demographics, tasks and workplace features, procurement practices, and typical outcomes of the POs and their offices. At the end of the survey, we analyze how officials' preferences and career incentives (i.e., perceived office interests) are aligned concerning tender outcomes.

We produce several interesting findings. First, for both the price and supplier reputation, the size of the effect in case of negative realizations (or "threats" to the organization) is considerably stronger than for positive realizations (or "opportunities", respectively). In other words, avoiding threats along these dimensions is a much stronger driver of the observed choices than possible opportunities for the respective organization. This effect may be explained by the desire to avoid criticism prompting otherwise public-spirited bureaucrats to behave inefficiently (Leaver, 2009). Such a *"prevention focus"* of bureaucrats has been also previously discussed by public administration scholars (e.g., Jackson and Dutton, 1988 and more recently, Kuehnhanss et al., 2017). The latter argue that bureaucrats, in particular those with strong mission-based views and aligned personal- and office-level interests, focus on avoiding negative outcomes for their office. Second, our baseline results show distinct and very robust priorities in POs' choices regarding the randomized attributes. Winners' low reputation, in particular due to previous negative interactions, is valued extremely negatively by respondents. Likewise, unexpectedly high prices are seen as an extremely unfavorable purchase condition, but the effect is only about two-thirds as adverse as a winner with a bad reputation. The results concerning competition show that POs have non-linear preferences towards additional bid submissions with a flattening curve after about four offers. We attribute this finding to the trade-off of costs and benefits of competition in the eyes of POs. We also show that litigation is valued negatively, but the related effects are fairly small. Lastly, whether or not the firm is located in the same region as the PO (i.e., to buy local) does not have a strong influence on the stated choices of public buyers.

[5] A huge benefit of conjoints compared to merely inquiring bureaucrats about their work priorities is that we can uncover latent or implicit preferences (Cunningham and de Quidt, 2015). POs may mix up their preference over specific outcomes in complex procurement decisions when being asked directly. For instance, POs may have implicit associations over the promised quality of the bid, trust in the supplier, and their regionality. Asking POs directly about their preferences over these attributes may then lead to false inference. Specifically, POs state that their strongest preference is towards ensuring high-quality purchases (see Figure A.1 in Appendix A)) but our choice experiment shows that—on ceteris paribus grounds, e.g., holding the difference between price and quality with PO's expectations constant—trust in the supplier is the most preferred attribute such that POs have an explicit preference for quality but an implicit preference for trust in suppliers. There are other benefits to conducting conjoint experiments. Stantcheva (2023) mentions high statistical power since individual respondents decide over multiple scenarios and address causal effects on choices in realistic, but yet hypothetical scenarios.

[6] The public procurement market is relatively larger in Germany and Finland (on average 17% of GDP and 33% of government spending) than the US (10% and 28% respectively) (OECD, 2019). Both countries share good regulatory practice, average levels of discretion for officers and high process efficiency and quality (Bosio et al., 2022). For both cases motivation to work is similarly higher in the public sector as compared to the private sector (Cowley and Smith, 2014).

Our paper engages in an in-depth discussion of the potential underlying factors driving the observed choice behavior. The headline finding concerns the striking absence of heterogeneities in reported choices based on all observables, including country, socio-demographics, task-, and office-level features. We examine various potential explanations for this lack of preference heterogeneity. First, we find no evidence that variation in external incentive schemes in the form of career concerns significantly influence choice behavior. Second, variations in office size, which might approximate the intensity of office-level norms, show no correlation with elicited priorities. Third, we observe a strong alignment between bureaucrats' preferences and the perceived interests of their offices, suggesting efficient yet conservative bureaucracies (Besley and Ghatak, 2005). Although strong mission-based motivation among POs provides a plausible explanation for homogeneous bureaucratic preferences, our data cannot conclusively determine the underlying mechanisms. Specifically, our one-shot survey design on public sector workers cannot determine whether bureaucratic preferences are innate or shaped by institutional incentives. As a result, we avoid drawing firm conclusions about the origin of these preferences.

Our results have several policy implications. First, more focus may be given to lower prices in the procurement process. Second, the procurement rules could be altered to better incorporate past performance criteria into the selection framework. Third, the observed alignment of priorities between bureaucrats and offices speak in favor of granting POs more discretion.

Related literature. Our paper contributes to various distinct strands of the literature. One broad literature we contribute to is the personnel economics of the state (Finan et al., 2017; Besley et al., 2022), which investigates how the public sector's organizational and design features interact with individual bureaucrats.[7] A different approach focuses on the underlying fundamentals of individual bureaucrats in contrast to private sector workers.[8] Our paper concisely measures bureaucrats' preferences regarding the most relevant and frequent choices in their work. By digging into these preferences, we gauge how bureaucrats trade-off risks and benefits attached to a given work objective and quantify the relative importance of multiple objectives in their work environment in an experimental set-up. For instance, we show a *"prevention focus"* of public buyers. Moreover, we show that personal bureaucratic preferences and office interests are well aligned in our setup. Contrary to the previous contributions mostly targeting developing settings, we focus on two countries with high state capacity.

We also contribute to the economics of public procurement. First, we speak to studies about competition in this area which typically report a very low number of bids (e.g., Jääskeläinen and Tukiainen, 2019; Fazekas, 2017). The latter are typically associated with higher prices (e.g. Hyytinen et al., 2018). Recent studies relate low competition levels to procurement design, and draw conclusions regarding procurers' tendency to improve supplier selection or reduce administrative costs by raising entry barriers (Kang and Miller, 2022; Coviello et al., 2018a). Since our conjoint controls for key bid features (e.g., reputation, quality, price), we can isolate public procurers' unconfounded preferences for competitive tenders. Our findings suggest that POs' preferences *per se* are unlikely to be the main driver for low competitiveness. Second, we find that POs would welcome relational contracts through the possibility of rewarding or punishing bidders during the selection process based on past performance (Decarolis et al., 2016; Butler et al., 2020), but such policies are often not implemented for transparency and competition concerns.[9] Third, we add to papers on the spatial location of public buyers and their suppliers. Empirically, procurement contracts are often allocated to firms in the proximity of the contracting authority (Jääskeläinen and Tukiainen, 2019; Kutlina-Dimitrova and Lakatos, 2016). To some degree, this reflects natural entry costs, and explanations include stated policy goals of "buy local", political connections (Szucs, 2024; Baltrunaite, 2019; Baltrunaite et al., 2021; Ryan, 2020; Baránek and Titl, 2024), and prior relationships (Hyytinen et al., 2018). We provide direct evidence that public buyers do not have strong preferences to buy locally when controlling for key bid features such as bid price, quality, and supplier reputation. Lastly, we contribute to recent studies on the interactions of procurement activity and potential judicial complaints raised by suppliers (Coviello et al., 2018b), including policy work on how procurers adopt practices to reduce litigation exposure risks (Halonen and Tukiainen, 2020). Our results suggest that bureaucrats dislike litigations, but only to a small extent.

Lastly, our paper contributes to studies of preferences using survey experiments (see Haaland et al., 2021 for a review). Specifically, we employ conjoint experiments that simultaneously randomize hypothetical choice scenarios and various choice attributes. We employ this method to unravel distinct choice aspects of bureaucrats and thereby elicit their latent preferences. By studying the preferences of public buyers, we further relate to papers on survey data using special samples such as academics (Andre et al., 2022; DellaVigna and Pope, 2018a,b), politicians (e.g. Broockman and Skovron, 2018), economic experts (Gründler and Potrafke, 2020) and firm managers (Coibion et al., 2018; Link et al., 2023). We are the first to conduct choice experiments regarding the anatomy of desired work outcomes among procurement bureaucrats.[10]

[7] Topics include explicit incentive schemes (Bertrand et al., 2019; Khan et al., 2019; Burgess and Ratto, 2003; Prendergast, 2007), the optimal level of discretion for bureaucrats (Bosio et al., 2022; Bandiera et al., 2021; Rasul and Rogger, 2018; Duflo et al., 2018), and determinants of the bureaucratic selection process (Ashraf et al., 2020; Hanna and Wang, 2017; Xu, 2018; Colonnelli et al., 2020).

[8] Studies investigate, among others, intrinsic motivation (Besley and Ghatak, 2005; Prendergast, 2007; Friebel et al., 2019), risk aversion (Tepe and Prokop, 2018; Buurman et al., 2012), self-esteem (Ellingsen and Johannesson, 2008) and pro-sociality (Gregg et al., 2011).

[9] This finding stresses how relational contracts based on trust are perceived as efficient governance instruments for non-contractible dimensions *of quality* (see Levin, 2003; Andrews and Barron, 2016; Board, 2011) relating to the selected supplier rather than the specific tender (Calzolari and Spagnolo, 2009; Albano et al., 2023).

[10] Detkova et al. (2021a,b) study corruption during the covid pandemic and across genders, respectively, while employing list experiments among public procurement agents in Russia. A few other papers run choice experiments with public servants. For example, Kuehnhanss et al. (2017) analyze the prevention focus of civil servants of the European Parliament when policy proposals with distinct benefits and disadvantages are available. Meyer-Sahling et al. (2021) study differences in public sector motivation through different personal management practices through conjoint experiments and Oliveros and Schuster (2018) provide evidence on how bureaucratic structures affect work motivation, political services and corruption using the same method in the Dominican Republic.

2. Survey design and conjoint experiments

2.1. Sampling strategy

To study how bureaucrats form decisions about desirable tender outcomes in high-capacity contexts, where bureaucratic discretion is associated with better decisions (Bosio et al., 2022), we fielded large-scale surveys to real-life POs in Finland and Germany.

We deem such a cross-country analysis relevant for this study since nation-specific differences (e.g., culture, institutions, and population density—mirroring, e.g., in public bureaus of different sizes as stressed in Section 2.4) exist. The two countries are comparable for the bureaucratic decision-making process since as they are generally subject to the same procurement regulation, as both are part of the EU, which limits unobserved institutional features and allows for a meaningful comparison. We fielded an online survey between September and November 2020 using a commercial survey platform provider called Alchemer.[11] We conceptualized the survey in English and subsequently translated it to the respective native languages.[12] The survey was administered via email invitations to POs with unique web links to enter the survey platform. Bureaucrats' participation was entirely voluntary and was not incentivized in any way. In order to increase response rates, we sent two reminder emails. The survey had a median duration of about 15 min. For Finland, we draw on the universe of contracting authority emails registered in the administrative "Hilma" database, the official and mandatory online public procurement platform in Finland (see Jääskeläinen and Tukiainen, 2019 for more details). The survey invitations were sent directly by our implementation partner, the Finnish Competition and Consumer Authority (FCCA). Altogether, we contacted 1,353 POs and received a response rate of 29.8% (403) in Finland. To contact public procurers in Germany, we collaborated with the *Deutsches Vergabenetzwerk* (DVNW), a leading web platform for public procurers and associated experts, designed to exchange information and interact on all topics regarding procurement regulation, law, and related news. Overall, we invited 7247 POs through DVNW in Germany and received 530 completed responses (7.3% response rate). While the Finnish sample targets the universe of all unique PO addresses and we can show that it is indeed broadly representative for all Finnish public procurement regarding several office-level features such as office type, typical contract size, procured industries, regions, typical award procedures (for more details on non-response, see Appendix B), we do not know whether the German sample is representative of all German public buyers.[13] This is due to the fact that, to the best of our knowledge, there is no data available on public contracting personnel in Germany to benchmark the background answers.[14]

However, in both Finland and Germany, we observe a highly heterogeneous sample of bureaucrats in terms of demographic and educational background, office characteristics, and typically performed tasks. Moreover, our results themselves are extremely similar, not only across both countries (see Section 3) but also across all our demographic, office-, and job-related observables (Section 4). These observations can, however, ultimately not fully mitigate potential worries regarding sample selection among German procurers; nevertheless, and perhaps more importantly, both our very heterogeneous samples of individual bureaucrats and the highly similar results across these sample characteristics speak strongly in favor of the external validity of our findings. This applies primarily to public procurement practices in similar high-capacity settings. However, since public procurement is an omnipresent feature associated with all governmental tiers and activities and our respondents span all of these areas, we argue that our results have some degree of applicability to bureaucratic behavior in general beyond public procurement.

2.2. Survey structure

This section provides a brief description of the survey structure. Participants enter the survey through the individual web-link and are shown a starting page (see Figure F.1 in Appendix F for a screenshot) which introduces the academic partners involved and explains the general purpose of the survey, namely identifying potential improvement possibilities in the respective national public procurement policies from the viewpoint of POs. We assure POs anonymity. Survey participation is entirely voluntary and is dealt with confidentiality. We also appeal to answer honestly and state that the web link can only be used once. Moreover, subjects are asked whether they agree to these conditions and wish to take part in the survey. Upon agreeing, POs then enter our survey.

The survey comprises three distinct blocs of questions (*Bloc A–C*). Bloc A specifically queries respondents about their individual background characteristics concerning their socio-demographics and place of work (*Bloc A.1*), their work environment and task structure (*Bloc A.2*), as well as procurement practices (*Bloc A.3*). Bloc A enables us to study the sample composition and estimate the influence of various individual traits on PO choices from the conjoint experiments. Bloc B proceeds with a brief attention check followed by our main conjoint experiments eliciting POs' preferences (for a detailed account, see the next section). Finally, bloc C presents some short follow-up questions about the role of our treatment attributes. A detailed account of the survey structure and the exact wording of questions is provided in Appendix E and F, respectively.

[11] To test our survey modules, we fielded a pilot for a small sample of Finnish POs in June 2020. The pilot contacted 100 POs working in 36 contracting authorities and received twenty-one responses. We randomly sampled these contracting authorities from the universe of offices with 2–5 contact emails in the Hilma database. We do not use these responses or their offices in the main analysis. The pilot led to some changes in the follow-up questions after the conjoint experiments.

[12] The native languages include German, Finnish, Swedish (for some regions in Finland), and Italian. The translations were done by the authors who are native speakers using a four eye-principle. Moreover, we checked the translations with national procurement experts to make sure they fit the institutional contexts. Here we acknowledge help from Tim Bauckloh, Jan Buchholz, Max Jahnsson, and Emmi Silvo.

[13] We originally fielded the survey also in Italy, unfortunately with very few responses. The response rate was extremely low, potentially due to a distribution via a commercial provider. Appendix C provides details and shows that the qualitative results are nonetheless quite similar to our main findings.

[14] For more details on sampling, see Appendix D.

2.3. The conjoint experiments

Set-up. Bloc B includes our conjoint experiments consisting of repeated choices between pairs of fictitious tender outcome scenarios which are randomized along key attributes. In the design and framing of the experiment, our guiding motivation was to generate responses as comparable as possible given the large differences in the institutions, types of goods or services purchased and task structures of the respondents. We present the decision scenario as a situation in which the contract has just been awarded. Specifically, all bids have been placed and compared against one another, the winner has been chosen, and losing bidders have had the time to litigate if they wished. However, the actual good or service has not yet been delivered, and we ask respondents to evaluate the tender outcomes based on the information contained in the selected bid (for the introductory text, see Figure F.3 in Appendix F). We refer the reader to Appendix G for a visual representation of the timing of our experiment along a schematic timeline of the procurement process. This timing prompts respondents to reflect retrospectively on the very outcome of their work (i.e., the contract award) along the various completed steps of the procurement process before the successful bidder (and sometimes other oversight officials) steps in and the execution phase begins.

We chose this decision scenario because the situation is relatable and familiar to respondents, relevant for all types of public procurement, and comparable across different job tasks. This makes the elicited choices realistic and relevant for as many procurers as possible. Asking to evaluate post-construction or delivery contract outcomes would have been challenging as the execution stage varies greatly across different types of goods and services. For example, some goods are delivered immediately and the realizations of price and quality are observed almost simultaneously. Other goods, like construction works, may take years to complete, and the actual quality is revealed even later, if at all. Alternatively, we could have designed the conjoint experiments about choices regarding tender procedures and auction mechanisms. However, since the procedures and mechanisms vary greatly between settings and are more of a means to an end, we believe it is more useful to examine preferences about tender outcomes directly.

Our set–up mimics the day-to-day choices that the respondents face in their work fairly well, which is key in conjoints (Stantcheva, 2023). Moreover, as multiple attributes need to be evaluated and it is thus difficult to respond strategically, conjoints have been shown to reduce social desirability bias (Hainmueller et al., 2014; Horiuchi et al., 2020).

Design. The instructions explicitly convey that respondents have to make six repeated choices, each between two hypothetical tender outcome scenarios (also referred to as cards or profiles).[15] The instructions also clearly state that respondents should think about these scenarios in the context of purchases they are usually involved with in their professional life and each scenario represents a new purchase. They are also instructed to choose tender outcomes they would personally prefer in their function as a procurement employee (see Figure F.3 in Appendix F for details). After the instructions, subjects are exposed consecutively to each of the six pairs of tender outcome scenarios. Participants have to make individual choices about which tender outcome scenario they would prefer, respectively, and were also forced to choose between them.[16] Thus, each respective participant evaluated a total of 12 tender outcome scenarios. Each scenario differs along six attribute dimensions. Fig. 1 gives an example of one possible pair of outcome scenarios (i.e., a pair of cards or profiles) that a respondent chooses between.[17] Each decision scenario of paired hypothetical tender outcomes was placed on a separate screen. First, we fully randomized all attribute levels for each of the six attributes by which tender outcome scenarios differ. Second, we also randomized the order in which we present the attributes in order to avoid primacy or recency effects but kept the ordering unchanged for a given respondent to make responding easier.

Choice of attributes and expected effects. Table 1 reports all the possible attribute realizations (levels) relevant to our choice experiments. The attributes we include are the following: price (5 levels), a generic quality indicator (3 levels), the amount of competition (4 levels), quality arising from the type of winner in terms of familiarity (3 levels), whether the tender was litigated (2 levels), and whether the winner is a local firm (2 levels). Experienced practitioners (i.e., lawyers and public procurers), academic papers and policy work (see below for details) which we consulted prior to the study, acknowledge these dimensions as relevant tender outcomes. Moreover, the bureaucrats themselves also stated in our pilot study (see footnote 11), that these attributes account for and capture all the most relevant features of tender outcomes.

The choice of attributes and their realizations are well-founded in the public procurement literature. Since value for taxpayers' money is the primary procurement objective, both price and quality are, in fact, relevant bidding dimensions (Asker and Cantillon, 2010). We expect that a higher price and quality as stated in the winning bid are ceteris paribus negatively and positively related to individual support of bureaucrats, respectively. Moreover, from the perspective of social optimum, lower prices and higher quality should be desired, and more specifically, higher prices should be avoided with the same magnitude as lower prices desired given that the taxpayers' utility for money is linear.

The hypothesis regarding preferences for competition is more complex. While POs should prefer an unconditional increase in competition for its instrumental role in lowering prices and improving quality, the preference in our experimental design is

[15] The number of choice tasks was chosen conservatively to not cause respondent fatigue and erroneous answer behavior due to an overwhelming number of tasks. Bansak et al. (2018) and Bansak et al. (2019) show that for standard online survey platforms, response quality does not deteriorate for high numbers of choice tasks and can well be justified for up to 15 tasks. In line with this, we have only little attrition in the choice tasks (see Appendix D).

[16] Specifically, respondents were asked on each profile to carefully look at the presented pair of hypothetical tender outcome scenarios and then decide which outcome scenario they prefer in each round, respectively. See Fig. 1 for the exact question and a graphical illustration of an exemplary card.

[17] Including six attributes in the choice experiments is well in line with the recommendations from Bansak et al. (2019) in order not to overwhelm participants. We also chose the table format of the experiment according to Hainmueller et al. (2015), because it appears to outperform the text format of the choice experiment, i.e., so called vignettes. Also pairwise comparisons represent real-world choice better than decisions which are solely based on a single hypothetical scenario (Hainmueller et al., 2015).

Table 1
Tender outcome attributes and levels for the conjoint experiments.

Conjoint Experiments		
Attributes	Attribute Explanation	Possible Attribute Realizations (Levels)
Price	The price as stated in the winning bid is	much lower than I expected; a bit lower than I expected; what I expected; a bit higher than I expected; much higher than I expected
Quality	The quality of the purchase as promised in the winning bid is	as I expected; a bit better than I expected; much better than I expected
Number of bids	The tender received	1 bid; 2 bids; 4 bids; 8 bids
Familiarity	The selected winner is	a firm that was unknown to me through previous tenders; a firm I already knew from previous tenders and trusted; a firm that I already had a bad experience with
Regionality	The selected winner is	a local bidder from your region; a non-local bidder that does not come from your region
Legal Complaints	After awarding the contract, … legal complaint has been filed against the tender.	· … · no · … · [weighted probabilities of 90%]; · … · a · … · [weighted probabilities of 10%]

Notes: This table shows the attributes and their potential realizations (levels) used in the conjoint experiments.

conditional on these attributes. Nevertheless, our approach to framing price and quality attributes in terms of buyers' expectations may introduce some residual uncertainty about the accuracy of these expectations. As competition intensifies, buyers become more certain they have received the true minimizing cost-quality bid. In other words, POs may conditionally prefer more competition because it is likely to narrow the gap to the most cost-effective or quality-optimized bid, thereby reducing uncertainty in their eyes. Furthermore, policymakers typically view competition as a method of increasing process efficiency, and thus, a high number of bids as an independent indicator of healthy procurement markets (e.g., Fazekas, 2017). Consequently, we expect that public buyers internalize this preference and favor increased conditional competition also along this dimension. On the other hand, more competition also means a higher administrative burden due to the need to evaluate and compare additional bids, suggesting a potential trade-off (Coviello et al., 2018a). Thus, we hypothesize that POs have inverted U-shaped preferences concerning conditional competition, as costs tend to outweigh benefits at some point. Instead, a social planner would be less concerned with these administrative costs, as they are generally small compared to the potential welfare gains from increased competition. Thus, while the social planner's preferences may also follow an inverse U-shape, the diminishing value of competition likely necessitates more bids than those indicated within our experiment, suggesting that more is better for the social planner.

The reputational forces linking future business to past performance (i.e. relational contracts) are key pillars of any economic sector (e.g., Levin, 2003). Because auctions can be a problematic mechanism in the context of procurement with incomplete contracts (Calzolari and Spagnolo, 2009), past performance indicators are widely employed for supplier selection in private procurement. More generally, private sector relies on business relationships and partnership sourcing and base their transactions more on mutual trust and information sharing (Stentoft Arlbjørn and Vagn Freytag, 2012). Public regulations, instead, restrict buyers' leeway to exclude suppliers with a poor track record to reduce room for favoritism and corruption. Nonetheless, the use of reputational metrics is found to improve performances (Butler et al., 2020; Decarolis et al., 2016). Therefore, we conjecture that public buyers consider reputation based on past performance as a non-contractible quality dimension relating to the supplier rather than the specific tender. Thus, they have a preference for suppliers they have worked with and can trust and this can be seen as a component of the quality realization. Preference for reliable supplier is likely to be in the interest of both public officials and the social planner.

Moreover, interviews conducted with procurers highlight how litigation exposure is a major concern (Halonen and Tukiainen, 2020), although they are empirically not supposed to introduce much wastefulness if the judicial system is well-functioning overall (Decarolis et al., 2020b)—as in our setting. We want to test for POs' perception of the litigation issue and expect that public buyers favor procurement outcomes without realized litigation. Given that the threat of litigation is one way to hold public officials accountable, it is also sensible from a welfare point that POs dislike litigation.

Finally, lawmakers usually set regulations favoring local players in the national procurement market, mostly due to small business concerns. Moreover, there could also be political pressure to favor local suppliers. We thus have an interest in whether POs perceive the winner's geographical proximity (i.e., to buy local) as a relevant outcome when conditioning on other tender features, in particular quality and buyer–supplier relations. We expect that public buyers have a preference for local suppliers in their vicinity. By contrast, it would be socially optimal that local suppliers are not favored.

Although we can theoretically form expectations on the respective signs of the estimates for each attribute to the outcomes of the tender process, their related relative magnitudes are to be tested empirically. Exactly herein lies the core strength of the conjoint experiment method, namely, to be able to test the relative preference of respondents for each of these individual attributes. Thus, our empirical setting allows us to estimate the priorities of public buyers for these procurement outcome features in a causal manner.

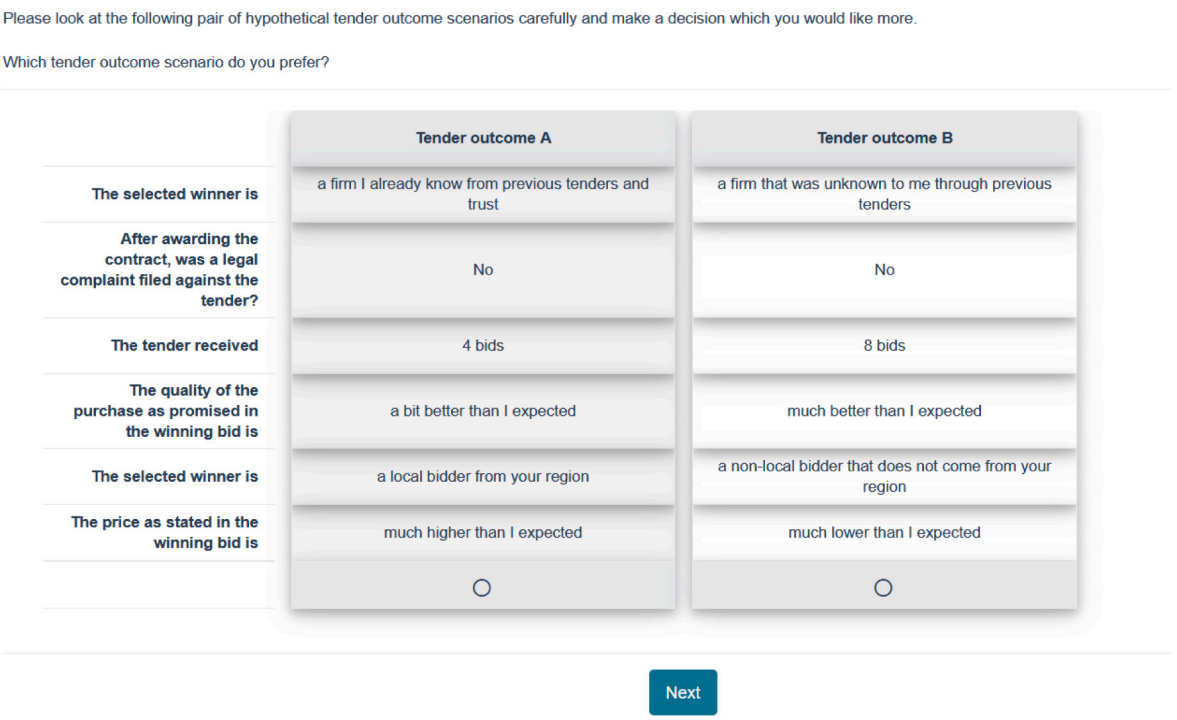

Fig. 1. Example of a Conjoint Scenario. *Notes:* This figure illustrates the binary comparisons presented to respondents in the conjoint experiments.

Choice of attribute realizations. We randomly drew realizations of our attributes and did not exclude any constellation of attribute levels ex-ante from the choice set. Altogether, we have 720 possible combinations of attribute realizations. All (with one exception) attribute levels are shown with equal probabilities. This design has the merit of maximal statistical power to detect bureaucratic preferences. Moreover, the risk that respondents would surmise preferred outcomes for specific attributes from us surveyors is low due to equal probabilities and other properties of conjoint experiment design. Still, one might be concerned that our independent cross-randomization of various attribute realizations might lead to some unrealistic tender outcomes. In particular, respondents might spot unusual or inconsistent attribute combinations (Kessler et al., 2019), leading to a possible bias in our results. To prevent this, we have undertaken several experimental design decisions to enhance realism. First, we included the most relevant and familiar tender components. Second, the litigation odds are set at 10%, because based on data collected from the Finnish Market Court, litigation occurs in less than 10% of Finnish tenders. We did not set it below 10% as we were concerned that a lower realization rate would have risked too low power. Third, the average number of bids displayed is 3.75, which is very close to administrative records from Finland (see Table B.1 in Appendix A). The same applies to regional proximity of the winner, as the cross-sector chance of a winner located in the same region of the buyer is approximately 50% in Finland (Jääskeläinen and Tukiainen, 2019). Fourth, despite benchmarking the price, quality, and reputation to personal expectations and therefore being unable to compare these dimensions with observational data, the heterogeneity of the respondents and the different daily procurement processes in their mindset makes all combinations possible in theory and in practice, even though possibly with different probabilities. Therefore, we also did not exclude any constellation of attribute levels ex-ante from the choice set.[18]

We deliberately defined attribute levels in a very general way since we target a heterogeneous set of participants with potentially varying beliefs about price, quality, and other tender aspects. Due to respondents making repeated choices, we can account for heterogeneous individual beliefs about price or quality. Importantly, we designed the attribute levels regarding price and familiarity in a symmetric way in order to study potential differences in POs' valuations of opportunities and threats (e.g., Jackson and Dutton (1988) or Kuehnhanss et al. (2017) on the "prevention focus" of bureaucrats). Concerning price, the possible attribute realizations are "much lower" or "a bit lower" versus "much higher" or "a bit higher"; for familiarity, they are "a firm I already knew from previous tenders and trusted" versus "a firm that I already had a bad experience with". Both compare to a neutral reference category

[18] According to Bansak et al. (2019) it is also not recommended to make attributes depend on each other unless there is a clear and unambiguous theoretical reason to do so, because it complicates both the estimation and interpretation. However, it is very hard to entirely exclude the possibility of any combination of attribute levels. For example, even unexpectedly high prices can arise when there are many bidders, e.g., due to common or affiliated valuations (Hong and Shum, 2002; Pinkse and Tan, 2005). Thus, we decide not to omit any attribute level constellations ex ante.

("what I expected" and "a firm that was unknown to me through previous tenders").[19] Lastly, our set-up distinguishes between three key quality dimensions: quality of the good or service as promised in the bid, quality signals about past performance reputation, and the regionality of suppliers. Note that the first general quality attribute subsumes a range of relevant potential factors depending on the good or service provided (e.g., the respective delivery time).[20]

2.4. Sample characteristics

This section summarizes socio-demographic and workplace characteristics of our Finnish and German respondents (see Table A.1 in Appendix A for socio-demographics). The majority of our respondents are older than 41, with the age bins 41–50, 51–60, and above 60 years being reported by 30%, 36%, and 10% of POs, respectively. Only 5% of respondents are in their twenties. About half of our respondents are female. Educational attainment is relatively high, with 71% having a college education and 49% with a Master degree. The multidisciplinarity inherent in procurement activity is underscored by the diverse major mix, including engineering, public administration, business administration, and accounting. While general educational patterns are very similar across countries, Finland relies more on engineers while Germany more on public administrators.

Heterogeneous workplaces of POs are detailed in Table A.2 in Appendix A. In both countries, most POs work for municipal authorities (40% on average). Other employers include publicly owned companies (18%), state/regional government (17%), and federal/central government (14%). While the relative employment across government agencies is quite comparable across both countries, we observe different size distributions of the procurement offices our respondents work for, with Finnish offices being smaller on average. In both countries, the most frequently chosen size group is 2–4 POs per office. However, while about 50% of German respondents work in offices with at most ten employees, this share is 70% in Finland. Likewise, Finnish POs report office sizes of more than 100 employees in only 5% of cases, while in Germany, it is 16%.

The observed work contracts indicate high job stability levels since only 4% have temporary employment contracts. Subjects are also relatively likely to engage in managerial tasks, with 40% of all participants reporting personnel responsibilities. On average, we observe a large heterogeneity of tasks that POs engage in, and many report multiple tasks (see Figure A.2 in Appendix A). Participating bureaucrats are also relatively experienced, having collected about 8.5 years of experience in their current office and 11.6 years in public procurement in general. The difference with respect to these two types of experience suggests a shared pattern of rather low job mobility within the public sector on average. The majority of respondents are satisfied with the competence of their own offices, with 67% choosing to agree or strongly agree to this notion.

While most responding POs typically use open tendering procedures (i.e., unrestricted call for bids), Finnish participants use them substantially more than their German counterparts (81% vs. 55%). On the other hand, the use of awarding procedures remain very similar between the two countries, with scoring (best price-quality ratio, also known as "most economically advantageous tender") allocations slightly outranking the lowest price mechanisms. Slightly less than half of our respondents typically work on contracts valued above the EU regulatory thresholds or subject to secondary objectives (mostly additional environmental concerns or support to SMEs). The median number of bids reported was about four and is thus fairly high compared to data for the EU (3 in the period 2007–2017) and the US federal procurement (2 in the period 2008–2018).[21] On average, 39% of participants in both countries have managed tenders that were exposed to litigation, mostly due to bid protests that challenged the choice of the awarded firm.

3. Experimental analysis

This section introduces the empirical method and presents the main results.

3.1. Empirical method

As is standard in the conjoint experiment literature, we use the estimation procedure proposed by Hainmueller et al. (2014). We estimate an OLS regression where the unit of observation is a given scenario (i.e., a card or profile) within a conjoint comparison pair. Since our experiment includes six choices over scenario pairs, each respondent i chooses from $m = 1, \ldots, 12$ scenarios. Our outcome variable y_{im} is a dummy denoting whether or not a given scenario (profile) m was chosen by respondent i. We regress this binary outcome on a large set of indicators X_{imal} providing a complete description of a given conjoint profile m with dummies for each attribute a (e.g., price) and its randomly assigned realization l (e.g., "lower than expected"). The regression leaves out one level dummy for each attribute to provide the baseline. We cluster the standard errors at the respondent level to deal with two

[19] There are two possible criticisms of our choice of attribute realizations. On the one hand, respondents may have different perceptions of what, for example, "much" and "a bit" mean—the so-called "interpersonal incomparability" in survey research (King and Wand, 2007). On the other hand, respondents may have different baseline expectations on, e.g., price and quality. Such heterogeneity of expectations on the outcomes could also contribute to the heterogeneity in responses. To rule out such concerns of spurious heterogeneity and map the results to the actual behavior of POs, we emphasize that our repeated conjoints allow estimates not only between respondents but also within respondents. This design lets us include respondent fixed effects in our model (see Section 3.1), which allows us to control for unobserved heterogeneity in subjective wording interpretations and attribute expectations.

[20] In our conjoint design, lower-than-expected quality scenarios were excluded, reflecting real-world procurement standards that prescribe a minimum quality. Awarding contracts for inferior quality is not feasible in practice. This approach ensures that our experiment is consistent with realistic procurement decisions and focuses on scenarios where quality meets or exceeds expectations.

[21] Figures on EU and US public procurement are sourced from Tenders Electronic Daily (https://ted.europa.eu/TED/main/HomePage.do) and USASpending (https://www.usaspending.gov/).

types of potential error correlations: A mechanical correlation within each choice task (once profile A is chosen, it is impossible to choose profile B), and a second correlation regarding similar attribute valuation within the repeated choices of a given respondent. In particular, we estimate the following linear probability model:

$$y_{im} = \alpha + \sum_{al} \delta_{al} X_{imal} + \varepsilon_{im}. \quad (1)$$

In the specification, α is a constant and ε_{im} the error term. Hainmueller et al. (2014) show that our coefficient of interest, δ_{al}, identifies the AMCE of the realization l of a given attribute a under specific assumptions (see below). This procedure averages the marginal effect of a specific realization over the joint distribution of the remaining attributes. It is thereby a measure for the relative importance of a given attribute and its realization (e.g., "price lower than expected") in driving the overall profile choice. The individual effect should be interpreted as conditional on the other attributes and relative to the omitted baseline attribute level. Including several attributes in the set-up allows for a simultaneous and comparable estimate within a multiple objectives trade-off.[22]

Three assumptions are required for identification. First, the attributes' levels need to be randomly assigned to each profile. This is true by research design. Second, there cannot be carryover effects for the potential outcomes. This means that the potential outcomes remain stable across the choice tasks (i.e., no period effect) and that treatments given to a respondent in their other choice tasks do not affect their response in the current task. Third, we assume that there are no profile-order effects; that is, the ordering of profiles within a choice task does not affect responses. We present the test for both carryover and profile-order effects in Appendix H. Moreover, we randomize the order in which the attributes are presented but keep the ordering unchanged for a given respondent. We also present the test for these types of attribute ordering effects in Appendix H. The latter also shows that our effects are robust to multiple-hypothesis testing and various alternative specifications, including different sets of controls at the respondent level (socio-demographics and/or workplace features), different sets of fixed effects as well as using Logit as an estimator instead of binary OLS. In Appendix E we also discuss the role of careless answers and argue on the basis of self-reported attention flags and analysis of answering times that they are unlikely to explain our results. To estimate heterogeneous treatment effects, we ultimately split the data into subsamples and separately estimate AMCEs, providing us with estimates conditional on a specific set of observables.

3.2. Baseline results

Fig. 2 illustrates the main experimental results across both countries. As previously mentioned, we estimate the average causal effect of treatment attribute levels compared to the baseline attribute level—say, much higher than expected prices compared to prices as expected—on the probability of choosing a certain tender outcome profile, holding all other attribute realizations constant. The figure depicts the respective estimates for each experimental attribute realization separately for each country. Altogether, our estimates are based on a pooled sample of 11,196 choices (4836 for Finland and 6360 for Germany) elicited from 403 and 530 distinct respondents, respectively.

First and most important, the results indicate that public buyers care more about avoiding certain negative realizations than achieving positive related realizations. That is, they exert a so called *prevention focus* regarding negative threats to their organization. In particular, both the attributes price and familiarity exhibit economically and statistically strong negative effects in case of adverse realizations (higher than expected price and previous bad experiences, respectively); however, favorable realizations of the same attributes give rise to much smaller effects in absolute terms. In the case of prices, the positive effects are not even statistically significant, and in the case of familiarity they only account for 19% of the size of the negative effect. This decision pattern can be observed in both countries. Please note that our findings may also speak for interpretations consistent with loss aversion (Kahneman and Tversky, 1979) or a form of negativity bias (e.g. Rozin and Royzman, 2001) where a higher weight is placed on negative outcomes than on related gains.[23]

Second, distinct and robust patterns about the priorities of public buyers emerge from our experimental results. POs place the highest priority on avoiding bidders with which they have previously had a bad experience. Compared to an unknown firm (the omitted category), POs select a given tender profile 41 (Finland) or 47 (Germany) percentage points less often if the firm previously performed badly, respectively. The second priority of POs relates to avoiding unexpected price hikes, which decreases the support for a given tender outcome by 27 (29) percentage points in Finland (Germany) in the case of a much higher price than expected. In comparison with the pooled sample, very high price-hikes—compared to expected prices—are approximately two-thirds as disliked

[22] The interpretation of AMCE estimates in conjoint experiments reflects the *average* preferences of respondents. Recently, Abramson et al. (2022) show that such a majoritarian interpretation of the AMCE fails when a minority of respondents with strong preferences is alongside a majority that prefers the opposite but less strongly. The greater the correlation between the direction and intensity of these two forces, the more misleading the interpretation of the AMCE as average preference. This critique boils down to the difference between an average and a median interpretation of the results, as a small minority with consistent, extreme beliefs could sway the average in its favor but not the median. We believe that a majoritarian interpretation fits our framework and is of interest to the scope of this work. First, contrary to political economy outcomes such as voting addressed in Abramson et al. (2022), a median PO is not decisive for procurement outcomes. Second, as shown in Section 4, PO's preferences are mostly aligned with their own office priorities, and we do not find strong effect heterogeneities overall. Therefore, our sample is unlikely to have groups with overly extreme views.

[23] This holds both for the fact that negative price and reputation attributes are weighed stronger than positive realizations as well as the observation that effects are larger the more negative the price shock (compared to one's expectation) but are constant regarding positive price realizations. These explanations are observationally equivalent and, ultimately, our experiment cannot disentangle them. More generally, one possible driver of this prevention focus could be the regulatory environment of bureaucrats, which focusses more on avoiding failures than rewarding gains.

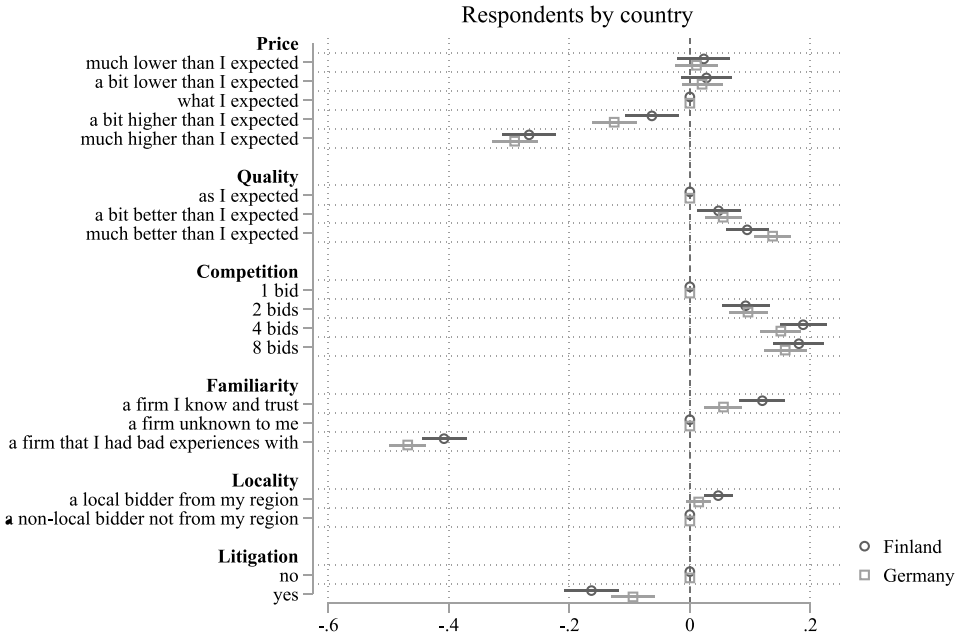

Fig. 2. Baseline results of Conjoint experiments.
Notes: Estimated coefficients of tender outcome attributes on the probability of deciding in favor of a tender outcome. Estimation details provided in Eq. (1). The horizontal lines indicate 95% cluster-robust confidence intervals. Points without these lines indicate the respective reference categories for the effects of the attributes. The respective point estimates are presented in Table A.3 in Appendix A.

as a bad-reputation winner—compared to an unknown winner. Positive realizations of the quality attribute appear to have significant positive effects on support but account for only a fraction of the negative price or reputation concerns, respectively.

Bureaucrats also seem to value a certain degree of competition in calls for tenders, although at a rapidly decreasing rate. In particular, receiving two bids compared to merely one increases the likelihood of support by about ten percentage points. An increase to four offers, however, only increases the same estimate up to 19 or 15 percentage points (Finland and Germany, respectively) as compared to one bid only. A further doubling of the number of bids to eight, in turn, has no statistically significant increase in support when compared to the four-bid scenario. In interpreting these results, we stress that the estimates are conditional on all other attributes, in particular holding the difference between price and quality with PO's expectations constant. This is highly relevant since competition is usually considered a means to an end with strong instrumental benefits for price and quality (Coviello et al., 2018a; Kang and Miller, 2022). Therefore, the observed effects should be interpreted to relate to the value of competition per se and point to a trade-off in the eyes of POs' as hypothesized in Section 2.3.

The results also show that being challenged by a legal complaint after awarding the contract significantly decreases support for a tender outcome in POs' eyes. To improve applicability to different settings, we abstain from specifying the exact type of litigation (usually either a bid protest or a solicitation challenge). Moreover, the choice scenarios specify whether or not litigation has actually occurred after awarding the contract. Since POs usually do not know beforehand whether litigation will occur, we interpret our results as providing an upper bound of the effect that perceived litigation risk has on POs' choices. And due to the fact that the estimates of actual litigation are dwarfed by those of the familiarity and price effects, we conclude that litigation risks have only relatively moderate negative implications for individual bureaucratic support.[24]

Lastly, the choice scenarios also feature whether the winning firm comes from the respondent's region or not. While we estimate a positive effect for a local bidder, the effect is very small and not economically meaningful. At first glance, this may be in contrast to observations of strong regionalism in procurement activity, but correlations between procurement and firm locations arise due to a number of reasons, including better (informal) information channels, past interactions, or even political favoritism and corruption (Baltrunaite, 2019; Ryan, 2020; Baránek and Titl, 2024). In our unique set-up, we control for most of these other channels through past interactions, price, quality, and litigation considerations. Our results suggest that pure geographic proximity (i.e., to buy local) does not play a large role in shaping bureaucratic preferences.

The results are qualitatively the same for both Finland and Germany. Overall, any differences are very small in magnitude. We observe small differences in how respondents value familiarity and locality of the bidder with Finnish respondent valuing

[24] Field evidence shows that the degree of court inefficiency increases the delay in public contract executions (Decarolis et al., 2020b). Our finding might signal that the judicial system is perceived as efficient in both countries, and thus a judicial complaint usually does not considerably interfere with the tender outcome or procurement process. The 2020 World Justice Project findings confirm that both Finland and Germany currently rank very high in terms of how the rule of law is experienced and perceived in practice. See https://worldjusticeproject.org/rule-of-law-index/ for more details.

knowing and trusting bidders slightly more. This could be due to some Finnish procurers, especially small municipalities, being more isolated than their German counterparts and thus rarely encountering bidders outside their area. Finnish procurers appear to be also somewhat less intolerant towards prices which are a bit higher than expected. However, we find no statistically significant differences in any coefficients between the two countries when adjusting the p-values for multiple testing. We report the p-values for differences in coefficients in Table A.3.

In Appendix H, we outline the checks conducted to ensure the robustness of our findings. We summarize that our results hold against potential carryover effects as well as profile and order effects. Moreover, we show that our results are robust to specification and plausibility checks and the risks of multiple hypothesis testing.

4. Discussion of mechanisms

This section sets out to understand why bureaucrats decided the way they did in our experiment. Bureaucratic preferences may be motivated internally or externally. Internal (or intrinsic) work motivation may stem, among others, from a good match between personal preferences and office-level interests, for instance through selection (e.g., Friebel et al., 2019). These mission-based views of preferences have to be distinguished from external forces. In case of the latter, bureaucrats may be indirectly influenced by explicit career incentives, which, although arguably low in this context, may still motivate public sector workers and shape their preferences towards organizational goals within a certain regulatory framework. Beyond this, bureaucrats may also be influenced by their peers (Besley and Persson, 2022) or general office culture—even beyond their own office or region. Deviating from informal norms at the workplace may be costly for individual POs and, thus, shape their preferences (Ellingsen and Johannesson, 2008).

In the following, after displaying that respondents are equipped with enough leeway to turn their preferences into relevant decisions at the work place, we discuss different channels behind our experimental findings on the priorities of POs. We argue that (i) results are strikingly similar across respondent and office characteristics, (ii) alternative external forces seem not to matter substantially, and (iii) internal and external (i.e., perceived office-level) interests are largely aligned between bureaucrats and their offices. We caution that none of the presented exercises can be interpreted in a causal manner, but rather rely on a battery of descriptive exercises, including the analysis of follow-up questions, performance-related views of POs, and heterogeneity analysis of the experiment.

4.1. Perceived discretion

The relevance of our study is related to the degree to which bureaucratic choices depend on bureaucratic preferences. Therefore, a necessary prerequisite in this regard is that bureaucrats have sufficient discretion in their daily work to turn idiosyncratic preferences into actual decisions. Thus, before turning to the results discussion, we provide descriptive evidence on the respondents' perception concerning this key aspect. In doing so, we also show that somewhat different views on national procurement work practices—on top of somewhat different institutional environments presented in Section 2.4—arise between the two countries despite sharing the same EU regulatory framework.

Figure A.3 in Appendix A shows that in our setting, most respondents perceive a high degree of discretion, although to a different extent across the two countries.[25] In particular, while 50% of German respondents feel they can influence tender outcomes "very much" or "somewhat", the share of their Finnish peers who feel the same comes in much higher at 82%. The higher discretion of Finnish respondents also translates into a lower share reporting to have "no influence" (4% in Finland vs. 15% in Germany).[26] Moreover, we asked POs about aspects that typically create problems in the procurement process according to their personal experience. Figure A.4 in Appendix A illustrates the respective results and shows notably that the biggest obstacle in the eyes of public buyers is too rigid regulation, which is seen as problematic by 38% of all respondents. This finding corroborates the importance of discretion from the viewpoint of POs. This evidence is also in line with findings on the benefits of bureaucratic autonomy in public contracting (Carril, 2021), which respondents seem to internalize in our setting.[27]

In Section 2.4, we displayed that German and Finnish POs work under somewhat different institutional environments. For example, we show that German bureaucrats appear to work in larger offices and use open tender procedures less often than Finnish POs. The survey answers reported here highlight less perceived discretion to German respondents. Altogether, on top of fewer extrinsic career incentives (see below), this evidence is indicative of (perceived) institutional differences between the two countries despite sharing the same EU regulatory framework.[28]

[25] See Q10 in Appendix F.

[26] To investigate whether perceptions of discretion are related to other demographic features, we regress a dummy of perceived discretion on respondents' background characteristics. The results are presented in the first column of Table A.4 in Appendix A. Having influence over procurement outcomes does not correlate with any background characteristics other than respondent's country, where we see a large negative correlation for German respondents.

[27] The second biggest perceived threat is that the contractor causes problems, which, in light of our experimental results, stresses how valuable reputation and trust with potential suppliers are in the procurement process. This is also somewhat of a bigger problem in Germany (see Table A.4 in Appendix A). Budgetary constraints and litigation risks (15% and 9%, respectively) are less problematic in the respondents' eyes.

[28] To assess whether the perceptions about threats to the procurement process are related to socio-demographics, we regress dummy variables for different perceived threats on respondents' background characteristics. The results are presented in columns 2–7 of Table A.4 in Appendix A. Most background characteristics do not have a statistically significant correlation with any of the perceived threats.

4.2. Lack of heterogeneity

As already highlighted in Section 3.2, we observe a striking similarity in results between Finland and Germany. These similarities in our experimental findings also extend to almost all observable features of respondents and their organizations, which is particularly interesting since individual demographic characteristics, task structures, and office-level features have been shown to affect bureaucratic performance in procurement (e.g., Decarolis et al., 2020a, 2021; Best et al., 2023).[29] Contrary to these previous findings, we provide evidence that these factors do not seem to influence preference formation in our experimental setting.

To this end, we first investigate a large number of potential effect heterogeneities in terms of socio-demographics, task structures, and office-level characteristics. Because we have no ex-ante expectations about potential differences in the direction or strength of the effects, this part of our analysis is more exploratory in nature. As such, our analysis must be presented with three important caveats. First, the background characteristics are not randomly assigned, which means that interpreting any given characteristic as having a causal effect on the difference in preferences could be problematic. Second, statistical power in sub-groups is a potential issue, although we usually have small confidence intervals around our estimates. Third, due to a large number of possible heterogeneities, we cannot completely rule out spurious differences arising from multiple hypotheses testing in our sub-group analysis.

In our heterogeneity analysis, we generally split the sample into two or more groups depending on the source of variation. In the case of continuous variables, we split the sample according to the median value of the variable of interest (e.g., years of experience). In the case of categorical variables, we typically assign groups based on selected elicited categories (e.g., type of award mechanism). We separately estimate the original regressions for these new sub-samples and assess whether the estimated effects differ. We conduct our analysis by comparing the marginal means of attributes in various sub-groups. The marginal means estimate of an attribute indicates how likely the respondents choose a tender outcome conditional on the attribute appearing in the profile. This approach goes back to Leeper et al. (2020), who show that simply comparing AMCEs between subgroups can be problematic because there may be (otherwise unobserved) differences in how subgroups value the baseline attribute. By following their procedure, our choice of reference group does not influence our heterogeneity analysis.[30]

In particular, we analyze heterogeneities based on respondents' age, gender, education level, education type, work experience, and exposure to EU regulation. Importantly, given that many POs feel that they have discretion and can influence their work outcomes (see the previous subsection), we test also whether differences in discretion explain observed differences in choice behavior. Figure A.5 of the Appendix provides evidence that discretion does not seem to be related to the priorities elicited from bureaucrats. Another potentially important issue affecting POs priorities could be the their task structure or the industry they typically procure in. Both factors, however, do not provide any heterogeneous results (Figures A.6 and A.7 in Appendix A). We carry out additional tests regarding prior exposure to litigations, typical awarding mechanism, use of secondary objectives, size of the regular contract, the typical number of bidders, perceived workload and discretion, and job responsibilities. Moreover, we explore heterogeneities concerning (perceived) office competence and government tier. We refer to Appendix I for a detailed description of all sub-group analyses.

The treatment effects are extremely stable across almost all sub-groups. This striking similarity across major socio-demographics, tasks, and office-related differences points to the existence of very robust priorities among procurement bureaucrats.

4.3. The role of external forces

In what follows, we discuss a number of empirical exercises indicating that external forces do not play a dominant role in our bureaucratic environment.

Incentive schemes. It could be that the government employs highly efficient incentive schemes towards its bureaucrats, and, since the fundamental goals of governments are similar across different settings, we observe similar results in terms of bureaucratic choices. Such incentives could come in the form of explicit pay-for-performance plans and career concerns (Bertrand et al., 2019). To address this issue, we study the incentives public sector workers face in our settings. In both countries, to the best of our knowledge, explicit monetary incentive schemes are essentially unavailable for public buyers.[31] Turning to career concerns, our descriptive evidence shows that POs in our survey perceive low levels of job-related incentives through individual career concerns. A large majority states that tender outcomes are "absolutely not" or "rather not" important for their career prospects—and even more so in Germany (80%) than in Finland (67%, see Figure A.9 in Appendix A). This is in line with a high job stability pattern of POs, as shown in Section 2.4.

[29] Even early work already pointed out that task structures and office characteristics may affect bureaucrat behavior (Christensen and Lægreid, 2004). Work on representative public administration (e.g. Bradbury and Kellough, 2011) also pointed towards the role of demographic characteristics of public servants (inasmuch as they represent the general public) for beneficial public sector outcomes.

[30] For example, in Fig. 2, which shows conditional AMCEs for Finland and Germany, Finnish respondents seem to have slightly stronger preferences for firms they know and trust than their German counterparts. However, looking at the marginal means in Figure A.8, we see that the underlying difference is based on the valuation of the baseline attribute: German POs see unknown firms as slightly more preferable than Finnish POs do. Nevertheless, this particular difference between subgroups is very small.

[31] This does not mean that public sector workers in other contexts do not face pay-for-performance schemes. In Norway, for instance, top local bureaucrats are financially rewarded (albeit modestly) for improving on budget deficits, but not when budget surpluses increase (Geys et al., 2017). Burgess et al. (2017) report team-level incentives for public sector workers in the UK. According to Checchi et al. (2021), only 10% or 25% of public sector workers in continental Europe (including Germany) or Europe as a whole are subject to some form of pay-for-performance schemes, respectively. This stands in contrast to about 2/3 of US public workers which face some sort of monetary incentives.

When directly inquired on the role of incentives in their daily work, only 26% of our respondents perceive tender outcomes to matter for their career prospects (20% in Germany and 33% in Finland). Importantly, when splitting the sample into subgroups depending on whether or not individuals view tender outcomes as important for their career prospects, we find no sizeable differences between the marginal means of any attributes (Figure A.10 in Appendix A). This even holds when only considering respondents with the strongest or weakest career concerns (Figure A.11 in Appendix A).

Nevertheless, our survey question regarding perceived career incentives might not capture all the indirect incentives individual bureaucrats face. One additional aspect of explicit or implicit incentives could be the relative job position already attained within the organization. As a proxy for such forces, we exploit information on whether a given bureaucrat carries personnel responsibility. A significant portion of the bureaucrats in our sample have some managerial responsibility for employees—40% in total, 45% in Germany, and 33% in Finland. Nevertheless, this personnel responsibility is also not a predictor of any of our results (Figure A.12 in Appendix A). In sum, neither respondents' perceived career concerns nor their hierarchical position within the institution are associated with their stated preferences. This makes it unlikely that the observed similarities in stated choices arise due to similarities in the objectives of government incentive schemes.

Office–level norms. Bureaucratic preferences could also be also shaped by office-level norms (Besley and Persson, 2022). Please note that since we cannot link office-level data to experimental choices due to data protection issues (see more on this in Appendix Section B), we cannot directly account for office-level norms by aggregating multiple survey responses per office or including office-level fixed effects. However, our survey asks respondents to report staff (office) size, which should adequately approximate office-level norms (Burgess et al., 2017). If strong peer effects shaped our respondents' decisions, then participants from single employee offices, small or large offices should have reacted differently in our experiment (e.g. Burgess et al. (2017) points to stronger peer effects in smaller than in larger offices). However, heterogeneity analysis using subsamples of different office sizes does not show that different office sizes matter for our experimental results (see Figure A.13 in Appendix A). Even buyers in single-officer units do not have other priorities than bureaucrats in multi-person offices. Also, the fact that a broad set of procurement practices (see bloc A.3 in Appendix E)) does not systematically affect experimental choices points towards office norms playing a rather minor role.

4.4. Alignment of personal preferences with office–level interests

As previously described in Section 2.2, we ask respondents follow–up questions after the experiment in order to distinguish how their preferences and career concerns (i.e., presumed office–level interests) may differ and drive our experimental results. Specifically, we ask subjects to state their most important and least important attribute—among those of the experiment—regarding their personal preferences or perceived importance to the office for career progress. Table A.5 in Appendix A shows the results for the respective shares of respondents who differ in their views on what dimension is most (least) important for themselves or the office. Overall and most importantly, it appears that the fraction of respondents who prefer (frown upon) individual attributes differently from their office is always relatively low—at most 33%. The fraction of respondents for whom all of the attributes are aligned between both margins is 51% and thus constitutes a majority. This indicates that most surveyed bureaucrats have aligned interests with what they think their office prioritizes the most or the least across all attributes. This points towards a good match of bureaucrats with their organizations' mission and likely represents conservative but efficient bureaucracies, according to Besley and Ghatak (2005).

It is important to note that attributes are evaluated with respect to their perceived importance to the organization only on a subset of respondents. This subset includes those who believe their career prospects are impacted by procurement outcomes, representing 241 out of 933 respondents (25.8%). For these respondents, the degree of alignment between personal and office-level priorities has a meaningful interpretation, as the office-level priorities have a direct effect on individual outcomes. However, this limitation of our data implies that we are not capturing the preferences and behaviors of POs who do not perceive their career prospects to be tied to procurement outcomes. Nevertheless, we have shown earlier that the degree of extrinsic motivations does not affect the answering behavior of POs in our conjoint experiment (refer to Figure A.10 in Appendix A). The finding that public buyers put focus on prevention also supports the idea that personal and office interests are substantially aligned among the bureaucrats interviewed.

To understand the mechanisms behind the observed choice behavior of bureaucrats, we have engaged in a battery of descriptive analyses linking choices to further information at our disposal and exploring potential driving factors. While we observe a lack of heterogeneity across contexts and find no substantial influence of external incentive schemes or office norms on choices, the strong alignment between bureaucrats' preferences and the perceived interests of their offices suggests that mission-based motivation play a role. However, we cannot conclusively determine which mechanisms are at play. It remains unclear whether this alignment reflects selection based on innate preferences for public service or gradual preference formation within the public sector. Since our data does not vary institutional contexts or track individual respondents over time, it does not provide sufficient evidence to distinguish between these mechanisms. Our findings suggest that bureaucrats exhibit relatively homogeneous preferences aligned with the mission of their organizations, but we refrain from drawing firm conclusions about the origins of these preferences.

5. Conclusions

Bureaucratic decision-making shapes policy implementation. In a context where officials have wide discretionary power, understanding bureaucratic preferences is arguably important and decisive for public goods provision. We tackle this vital issue in the context of public procurement and analyze the priorities of bureaucrats regarding desirable project outcomes. We field a

unique cross-country survey on public procurement practices to a large sample of real-world bureaucrats and combine it with choice experiments, where respondents repeatedly decide (and thus, trade off) between pairs of hypothetical tender outcomes that differ randomly across key tender features.

Our approach is abductive (Heckman and Singer, 2017) or explorative, as we did not have pre-determined hypotheses for how strong the preferences over attributes would be in relation to each other or what to expect from most of the heterogeneity analysis. Hence, new hypotheses arise in interaction with the survey results, which may also inform future theory work. Our findings provide the first view into the black box of bureaucrats' preferences on desirable work outcomes. Specifically, we learn the importance that procurement officials place on the bid price and quality, the degree of competition, supplier reputation, the geographical proximity of winners, and litigation risks at the tender stage. Our results may be interpreted as first-best choices the POs would make without frictions, such as rules, which our respondents indeed perceive as rigid.

Our first and main finding is that procurement bureaucrats value avoiding threats concerning both prices and supplier reputation more than grasping potential opportunities. This is consistent with a prevention focus (e.g., Jackson and Dutton, 1988) of public buyers against harm to their organization. This is also in line with findings that exceeding the budget is bad for the officials, but staying under budget is not necessarily good (Liebman and Mahoney, 2017). When viewing bureaucrats as agents of a principal such as the general public, a pure prevention focus against negative outcomes in prices is likely not socially optimal if we assume the general public to value gains and losses (averaged over many tenders) equally. Second, for POs, it is more important to avoid bidders with a bad reputation than to elude unexpectedly high prices. This resembles behavior in private sector procurement, where mutual trust and good long-term relationships are important (Stentoft Arlbjørn and Vagn Freytag, 2012), and can be seen as relating to addressing concerns over non-contractable quality, and thus, be a socially valuable preference. Moreover, officials also value the direct (contractable) quality attribute. Third, we show that POs value competition above the expected price and quality implications of it, suggesting it being an independent measure of a successful tender. However, competition is only valued up to a limit, suggesting that administrative costs might be an issue and may be of some concern regarding the incentives of officials to invest in attracting competition. Fourth, litigation concerns are secondary dimension in the evaluation of procurement outcomes. The litigation result is slightly concerning as the threat of litigation is meant to hold officials accountable and incentivize legal purchasing procedures. Finally, there does not seem to be regional favoritism, which is good news for the social planner.

Given the lack of career and pay incentives, the absence of discernible effects of bureau size as a proxy for office-level norms on experimental choices, and the strong alignment between personal and perceived office interests, our findings suggest that bureaucratic decision-making in our setting may be influenced by mission-based preferences of public buyers. This aligns with findings from the World Value Survey, which indicates high levels of mission motivation among public servants in Finland and Germany (Cowley and Smith, 2014). However, we caution against drawing firm conclusions, as institutional factors (e.g., incentives or workplace culture) may also play a role. More broadly, our data does not allow us to determine whether bureaucratic preferences are innately rooted or shaped by institutional influences.

Our work opens the door to new avenues of organizational research in the public sector more generally and in public procurement more specifically. For instance, conducting repeated surveys across sectors could help disentangle the effect mechanisms and provide a more comprehensive understanding of the origins of bureaucratic preferences. While conjoints may be of help here again, their results may not necessarily apply to real-world bureaucratic practice. Thus, these valuable survey approaches may be complemented with targeted field experiments to study how preferences of bureaucrats drive their outcomes. Specifically for public procurement, our evidence indicates, for instance, that the lack of competition observed in many settings is unlikely to predominantly arise from the preferences of public buyers as they appear to value competition in our experiment more than is typically assumed. Thus, future experimental research could address the relevant entry barriers for firms in public procurement markets in more detail and how they could be reduced. Also, further research is necessary to investigate the alignment of preferences with their offices among officers. This exploration could yield a more comprehensive understanding of how intrinsic and extrinsic motivations interact across various bureaucratic contexts. Also, future work should tease out the relative importance of different motives and channels in bureaucratic preferences and show their relevance for economic outcomes.

Policy implications. This paper draws a number of policy implications highly relevant not only for Finland and Germany, but also for countries with similarly high levels of public sector capacity (Bosio et al., 2022). These policy implications are not limited to the realm of public procurement, but also include some general insights for environments where bureaucrats face low incentives and wield sufficient autonomy. We base this external validity of our results on several factors. For Finland, we compare our survey data to the universe of administrative procurement records, and find our respondents and their offices to be representative regarding various observables, including regions, office types, contract procedures, awarding mechanisms, and the typical number of bidders. While similar data does not exist for Germany, we note that the priorities elicited from German procurers are very similar to those elicited in Finland, and both are similarly robust across all observables. The procurers themselves in both countries comprise a heterogeneous set of individual buyers across a broad range of individual tasks, demographics, and office characteristics, and operate in the same EU regulatory framework for procurement purchases. Moreover, while Finland has excellent data to provide supportive evidence for our sample to be representative, the German economy and procurement market, represents the largest in the EU. The external validity of our results for public bureaucracies more generally comes from the fact that public procurement is *per se* representative of a large share of bureaucrats. A considerable proportion of public employees will be involved in the procurement process at some point in their careers, since a large share of administrative tasks are in some way related to procurement.

Our study raises two major policy implications for the design and management of public sector organizations. First, there may be scope for guiding POs to focus more on achieving lower prices (especially in scoring auctions). This follows from our finding that

POs place a disproportionate focus on avoiding too high prices, while the government should be a risk-neutral agent valuing gains and benefits comparably. Second, optimal public sector management should respond to the high value POs give to the reputation of bidding firms by incorporating clear criteria on past performance into the tender choice framework. Introducing rewards or penalties during the selection process based on past performance is also proposed by Decarolis et al. (2016) and Butler et al. (2020), but is rarely implemented in practice due to transparency and competition concerns of legislators. Third, in line with recent evidence on the benefits of granting POs more discretion (Bosio et al., 2022), our finding of preference alignment among bureaucrats can be seen as a further reason to grant POs more discretion in the procurement process. Our results also have implications for supply management of firms interested in making business with public buyers. Notably, firms entering or expanding their operations in the public procurement market should pay particular attention towards establishing a good track record of procurement operations.

Declaration of competing interest

Blesse is thankful for funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project ID 139943784—SFB 884. Blesse, Bohne and Giuffrida are thankful to the Leibniz Association for its support through the project Market Design by Public Authorities. Jääskeläinen and Tukiainen are thankful for funding by the Yrjö Jahnsson foundation. Tukiainen is thankful to the Academy of Finland project funding 340045, and the European Union (ERC, INTRAPOL, 101045239). Views and opinions expressed are only those of the authors, however, and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

All authors are thankful for EconPol funding. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data availability

The data that has been used is confidential.

Appendix. Supplementary data

The Appendix to this article can be found as supplementary material online at https://doi.org/10.1016/j.jebo.2024.106716.

References

Abramson, S.F., Kocak, K., Magazinnik, A., 2022. What do we learn about voter preferences from conjoint experiments? Am. J. Political Sci. 66 (4), 1008–1020.
Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. J. Polit. Econ. 105 (1), 1–29.
Albano, G.L., Cesi, B., Iozzi, A., 2023. Teaching an old dog a new trick: reserve price and unverifiable quality in repeated procurement. J. Econ. & Manag. Strat. 32 (2), 377–399.
Alesina, A., Tabellini, G., 2008. Bureaucrats or politicians? Part II: Multiple policy tasks. J. Public Econ. 92 (3), 426–447.
Ameriks, J., Briggs, J., Caplin, A., Shapiro, M.D., Tonetti, C., 2020. Long-term-care utility and late-in-life saving. J. Polit. Econ. 128 (6), 2375–2451.
Andre, P., Pizzinelli, C., Roth, C., Wohlfart, J., 2022. Subjective models of the macroeconomy: Evidence from experts and representative samples. Rev. Econ. Stud. 89 (6), 2958–2991.
Andrews, I., Barron, D., 2016. The allocation of future business: Dynamic relational contracts with multiple agents. Amer. Econ. Rev. 106 (9), 2742–2759.
Ash, E., MacLeod, W.B., 2015. Intrinsic motivation in public service: Theory and evidence from state supreme courts. J. Law Econ. 58 (4), 863–913.
Ashraf, N., Bandiera, O., Davenport, E., Lee, S.S., 2020. Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. Amer. Econ. Rev. 110 (5), 1355–1394.
Asker, J., Cantillon, E., 2010. Procurement when price and quality matter. Rand J. Econ. 41 (1), 1–34.
Baltrunaite, A., 2019. Political contributions and public procurement: Evidence from Lithuania. J. Eur. Econom. Assoc. 18 (2), 541–582.
Baltrunaite, A., Giorgiantonio, C., Mocetti, S., Orlando, T., 2021. Discretion and supplier selection in public procurement. J. Law Econ. Organ. 37 (1), 134–166.
Bandiera, O., Best, M.C., Khan, A.Q., Prat, A., 2021. The allocation of authority in organizations: A field experiment with bureaucrats. Q. J. Econ. 136 (4), 2195–2242.
Bansak, K., Hainmueller, J., Hangartner, D., 2016. How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. Science 354 (6309), 217–222.
Bansak, K., Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2018. The number of choice tasks and survey satisficing in conjoint experiments. Political Anal. 26 (1), 112–119.
Bansak, K., Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2019. In: James, N., Donald, P.G. (Eds.), "Conjoint Survey Experiments" For Druckman. In: Cambridge Handbook of Advances in Experimental Political Science, Cambridge University Press, New York.
Bansak, K., Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2021. Beyond the breaking point? Survey satisficing in conjoint experiments. Political Sci. Res. Methods 9 (1), 53–71.
Baránek, B., Titl, V., 2024. The cost of favoritism in public procurement. J. Law Econ. 67 (2), 445–477.
Bechtel, M.M., Hainmueller, J., Margalit, Y., 2014. Preferences for international redistribution: The divide over the Eurozone bailouts. Am. J. Political Sci. 58 (4), 835–856.
Bechtel, M.M., Scheve, K.F., 2013. Mass support for global climate agreements depends on institutional design. Proc. Natl. Acad. Sci. 110 (34), 13763–13768.
Beetsma, R., Burgoon, B., Nicoli, F., de Ruijter, A., Vandenbroucke, F., 2022. What kind of EU fiscal capacity? Evidence from a randomized survey experiment in five European countries in times of corona. Econ. Policy 37 (111), 411–459.
Bertrand, M., Burgess, R., Chawla, A., Xu, G., 2019. The glittering prizes: Career incentives and bureaucrat performance. Rev. Econ. Stud. 87 (2), 626–655.
Besley, T., Burgess, R., Khan, A., Xu, G., 2022. Bureaucracy and development. Ann. Rev. Econom. 14.
Besley, T., Ghatak, M., 2005. Competition and incentives with motivated agents. Amer. Econ. Rev. 95 (3), 616–636.
Besley, T., Persson, T., 2022. Organizational dynamics: culture, design, and performance. J. Law Econ. Org. 40 (2), 394–415.
Best, M.C., Hjort, J., Szakonyi, D., 2023. Individuals and organizations as sources of state effectiveness. Amer. Econ. Rev. 113 (8), 2121–2167.

Bloom, N., Propper, C., Seiler, S., Van Reenen, J., 2015. The impact of competition on management quality: evidence from public hospitals. Rev. Econ. Stud. 82 (2), 457–489.
Board, S., 2011. Relational contracts and the value of loyalty. Amer. Econ. Rev. 101 (7), 3349–3367.
Boland, M., Godsell, D., 2021. Bureaucratic discretion and contracting outcomes. Account. Organ. Soc. 88, 101–173.
Bosio, E., Djankov, S., Glaeser, E., Shleifer, A., 2022. Public procurement in law and practice. Amer. Econ. Rev. 112 (4), 1091–1117.
Bradbury, M., Kellough, J.E., 2011. Representative bureaucracy: Assessing the evidence on active representation. Am. Rev. Public Adm. 41 (2), 157–167.
Broockman, D.E., Skovron, C., 2018. Bias in perceptions of public opinion among political elites. Am. Political Sci. Rev. 112 (3), 542–563.
Burgess, S., Propper, C., Ratto, M., Tominey, E., 2017. Incentives in the public sector: Evidence from a government agency. Econ. J. 127 (605), 117–141.
Burgess, S., Ratto, M., 2003. The role of incentives in the public sector: Issues and evidence. Oxf. Rev. Econ. Policy 19 (2), 285–300.
Butler, J.V., Carbone, E., Conzo, P., Spagnolo, G., 2020. Past performance and entry in procurement: An experimental investigation. J. Econ. Behav. Organ. 173, 179–195.
Buurman, M., Delfgaauw, J., Dur, R., Van den Bossche, S., 2012. Public sector employees: Risk averse and altruistic? J. Econ. Behav. Organ. 83 (3), 279–291.
Calzolari, G., Spagnolo, G., 2009. Relational contracts and competitive screening. CEPR Discussion Paper No. 7434.
Carril, R., 2021. Rules versus discretion in public procurement. Department of Economics and Business, Universitat Pompeu Fabra Working Paper No. 1765.
Checchi, D., Fenizia, A., Lucifora, C., 2021. Public sector jobs: Working in the public sector in Europe and the US. NBER Working Paper No. 23350.
Christensen, T., Lægreid, P., 2004. Public administration research in Norway: Organization theory, institutionalism and empirical studies in a democratic context. Public Adm. 82 (3), 679–690.
Coibion, O., Gorodnichenko, Y., Kumar, S., 2018. How do firms form their expectations? New survey evidence. Amer. Econ. Rev. 108 (9), 2671–2713.
Colonnelli, E., Prem, M., Teso, E., 2020. Patronage and selection in public sector organizations. Amer. Econ. Rev. 110 (10), 3071–3099.
Coviello, D., Guglielmo, A., Spagnolo, G., 2018a. The effect of discretion on procurement performance. Manage. Sci. 64 (2), 715–738.
Coviello, D., Moretti, L., Spagnolo, G., Valbonesi, P., 2018b. Court efficiency and procurement performance. Scand. J. Econ. 120 (3), 826–858.
Cowley, E., Smith, S., 2014. Motivation and mission in the public sector: Evidence from the world values survey. Theory and Decision 76 (2), 241–263.
Cunningham, T., de Quidt, J., 2015. Implicit preferences inferred from choice. mimeo.
Decarolis, F., de Rassenfosse, G., Giuffrida, L.M., Iossa, E., Mollisi, V., Raiteri, E., Spagnolo, G., 2021. Buyers' role in innovation procurement: Evidence from US military R&D contracts. J. Econ. Manag. Strat. 30 (4), 697–720.
Decarolis, F., Giuffrida, L.M., Iossa, E., Mollisi, V., Spagnolo, G., 2020a. Bureaucratic competence and procurement outcomes. J. Law Econ. Organ. 36 (3), 537–597.
Decarolis, F., Mattera, G., Menon, C., 2020b. Delays at the border: Court efficiency and delays in public contracts. CEPR Discussion Paper No. 14856.
Decarolis, F., Spagnolo, G., Pacini, R., 2016. Past performance and procurement outcomes. NBER Working Paper No. 22814.
DellaVigna, S., Pope, D., 2018a. Predicting experimental results: Who knows what? J. Polit. Econ. 126 (6), 2410–2456.
DellaVigna, S., Pope, D., 2018b. What motivates effort? Evidence and expert forecasts. Rev. Econ. Stud. 85 (2), 1029–1069.
Detkova, P., Pronin, P., Tkachenko, A., Yakovlev, A.A., 2021a. The changing perceptions of corruption during the COVID pandemic in Russia. mimeo.
Detkova, P., Tkachenko, A., Yakovlev, A., 2021b. Gender heterogeneity of bureaucrats in attitude to corruption: Evidence from list experiment. J. Econ. Behav. Organ. 189, 217–233.
Duflo, E., Greenstone, M., Pande, R., Ryan, N., 2018. The value of regulatory discretion: Estimates from environmental inspections in India. Econometrica 86 (6), 2123–2160.
Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: The human side of incentive theory. Am. Econ. Rev. 98 (3), 990–1008.
Fazekas, M., 2017. Assessing the quality of government at the regional level using public procurement data. Working Paper No. 12-2017.
Finan, F., Olken, B.A., Pande, R., 2017. The personnel economics of the developing state. In: Handbook of Economic Field Experiments, vol. 2, Elsevier, pp. 467–514.
Friebel, G., Kosfeld, M., Thielmann, G., 2019. Trust the police? Self-selection of motivated agents into the German police force. Am. Econ. J. Microecon. 11 (4), 59–78.
Fuster, A., Zafar, B., 2021. The sensitivity of housing demand to financing conditions: Evidence from a survey. Am. Econ. J.: Econ. Policy 13 (1), 231–265.
Gentry, M.L., Hubbard, T.P., Nekipelov, D., Paarsch, H.J., et al., 2018. Structural econometrics of auctions: A review. Found. Trends Econom. 9 (2–4), 79–302.
Geys, B., Heggedal, T.-R., Sørensen, R.J., 2017. Are bureaucrats paid like CEOs? Performance compensation and turnover of top civil servants. J. Public Econ. 152, 47–54.
Gregg, P., Grout, P.A., Ratcliffe, A., Smith, S., Windmeijer, F., 2011. How important is pro-social behaviour in the delivery of public services? J. Public Econ. 95 (7–8), 758–766.
Gründler, K., Potrafke, N., 2020. Experts and epidemics. CESifo Working Paper No. 8556.
Haaland, I., Roth, C., Wohlfart, J., 2021. Designing information provision experiments. J. Econ. Lit. 61 (1), 3–40.
Hainmueller, J., Hangartner, D., Yamamoto, T., 2015. Validating vignette and conjoint survey experiments against real-world behavior. Proc. Natl. Acad. Sci. 112 (8), 2395–2400.
Hainmueller, J., Hopkins, D.J., 2015. The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. Am. J. Political Sci. 59 (3), 529–548.
Hainmueller, J., Hopkins, D.J., Yamamoto, T., 2014. Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. Political Anal. 22 (1), 1–30.
Halonen, K.-M., Tukiainen, J., 2020. Competition and Litigation in Swedish Public Procurement. Swedish Competition Authority Report 2020:1.
Hanna, R., Wang, S.-Y., 2017. Dishonesty and selection into public service: Evidence from India. Am. Econ. J.: Econ. Policy 9 (3), 262–290.
Heckman, J.J., Singer, B., 2017. Abducting economics. Amer. Econ. Rev. 107 (5), 298–302.
Hong, H., Shum, M., 2002. Increasing competition and the winner's curse: Evidence from procurement. Rev. Econ. Stud. 69 (4), 871–898.
Horiuchi, Y., Markovich, Z.D., Yamamoto, T., 2020. Does conjoint analysis mitigate social desirability bias? MIT Political Science Department Research Paper No. 2018-15.
Hyytinen, A., Lundberg, S., Toivanen, O., 2018. Design of public procurement auctions: Evidence from cleaning contracts. Rand J. Econ. 49 (2), 398–426.
Jääskeläinen, J., Tukiainen, J., 2019. Anatomy of public procurement. VATT Institute for Economic Research Working Paper No. 118.
Jackson, S.E., Dutton, J.E., 1988. Discerning threats and opportunities. Adm. Sci. Q. 370–387.
Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.
Kang, K., Miller, R.A., 2022. Winning by default: Why is there so little competition in government procurement? Rev. Econ. Stud. 89 (3), 1495–1556.
Kessler, J.B., Low, C., Sullivan, C.D., 2019. Incentivized resume rating: Eliciting employer preferences without deception. Amer. Econ. Rev. 109 (11), 3713–3744.
Khan, A.Q., Khwaja, A.I., Olken, B.A., 2019. Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. Amer. Econ. Rev. 109 (1), 237–270.
King, G., Wand, J., 2007. Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. Political Anal. 15 (1), 46–66.
Kuehnhanss, C.R., Murdoch, Z., Geys, B., Heyndels, B., 2017. Identity, threat aversion, and civil servants' policy preferences: Evidence from the European parliament. Public Admin. 95 (4), 1009–1025.
Kutlina-Dimitrova, Z., Lakatos, C., 2016. Determinants of direct cross-border public procurement in EU member states. Rev. World Econ. 152 (3), 501–528.
Leaver, C., 2009. Bureaucratic minimal squawk behavior: Theory and evidence from regulatory agencies. Amer. Econ. Rev. 99 (3), 572–607.

Leeper, T.J., Hobolt, S.B., Tilley, J., 2020. Measuring subgroup preferences in conjoint experiments. Political Anal. 28 (2), 207–221.
Levin, J., 2003. Relational incentive contracts. Amer. Econ. Rev. 93 (3), 835–857.
Liebman, J.B., Mahoney, N., 2017. Do expiring budgets lead to wasteful year-end spending? Evidence from federal procurement. Amer. Econ. Rev. 107 (11), 3510–3549.
Link, S., Peichl, A., Roth, C., Wohlfart, J., 2023. Information frictions among firms and households. J. Mon. Econ. 135, 99–115.
Meyer-Sahling, J., Mikkelsen, K.S., Schuster, C., 2021. Merit recruitment, tenure protections and public service motivation: Evidence from a conjoint experiment with 7,300 public servants in Latin America, Africa and Eastern Europe. Public Admin. 99 (4), 740–757.
OECD, 2019. Reforming Public Procurement. p. 116.
Oliveros, V., Schuster, C., 2018. Merit, tenure, and bureaucratic behavior: Evidence from a conjoint experiment in the Dominican Republic. Comparative Political Stud. 51 (6), 759–792.
Pinkse, J., Tan, G., 2005. The affiliation effect in first-price auctions. Econometrica 73 (1), 263–277.
Prendergast, C., 2007. The motivation and bias of bureaucrats. Amer. Econ. Rev. 97 (1), 180–196.
Rakotonarivo, O.S., Schaafsma, M., Hockley, N., 2016. A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. J. Environ. Manag. 183, 98–109.
Rasul, I., Rogger, D., 2018. Management of bureaucrats and public service delivery: Evidence from the Nigerian Civil Service. Econ. J. 128 (608), 413–446.
Rasul, I., Rogger, D., Williams, M.J., 2020. Management, organizational performance, and task clarity: Evidence from Ghana's civil service. J. Public Admin. Res. Theory.
Rozin, P., Royzman, E.B., 2001. Negativity bias, negativity dominance, and contagion. Personal. Soc. Psychol. Rev. 5 (4), 296–320.
Ryan, N., 2020. Contract enforcement and productive efficiency: Evidence from the bidding and renegotiation of power contracts in India. Econometrica 88 (2), 383–424.
Stantcheva, S., 2023. How to run surveys: A guide to creating your own identifying variation and revealing the invisible. Annu. Rev. Econ. 15, 205–234.
Stentoft Arlbjørn, J., Vagn Freytag, P., 2012. Public procurement vs private purchasing: Is there any foundation for comparing and learning across the sectors? Int. J. Public Sector Manag. 25 (3), 203–220.
Szucs, F., 2024. Discretion and favoritism in public procurement. J. Eur. Econom. Assoc. 22 (1), 117–160.
Tepe, M., Prokop, C., 2018. Are future bureaucrats more risk averse? The effect of studying public administration and PSM on risk preferences. J. Public Admin. Res. Theory 28 (2), 182–196.
Weber, M., 1921. Economy and Society: An Outline of Interpretive Sociology (1921; German: Wirtschaft und Gesellschaft. Grundrißder Verstehenden Soziologie), vol. 1, Univ of California Press (1978).
Xu, G., 2018. The costs of patronage: Evidence from the British empire. Amer. Econ. Rev. 108 (11), 3170–3198.