

What Are the Priorities of Bureaucrats? Evidence from Conjoint Experiments with Procurement Officials

Janne Tukiainen, Sebastian Blesse, Albrecht Bohne, Leonardo M. Giuffrida, Jan Jääskeläinen, Ari Luukinen, Antti Sieppi

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

What Are the Priorities of Bureaucrats?

Evidence from Conjoint Experiments with Procurement Officials

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 but not 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 play minor roles. Personal preferences and office interests appear well-aligned among bureaucrats.

JEL-Codes: D730, D900, H110, H570, H830, K410, M540.

Keywords: bureaucrats, public procurement, preferences, conjoint experiments.

*Janne Tukiainen**
University of Turku / Finland
janne.tukiainen@utu.fi

Sebastian Blesse
Ludwig Erhard ifo Center for Social Market
Economy and Institutional Economics
blesse@ifo.de

Albrecht Bohne
ZEW Mannheim / Germany
albrecht.bohne@zew.de

Leonardo M. Giuffrida
ZEW Mannheim / Germany
leonardo.giuffrida@zew.de

Jan Jääskeläinen
Aalto University / Helsinki / Finland
jan.jaaskelainen@kkv.fi

Ari Luukinen
FCCA

Antti Sieppi
FCCA

*corresponding author

This version: December 6, 2022

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, Vítězslav 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 Vergabernetzwerk (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. 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>).

1 Introduction

The functioning of public bureaucracies is considered a principal driver of state capacity, an observation going back to Weber (1921). Arguably, the most important input into bureaucracies is individual bureaucrats and their behavior. However, this insight has received only little attention in economics as a stand-alone aspect of government effectiveness (Besley et al., 2021). Contrary to conventional wisdom, public servants usually possess sizeable leeway in their daily work. This consideration puts the spotlight on bureaucratic decision-making, for which recent evidence points to the benefits of granting discretion, particularly in high public-sector-capacity countries (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 on 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, such incentives are usually weak in the public sector (Burgess and Ratto, 2003; Prendergast, 2007) or draw on non-standard mechanisms such as seniority (Bertrand et al., 2019) or locational postings (Khan et al., 2019). Recent evidence, however, points to the substantial intrinsic motivation of public sector employees (Friebel et al., 2019; Ash and MacLeod, 2015). With this in mind, particularly when considering settings with large discretion, we must look towards other fundamentals to understand the bureaucratic decision-making process. Learning about bureaucratic preferences is especially informative in this regard insofar as bureaucrats may not represent “*Weberian machines*” and their preferences represent a key element of their daily choices at work. Moreover, understanding whether and how bureaucratic preferences align with policy goals would help better understand frictions for government effectiveness. Policy makers may then, in turn, improve the selection of unaligned bureaucrats, increase motivation at the workplace if need be, or even implement policies that find the support of expert bureaucrats.

This paper identifies bureaucratic preferences through repeated online choice experiments with real-life bureaucrats. Our unique conjoint survey experiment isolates these underlying preferences from real-world constraints, which usually confound observational data. This causal evidence on the preferences of public employees contributes to unpacking bureaucratic choices, which are often still viewed as a “black box”.

The setting of our study is public procurement, one of the largest sectors of government activity and with some of the most frequent and high-stakes decisions made in the public

¹There 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).

sector.² Due to its cross-sectoral and cross-institutional nature, public procurement represents an ideal laboratory to study bureaucratic decision-making more generally.

We field our survey experiment with procurement officials (POs) in Finland and Germany, allowing for cross-country comparisons between two high public-sector-capacity settings.³ Investigating bureaucratic priorities is very relevant in our context since POs in both countries report low extrinsic incentives and sizeable discretion, but nonetheless view rigid regulation as one of the largest obstacles to desirable work outcomes.

Studying the “anatomy” of bureaucratic choices is difficult in general since observational data typically only show aggregate realizations of procurement decisions among different actors. Observational data on procurement outcomes may mask POs’ true preferences. Specifically, tender outcomes are affected by the choices of both POs and bidders.⁴ Furthermore, 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. In a nutshell, POs can affect the outcome of a procurement, but typically only indirectly. To gain insights into the preferences underlying bureaucratic behavior, we field so-called conjoint survey experiments (e.g., Hainmueller et al., 2014; Bansak et al., 2019), a specific type of stated preference experiment (see more details below).⁵ Specifically, we make real-life POs repeatedly choose between pairs of hypothetical

²For 2018, the expenditures for public procurement of the median OECD country amount to 13% of GDP and 41% of total government spending (OECD, 2019).

³The public procurement market is relatively larger in Germany and Finland than, for instance, in the US, representing 17% of GDP and 33% of government spending against 10% and 28%, respectively (OECD, 2019). Both contexts share good regulatory practice, average levels of discretion for officers and high process efficiency and quality (Bosio et al., 2022). For both countries motivation to work is similarly higher in the public sector as compared to private works (Cowley and Smith, 2014).

⁴Observable 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 choose 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 rule for a specific tender, or make mistakes in specifying the object or the rules. While structural models of observable auction data allow to estimate primitives such as bidder’s production or entry costs (Gentry et al., 2018), they would not help analyze preferences over tender outcomes.

⁵A huge benefit in comparison 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. 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, i.e. holding other quality criteria and expected prices 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.

tender outcome scenarios that randomly differ across various relevant attributes. The strength of such a survey approach—in contrast to a revealed preferences approach—lies in the fact that we can decompose and quantify the relative importance of single tender attributes for POs in their multi-faceted choice framework. The results, in turn, provide a clear characterization of the preferences underlying bureaucratic choices.

Every choice scenario randomly varies the following attributes: the price in the selected bid, the quality of the good or service (as promised in the selected bid), the degree of competition (as measured by 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. The framing of the choices is such that POs should consider 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. This timing allows us to analyze key 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). We focus on tender outcomes rather than direct choices regarding, e.g., tender procedures and auction mechanisms, because the final outcome of implemented projects is the most relevant margin for the general public. This setting is also very realistic and familiar for the respondents, as the end of the tender process is a natural stage to reflect upon their performance in their daily work. While the respondents' daily work may not involve such a direct choice over some of the attributes (e.g., the extent of competition), they typically make choices that can influence this attribute (e.g., marketing and contract planning).

Conjoint experiments originate in marketing, were formalized in political science (Hainmueller et al., 2014)⁶ and 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. Being agnostic about these assumptions is key since we provide first-time evidence on the priorities of real-life public buyers. Second, conjoint experiments 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 observed choices are driven by a given level of an attribute compared with an omitted base category of the same

⁶Among others, conjoint experiments 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).

attribute—e.g., comparing a higher than expected price to the expected benchmark price of the PO. Key to identification is a full randomization of all attribute levels across choice scenarios and respondents. Conjoint analysis has been validated and shown to predict real behavior (Hainmueller et al., 2015). Moreover, conjoint analysis alleviates concerns about experimenter demand bias and increases realism because subjects simultaneously trade-off various relevant factors of the tendering process. As is typical for conjoint analysis, we do not provide monetary incentives because we do not want to encourage any particular choices over others. Also, POs in our context do not have distinct pecuniary incentives for certain 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.

We contacted about 8,500 real-life POs and received 933 completed answers in both Finland and Germany. Notably, the Finnish sample is broadly representative based on detailed administrative data. We also elicited socio-demographics, task structures, workplace features, procurement practices, and typical outcomes of the POs’ and their offices. At the end of the survey, we analyze how personal preferences and career incentives (i.e., perceived office interests) are aligned regarding the outcomes of the tender process.

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 can be explained by a “*prevention focus*” of bureaucrats and has been 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, may adopt such a view and 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. Winner’s 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 bad-reputation winner. The results concerning competition show that POs have non-linear preferences towards additional bid submissions with a flattening curve after about four offers. We also show that litigation resulting from the bid process 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.

We find very robust treatment effects between the different countries, socio-demographics, task, and office-level features. We argue that bureaucrats working in different public pro-

curement tasks and contexts may have some intrinsic motivation driving their preferences when choosing tender outcomes, and thus their work results. An alternative explanation is that strong and identical institutional constraints are in place both in Finland and Germany. However, this is a less likely explanation as the respondents perceive their institutional environment somewhat differently in the two countries even within the same EU regulatory framework (see details in Section 2.4 and 4.1). Neither explicit incentives nor career concerns appear to be a main motive of bureaucratic decision-making in our context. While office-size and procurement practices do not appear to explain our findings either, we cannot entirely rule out some explanatory power of office culture or peer effects. Notably, we find in further descriptive analysis that personal preferences are highly aligned with (perceived) office interests which indicates efficient but conservative bureaucracies (Besley and Ghatak, 2005). This is also in line with a prevention focus of agents for potential harm to their organization (see above). However, this alignment also means that this study is ultimately unable to show how bureaucratic preferences are formed and what role internal compared to external forces play.

Related Literature and Implications. 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., 2021), 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 set-up. Contrary to the previous contributions mostly targeting developing settings, we focus on two high-capacity countries.

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

⁷Topics 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; Duffo et al., 2018), and determinants of the bureaucratic selection process (Ashraf et al., 2020; Hanna and Wang, 2017; Xu, 2018; Colonnelli et al., 2020).

⁸Studies 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).

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” (e.g., “Buy American Act”), political connections (Szucs, 2018; Baltrunaite, 2019; Baltrunaite et al., 2021; Ryan, 2020; Baranek and Titl, 2020), 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 litigation from the awarding process, but only to a small extent.

Lastly, our paper contributes to recent 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 a number of papers studying survey data using special samples such as politicians (e.g. Broockman and Skovron, 2018), academics (Andre et al., 2022; DellaVigna and Pope, 2018a,b), economic experts (Gründler and Potrafke, 2020) and firm managers (Coibion et al., 2018; Link et al., 2021). We are the first to conduct choice experiments

⁹This finding stresses how relational contracts based on trust are perceived as efficient governance instruments for non-contractible dimensions (see (Levin, 2003; Andrews and Barron, 2016; Board, 2011), particularly between suppliers and public buyers Calzolari and Spagnolo (2009); Albano et al. (2017).

regarding work outcomes among real-world procurement bureaucrats.¹⁰

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. However, the two countries are comparable for the bureaucratic decision-making process inasmuch 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 minutes. 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

¹⁰Detkova et al. (2021a) and Detkova et al. (2021b) study corruption during the Corona-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.

¹¹To 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.

¹²The 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.

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 Vergabernetzwerk* (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 7,247 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 cannot claim that 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 factors may not considerably 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-public-sector-capacity settings. However, since public procurement is an omnipresent feature associated with all governmental tiers and our respondents span all of these areas, we argue that our results have some degree of applicability to bureaucratic behavior in general above and 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

¹³We 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.

¹⁴For more details on sampling, see Appendix D.

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 surveys 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.

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 large differences in the institutions 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 grandly 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. Moreover, given the randomization of multiple attributes and the subsequent difficulty to respond strategically, conjoint experiments have been shown to reduce social desirability bias (Hainmueller et al., 2014; Bansak et al., 2019; Horiuchi et al., 2020).

Figure 1: Example of a Conjoint Scenario

Please look at the following pair of hypothetical tender outcome scenarios carefully and make a decision which you would like more.

Which tender outcome scenario do you prefer?

	Tender outcome A	Tender outcome B
The selected winner is	a firm I already know from previous tenders and trust	a firm that was unknown to me through previous tenders
After awarding the contract, was a legal complaint filed against the tender?	No	No
The tender received	4 bids	8 bids
The quality of the purchase as promised in the winning bid is	a bit better than I expected	much better than I expected
The selected winner is	a local bidder from your region	a non-local bidder that does not come from your region
The price as stated in the winning bid is	much higher than I expected	much lower than I expected
	<input type="radio"/>	<input type="radio"/>

[Next](#)

Notes: This figure illustrates the binary comparisons presented to respondents in the conjoint experiments.

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

¹⁵The number of choice tasks was chosen conservatively in order 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).

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. Figure 1 gives an example of one possible pair of outcome scenarios (i.e., a pair of cards or profiles) that a respondent chooses over.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.

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 our attributes and their realizations are well-founded in the literature on public procurement. 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 hypothesize 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. Also, policymakers generally view competition among poten-

¹⁶Specifically, 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 Figure 1 for the exact question and a graphical illustration of an exemplary card.

¹⁷Including 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) since 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).

tial suppliers as an instrument to increase efficiency and few bids may be viewed with concern (e.g., Fazekas, 2017). Thus, we would expect public buyers to favor competition, at least to some extent, as the increased administrative burden of screening additional bids may lead to a trade-off (Coviello et al., 2018a). Also, 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 vastly employed for supplier selection in private procurement. Public regulations, instead, restrict buyer’s leeway to exclude suppliers with a poor track record to reduce room for corruption. Nonetheless, the use of reputational metrics is found to improve performances (Butler et al., 2020; Decarolis et al., 2016). Against this background, we conjecture that public buyers have a preference for suppliers they have worked with and can trust. Moreover, interviews conducted with procurers highlight how litigation exposure is a major concern (Halonen and Tukiainen, 2020), although empirically litigations are not supposed to introduce much wastefulness when the judicial system is well-functioning overall (Decarolis et al., 2020b)—like in our setting. We want to test for PO’s perception of the litigation issue in their environment and expect that public buyers favor procurement outcomes without realized litigations. Finally, lawmakers usually set regulations favoring local players in the national procurement market, mostly due to small business concerns. 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 all the other tender features, in particular the quality and buyer-supplier relationship. We expect that public buyers have a preference for local suppliers in their vicinity.

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 need 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.

Choice of attribute realizations. We randomly drew realizations of our attributes, and we 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 interpret that we would have a specific interest in extracting preferences for some specific attributes is low due to equal probabilities and other properties of conjoint experiment design. However, one might be concerned that our independent

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.

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. However, several of our design decisions 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 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 at different odds. Therefore, we also did not exclude any constellation of attribute levels *ex ante* from the choice set.¹⁸

¹⁸According 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*.

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 (“what I expected” and “a firm that was unknown to me through previous tenders”). 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).

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 multidisciplinary 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 consists of 2-4 POs in a given 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, this share 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 a 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%). However, the use of awarding procedures is 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 whole EU (3 in the period 2007-2017) and the US federal procurement (2 in the period 2008-2018).¹⁹ On average, 39% of participants in both countries have managed tendering processes 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 $m = 1, \dots, 12$ 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

¹⁹Figures 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/>).

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 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.

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 likely not 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

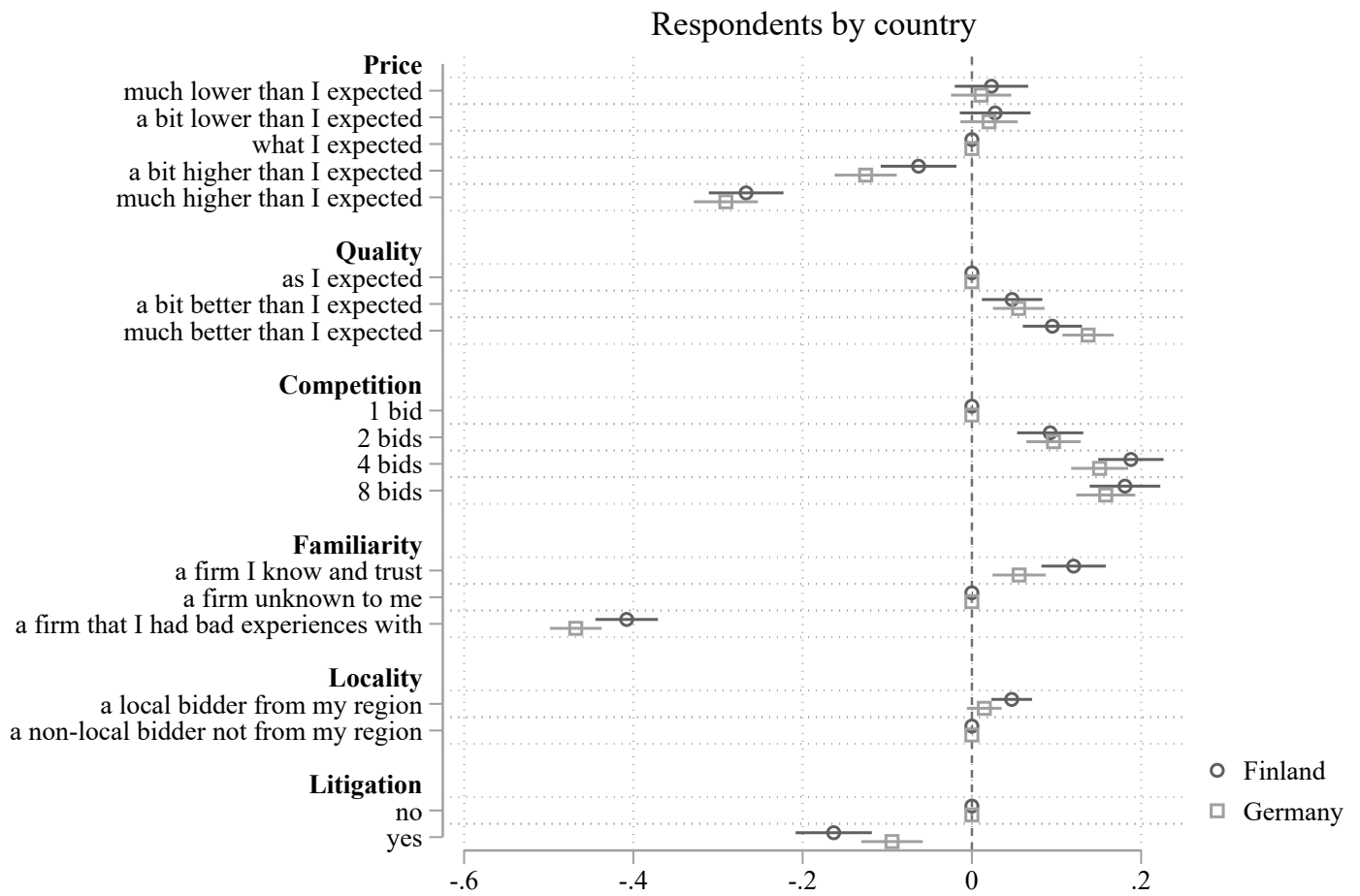
Figure 2 illustrates the main experimental results across both countries. As previously mentioned, we estimate the average causal effect of treatment attribute levels compared to the baseline attribute level—say, much higher than expected prices compared to prices as expected—on the probability of choosing 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 (4,836 for Finland and 6,360 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 facing positive related realizations when deciding which tender outcomes they deem desirable. 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.²⁰

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. Put in comparison in the pooled sample, unexpectedly high price-hikes—compared to expected prices—are approximately two-thirds as disliked as a bad-reputation winner—compared to an unknown winner. Positive realizations of the

²⁰This 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.

Figure 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 Equation 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. .

quality attribute appear to have significant positive effects on support but account for only a fraction of the negative price or reputation concerns.

Bureaucrats also seem to value a certain degree of competition in calls for tender, 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 price and quality 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, assuming the respondents correctly understood the conditioning on prices and quality, the observed effects should be interpreted to relate to the value of competition per se, such as administrative costs or an intrinsic value of procurement performance. Given that administrative costs associated with more bids should lead to a preference for less competition, the observed effects seem to indicate a strong (but rapidly decreasing) intrinsic valuation of POs towards competition.²¹

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.²²

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 economically not meaningful. At first glance, this may be in contrast to observations of strong regionalism in procurement activity. However, correlations be-

²¹The buyer's perception that adding an additional bidder reduces expected welfare is also consistent with the results of auction theory. See, e.g. Compte and Jehiel (2002).

²²Field 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 confirms 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.

tween 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; Baranek and Titl, 2020). 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.

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 are arguably low in the given context, but 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 a general office culture—not necessarily only within 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 these 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. All in all, these results point to very robust priorities among procurement bureaucrats and that mission-based preferences may be a mechanism at play. 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 that 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 display that somewhat different views on national procurement work practices—on top of somewhat different institutional environment 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 different extent across the two countries.²³ In particular, 82% of Finnish and 50% of German respondents feel that they can influence tender outcomes “very much” or “somewhat”. The higher discretion of Finnish respondents also translates into a lower share reporting to have “no influence” (4% in Finland vs. 15% in Germany).²⁴ Moreover, we asked contracting officials 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 recent findings on the benefits of bureaucratic autonomy in public contracting (Carril, 2021), which respondents seem to internalize in our setting.²⁵ 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.²⁶

4.2 Lack of Heterogeneity

As already highlighted in Section 3.2, we observe a striking similarity in results between Finland and Germany. This section documents that this similarity in our experimental

²³See Q10 in Appendix F.

²⁴To 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.

²⁵The 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.

²⁶To 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.

findings also extends to almost all observable features of respondents and their organizations. Investigating such differences in responses to the treatments 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., 2017).²⁷ 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 along with observable characteristics in terms of socio-demographics, task structures, and office-level characteristics. This part of our analysis is explorative in nature since we have no ex-ante expectations about potential differences in the direction or strength of the effects. There are three important caveats to this part of our analysis. First, the background characteristics are evidently 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.²⁸

²⁷Even early work already pointed out that task structures and office characteristics may affect bureaucrat behavior (Christensen and Lægveid, 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.

²⁸For example, in Figure 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.5, 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.

In particular, we analyze heterogeneities based on respondent’s age, gender, education level, education type, work experience, and exposure to EU regulation. Regarding the task structure, some of the heterogeneities we analyze include the type of most common purchase (i.e., goods, services, construction), 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.²⁹ 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.6 in Appendix A). This is in line with a high job stability pattern of POs, as shown in Section 2.4. 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 out-

²⁹This 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.

comes as important for their career prospects, we find no sizeable differences between the marginal means of any attributes (Figure A.7 in Appendix A). This even holds when only considering respondents with the strongest or weakest career concerns (Figure A.8 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.9 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 responses in our experiment due to data protection issues (see more on this in Section B), we cannot directly account for office-level norms by aggregating multiple survey responses per office or including office-level fixed effects in our analysis. However, our survey asks for staff (office) size which should well 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, running heterogeneity analysis using subsamples of different office sizes does not show that different office sizes matter for our experimental results (see Figure A.10 in Appendix A). Even buyers in single-officer units do not have different priorities from 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 negligible role.

4.4 Alignment of Personal Preferences with Office–Level Interests

As previously described in Section 2.2, we asked respondents follow-up questions after the experiment in order to distinguish how personal 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 im-

portance to the office to make a career. Note that attributes are only evaluated with respect to their perceived importance to the organization if respondents state that their career prospects indeed depend on procurement outcomes. As stated before, this represents only a minority of overall answers (241/933, i.e., 25.8%). Subsequently, though, only these answers can be used to compare the alignment of personal and perceived office-level preferences across attributes. 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). However, due to a lack of data we cannot explicitly show this alignment of preferences also for officers who do not report extrinsic motivation. Our finding of a prevention focus of public buyers also supports a substantial alignment of personal and office-level interests among bureaucrats in procurement in our setting.

Given the lack of heterogeneous experimental results across different institutional contexts and the fact that neither incentive schemes nor office norms seem to substantially matter, the strong correspondence between bureaucrats’ individual preferences and the perceived interests of their respective offices may in fact suggest a role of mission-based motivation in our sample. However, this alignment also means that we are unable to show how bureaucratic preferences are formed and what role external forces play. Essentially, we cannot say whether the observation of mission alignment of bureaucrats comes from intrinsic preferences or whether officers merely know what is expected of them in their office not necessarily to advance in career but also to avoid risks.

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 when performing complex work tasks. 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 can be viewed as abductive (Heckman and Singer, 2017) or explorative, as we did not have pre-determined hypotheses for how strong the preferences over the attributes are in relation to each other and what to expect from most of the heterogeneity analysis. Therefore, 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 on bureaucrats' preferences. 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 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 for harm to their organization. This is also in line with findings that exceeding the budget is bad but coming under the budget is not necessarily good (Liebman and Mahoney, 2017). When viewing bureaucrats as agents of a principal such as the general public, a behavior consistent with prevention focus of negative outcomes in prices is likely not in its interest; that is if we assume the general public to value gains and losses (averaged over many tenders) equally and to be risk-averse as stated in Alesina and Tabellini (2008). Second, for POs, it is more important to avoid bidders with bad reputation than to elude unexpectedly high prices. This result suggests that POs would welcome the possibility of reward or punishment during the selection process based on past performance, as proposed by Decarolis et al. (2016) and Butler et al. (2020), but is often not implemented in practice for transparency and competition concerns. Third, we show that POs value a certain degree of competition while litigation concerns and regional favoritism are secondary dimensions in the evaluation of procurement outcomes.

Assuming that organizations with matched preferences of public servants and their respective offices may indeed lead to a more productive public sector (Besley and Ghatak, 2005), aligned missions are preferable from a societal point of view. While our analysis shows that priorities over procurement outcomes are indeed aligned for a majority of officers and offices, governments may even want to take up policies for which there is both expert support and evidence of policy effectiveness, such as the installment of past performance criteria for supplier selection in low corruption contexts (see above). Our paper specifically points to such support from procurement experts.

Our findings may have external validity towards procurement practices in countries with similarly high levels of public sector capacity (Bosio et al., 2022). Moreover, some of our findings provide general insights into public bureaucracies and contexts in which bureaucrats face similar environments of low incentives and sufficient autonomy. Several factors corroborate this conclusion. First, by comparing our survey data to the universe of Finnish procurement contracts from administrative data sources, we find our survey broadly representative regarding various observables, including regions, office types,

contract procedures, awarding mechanisms, and the typical number of bidders. Second, procurement is *per se* representative of a large share of bureaucrats. A considerable proportion of public employees will, at some point in their careers, be involved in the acquisition process since a large share of administrative tasks are in some way related to procurement. Third, our experimental results are robust across countries, socio-demographics, and job and workplace characteristics (e.g., government tiers). Fourth, even though the EU regulatory framework provides an umbrella for general procurement rules, Finnish and German bureaucrats comprise a heterogeneous set of individual buyers across a broad range of individual tasks, demographic features, and office characteristics.

Given the observed lack of career and pay incentives, no discernible effects of bureau size as a proxy for office-level norms on experimental choices, as well as the fact that personal and (perceived) office interests are largely aligned, our findings provide suggestive evidence for the fact that bureaucratic decision-making in our setting may be based on mission-based preferences of public buyers. These arguments align well with the findings from the World Value Survey regarding high levels of mission-motivation among public servants in Finland and Germany (Cowley and Smith, 2014). Nonetheless, we cannot entirely rule out the role of the regulatory framework as an alternative explanation. However, the latter explanation is unlikely as respondents perceived some differences in the regulatory environment in Germany and Finland yet responded almost identically. Also, office culture or peer effects may be an explanation to some degree, even though neither office size nor procurement practices appear to drive our estimates.

Our work opens the door to new avenues of organizational research in the public sector more generally and in public procurement more specifically. For example, future work should tease out the relative importance of different motives and channels in bureaucratic preferences and show their relevance for economic outcomes. For instance, one may conduct similar experiments in more countries and for different types of bureaucrats to isolate intrinsic motivation from external forces of peers or the office as well as other channels for bureaucratic decision-making. Specifically for the context of public procurement, our evidence indicates, for instance, that the lack of competition observed in many settings is unlikely to dominantly arise from the preferences of public buyers because they appear to value competition in our experiment more than is typically observed in the real world. Thus, future research could address the relevant entry barriers for firms in public procurement markets in more detail and how they can be reduced.

References

- Aghion, P. and Tirole, J. (1997). Formal and Real Authority in Organizations. *Journal of Political Economy*, 105(1):1–29.

- Albano, G. L., Cesi, B., and Iozzi, A. (2017). Teaching an Old Dog a New Trick: Reserve Price and Unverifiable Quality in Repeated Procurement. CEIS Working Paper No. 404.
- Alesina, A. and Tabellini, G. (2008). Bureaucrats or Politicians? Part II: Multiple Policy Tasks. *Journal of Public Economics*, 92(3):426–447.
- Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D., and Tonetti, C. (2020). Long-term-care Utility and Late-in-life Saving. *Journal of Political Economy*, 128(6):2375–2451.
- Andre, P., Pizzinelli, C., Roth, C., and Wohlfart, J. (2022). Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples. *The Review of Economic Studies*. Forthcoming.
- Andrews, I. and Barron, D. (2016). The Allocation of Future Business: Dynamic Relational Contracts with Multiple Agents. *American Economic Review*, 106(9):2742–59.
- Ash, E. and MacLeod, W. B. (2015). Intrinsic Motivation in Public Service: Theory and Evidence from State Supreme Courts. *The Journal of Law and Economics*, 58(4):863–913.
- Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. S. (2020). Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. *American Economic Review*, 110(5):1355–94.
- Asker, J. and Cantillon, E. (2010). Procurement When Price and Quality Matter. *The RAND Journal of Economics*, 41(1):1–34.
- Baltrunaite, A. (2019). Political Contributions and Public Procurement: Evidence from Lithuania. *Journal of the European Economic Association*, 18(2):541–582.
- Baltrunaite, A., Giorgiantonio, C., Mocetti, S., and Orlando, T. (2021). Discretion and Supplier Selection in Public Procurement. *The Journal of Law, Economics, and Organization*, 37(1):134–166.
- Bandiera, O., Best, M. C., Khan, A. Q., and Prat, A. (2021). The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats. *The Quarterly Journal of Economics*, 136(4):2195–2242.
- Bansak, K., Hainmueller, J., and 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., and Yamamoto, T. (2018). The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments. *Political Analysis*, 26(1):112–119.

- Bansak, K., Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2019). “Conjoint Survey Experiments” For Druckman, James N., and Donald P. Green, eds. *Cambridge Handbook of Advances in Experimental Political Science*. New York: Cambridge University Press.
- Bansak, K., Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2021). Beyond the Breaking Point? Survey Satisficing in Conjoint Experiments. *Political Science Research and Methods*, 9(1):53–71.
- Baranek, B. and Titl, V. (2020). The Cost of Favoritism in Public Procurement. mimeo.
- Bechtel, M. M., Hainmueller, J., and Margalit, Y. (2014). Preferences for International Redistribution: The Divide over the Eurozone Bailouts. *American Journal of Political Science*, 58(4):835–856.
- Bechtel, M. M. and Scheve, K. F. (2013). Mass Support for Global Climate Agreements Depends on Institutional Design. *Proceedings of the National Academy of Sciences*, 110(34):13763–13768.
- Beetsma, R., Burgoon, B., Nicoli, F., de Ruijter, A., and Vandenbroucke, F. (2022). What kind of EU Fiscal Capacity? Evidence from a Randomized Survey Experiment in Five European Countries in Times of Corona. *Economic Policy*. Forthcoming.
- Berinsky, A. J., Margolis, M. F., and Sances, M. W. (2014). Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-administered Surveys. *American Journal of Political Science*, 58(3):739–753.
- Bertrand, M., Burgess, R., Chawla, A., and Xu, G. (2019). The Glittering Prizes: Career Incentives and Bureaucrat Performance. *The Review of Economic Studies*, 87(2):626–655.
- Besley, T. and Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *American Economic Review*, 95(3):616–636.
- Besley, T. J., Burgess, R., Khan, A., and Xu, G. (2021). Bureaucracy and Development. NBER Working Paper No. 29163.
- Besley, T. J. and Persson, T. (2022). Organizational Dynamics: Culture, Design, and Performance. CEPR Discussion Paper No. 17382.
- Best, M. C., Hjort, J., and Szakonyi, D. (2017). Individuals and Organizations as Sources of State Effectiveness. NBER Working Paper No. 23350.
- Board, S. (2011). Relational Contracts and the Value of Loyalty. *American Economic Review*, 101(7):3349–67.

- Boland, M. and Godsell, D. (2021). Bureaucratic Discretion and Contracting Outcomes. *Accounting, Organizations and Society*, 88:101–173.
- Bosio, E., Djankov, S., Glaeser, E., and Shleifer, A. (2022). Public Procurement in Law and Practice. *American Economic Review*, 112(4):1091–1117.
- Bradbury, M. and Kellough, J. E. (2011). Representative Bureaucracy: Assessing the Evidence on Active Representation. *The American Review of Public Administration*, 41(2):157–167.
- Broockman, D. E. and Skovron, C. (2018). Bias in Perceptions of Public Opinion among Political Elites. *The American Political Science Review*, 112(3):542–563.
- Burgess, S., Propper, C., Ratto, M., and Tominey, E. (2017). Incentives in the Public Sector: Evidence from a Government Agency. *The Economic Journal*, 127(605):117–141.
- Burgess, S. and Ratto, M. (2003). The Role of Incentives in the Public Sector: Issues and Evidence. *Oxford Review of Economic Policy*, 19(2):285–300.
- Butler, J. V., Carbone, E., Conzo, P., and Spagnolo, G. (2020). Past Performance and Entry in Procurement: An Experimental Investigation. *Journal of Economic Behavior & Organization*, 173:179–195.
- Buurman, M., Delfgaauw, J., Dur, R., and Van den Bossche, S. (2012). Public Sector Employees: Risk Averse and Altruistic? *Journal of Economic Behavior & Organization*, 83(3):279–291.
- Calzolari, G. and 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., and Lucifora, C. (2021). Public Sector Jobs: Working in the Public Sector in Europe and the US. IZA Discussion Paper No. 14514.
- Christensen, T. and Læg Reid, P. (2004). Public Administration Research in Norway: Organization Theory, Institutionalism and Empirical Studies in a Democratic Context. *Public Administration*, 82(3):679–690.
- Coibion, O., Gorodnichenko, Y., and Kumar, S. (2018). How Do Firms Form Their Expectations? New Survey Evidence. *American Economic Review*, 108(9):2671–2713.
- Colonnelli, E., Prem, M., and Teso, E. (2020). Patronage and Selection in Public Sector Organizations. *American Economic Review*, 110(10):3071–99.

- Compte, O. and Jehiel, P. (2002). On the Value of Competition in Procurement Auctions. *Econometrica*, 70(1):343–355.
- Coviello, D., Guglielmo, A., and Spagnolo, G. (2018a). The Effect of Discretion on Procurement Performance. *Management Science*, 64(2):715–738.
- Coviello, D., Moretti, L., Spagnolo, G., and Valbonesi, P. (2018b). Court Efficiency and Procurement Performance. *The Scandinavian Journal of Economics*, 120(3):826–858.
- Cowley, E. and 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. and de Quidt, J. (2015). Implicit Preferences Inferred from Choice. mimeo.
- Decarolis, F., de Rassenfosse, G., Giuffrida, L. M., Iossa, E., Mollisi, V., Raiteri, E., and Spagnolo, G. (2021). Buyers’ Role in Innovation Procurement: Evidence from US Military R&D Contracts. *Journal of Economics & Management Strategy*, 30(4):697–720.
- Decarolis, F., Giuffrida, L. M., Iossa, E., Mollisi, V., and Spagnolo, G. (2020a). Bureaucratic Competence and Procurement Outcomes. *The Journal of Law, Economics, and Organization*, 36(3):537–597.
- Decarolis, F., Mattera, G., and Menon, C. (2020b). Delays at the Border: Court Efficiency and Delays in Public Contracts. CEPR Discussion Paper No. 14856.
- Decarolis, F., Spagnolo, G., and Pacini, R. (2016). Past Performance and Procurement Outcomes. NBER Working Paper No. 22814.
- DellaVigna, S. and Pope, D. (2018a). Predicting Experimental Results: Who Knows What? *Journal of Political Economy*, 126(6):2410–2456.
- DellaVigna, S. and Pope, D. (2018b). What Motivates Effort? Evidence and Expert Forecasts. *The Review of Economic Studies*, 85(2):1029–1069.
- Detkova, P., Pronin, P., Tkachenko, A., and Yakovlev, A. A. (2021a). The Changing Perceptions of Corruption During the COVID Pandemic in Russia. mimeo.
- Detkova, P., Tkachenko, A., and Yakovlev, A. (2021b). Gender Heterogeneity of Bureaucrats in Attitude to Corruption: Evidence from List Experiment. *Journal of Economic Behavior & Organization*, 189:217–233.
- Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2018). The Value of Regulatory Discretion: Estimates From Environmental Inspections in India. *Econometrica*, 86(6):2123–2160.

- Ellingsen, T. and Johannesson, M. (2008). Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review*, 98(3):990–1008.
- Fazekas, M. (2017). Assessing the Quality of Government at the Regional Level Using Public Procurement Data. European Commission, Directorate-General for Regional Policy: Working Paper No. 12-2017.
- Finan, F., Olken, B. A., and Pande, R. (2017). The Personnel Economics of the Developing State. *Handbook of Economic Field Experiments*, 2:467–514.
- Friebel, G., Kosfeld, M., and Thielmann, G. (2019). Trust the Police? Self-selection of Motivated Agents into the German Police Force. *American Economic Journal: Microeconomics*, 11(4):59–78.
- Fuster, A. and Zafar, B. (2021). The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey. *American Economic Journal: Economic Policy*, 13(1):231–65.
- Gentry, M. L., Hubbard, T. P., Nekipelov, D., Paarsch, H. J., et al. (2018). Structural Econometrics of Auctions: A Review. *Foundations and Trends in Econometrics*, 9(2-4):79–302.
- Geys, B., Heggedal, T.-R., and Sørensen, R. J. (2017). Are Bureaucrats Paid Like CEOs? Performance Compensation and Turnover of Top Civil Servants. *Journal of Public Economics*, 152:47–54.
- Gregg, P., Grout, P. A., Ratcliffe, A., Smith, S., and Windmeijer, F. (2011). How Important is Pro-social Behaviour in the Delivery of Public Services? *Journal of Public Economics*, 95(7-8):758–766.
- Gründler, K. and Potrafke, N. (2020). Experts and Epidemics. CESifo Working Paper No. 8556.
- Haaland, I., Roth, C., and Wohlfart, J. (2021). Designing Information Provision Experiments. *Journal of Economic Literature*, Forthcoming.
- Hainmueller, J., Hangartner, D., and Yamamoto, T. (2015). Validating Vignette and Conjoint Survey Experiments against Real-world Behavior. *Proceedings of the National Academy of Sciences*, 112(8):2395–2400.
- Hainmueller, J. and Hopkins, D. J. (2015). The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes toward Immigrants. *American Journal of Political Science*, 59(3):529–548.

- Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2014). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political Analysis*, 22(1):1–30.
- Halonen, K.-M. and Tukiainen, J. (2020). Competition and Litigation in Swedish Public Procurement. Swedish Competition Authority Report 2020:1.
- Hanna, R. and Wang, S.-Y. (2017). Dishonesty and Selection Into Public Service: Evidence from India. *American Economic Journal: Economic Policy*, 9(3):262–90.
- Heckman, J. J. and Singer, B. (2017). Abducting Economics. *American Economic Review*, 107(5):298–302.
- Hong, H. and Shum, M. (2002). Increasing competition and the winner’s curse: Evidence from procurement. *The Review of Economic Studies*, 69(4):871–898.
- Horiuchi, Y., Markovich, Z. D., and Yamamoto, T. (2020). Does Conjoint Analysis Mitigate Social Desirability Bias? MIT Political Science Department Research Paper No. 2018-15.
- Hyytinen, A., Lundberg, S., and Toivanen, O. (2018). Design of Public Procurement Auctions: Evidence from Cleaning Contracts. *The RAND Journal of Economics*, 49(2):398–426.
- Jääskeläinen, J. and Tukiainen, J. (2019). Anatomy of Public Procurement. VATT Institute for Economic Research Working Paper No. 118.
- Jackson, S. E. and Dutton, J. E. (1988). Discerning Threats and Opportunities. *Administrative Science Quarterly*, pages 370–387.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–91.
- Kang, K. and Miller, R. A. (2022). Winning by Default: Why is There So Little Competition in Government Procurement? *The Review of Economic Studies*, 89(3):1495–1556.
- Kessler, J. B., Low, C., and Sullivan, C. D. (2019). Incentivized Resume Rating: Eliciting Employer Preferences without Deception. *American Economic Review*, 109(11):3713–44.
- Khan, A. Q., Khwaja, A. I., and Olken, B. A. (2019). Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats Through Performance-based Postings. *American Economic Review*, 109(1):237–70.

- Kuehnhanss, C. R., Murdoch, Z., Geys, B., and Heyndels, B. (2017). Identity, Threat Aversion, and Civil Servants' Policy Preferences: Evidence from the European Parliament. *Public Administration*, 95(4):1009–1025.
- Kutlina-Dimitrova, Z. and Lakatos, C. (2016). Determinants of Direct Cross-border Public Procurement in EU Member States. *Review of World Economics*, 152(3):501–528.
- Leeper, T. J., Hobolt, S. B., and Tilley, J. (2020). Measuring Subgroup Preferences in Conjoint Experiments. *Political Analysis*, 28(2):207–221.
- Levin, J. (2003). Relational Incentive Contracts. *American Economic Review*, 93(3):835–857.
- Liebman, J. B. and Mahoney, N. (2017). Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement. *American Economic Review*, 107(11):3510–49.
- Link, S., Peichl, A., Roth, C., and Wohlfart, J. (2021). Information Frictions Among Firms and Households. CESifo Working Paper No. 8969.
- Meade, A. W. and Craig, S. B. (2012). Identifying Careless Responses in Survey Data. *Psychological Methods*, 17(3):437–455.
- Meyer-Sahling, J., Mikkelsen, K. S., and 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 Administration*, 99(4):740–757.
- OECD (2019). *Reforming Public Procurement*.
- Oliveros, V. and Schuster, C. (2018). Merit, Tenure, and Bureaucratic Behavior: Evidence from a Conjoint Experiment in the Dominican Republic. *Comparative Political Studies*, 51(6):759–792.
- Pinkse, J. and Tan, G. (2005). The affiliation effect in first-price auctions. *Econometrica*, 73(1):263–277.
- Prendergast, C. (2007). The Motivation and Bias of Bureaucrats. *American Economic Review*, 97(1):180–196.
- Rakotonarivo, O. S., Schaafsma, M., and Hockley, N. (2016). A Systematic Review of the Reliability and Validity of Discrete Choice Experiments in Valuing Non-market Environmental Goods. *Journal of Environmental Management*, 183:98–109.

- Rasul, I. and Rogger, D. (2018). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*, 128(608):413–446.
- Rasul, I., Rogger, D., and Williams, M. J. (2020). Management, Organizational Performance, and Task Clarity: Evidence from Ghana’s Civil Service. *Journal of Public Administration Research and Theory*.
- Rozin, P. and Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, 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.
- Szucs, F. (2018). Discretion and Corruption in Public Procurement. mimeo.
- Tepe, M. and Prokop, C. (2018). Are Future Bureaucrats More Risk Averse? The Effect of Studying Public Administration and PSM on Risk Preferences. *Journal of Public Administration Research and Theory*, 28(2):182–196.
- Weber, M. (1921). *Economy and Society: An Outline of Interpretive Sociology (1921; German: Wirtschaft und Gesellschaft. Grundriß der Verstehenden Soziologie)*, volume 1. Univ of California Press (1978).
- Westfall, P. H. and Young, S. S. (1993). *Resampling-based Multiple Testing: Examples and Methods for p-value Adjustment*, volume 279. John Wiley & Sons.
- Xu, G. (2018). The Costs of Patronage: Evidence from the British Empire. *American Economic Review*, 108(11):3170–98.

A Appendix: Additional Tables and Figures

Table A.1: Socio-demographic characteristics of respondents

	Pooled sample			Germany			Finland		
	n	Mean/Prop.	SD	n	Mean/Prop.	SD	n	Mean/Prop.	SD
Panel A: Socio-demographic characteristics									
Age group	915			517			398		
20-30 years		.05			.06			.04	
31-40 years		.20			.20			.19	
41-50 years		.30			.28			.31	
51-60 years		.36			.37			.34	
above 60 years		.10			.08			.12	
Female	899	.50	.50	511	.50	.50	388	.49	.50
Master's degree	933	.49	.50	530	.49	.50	403	.49	.50
Education	682			327			355		
Business Administration and Accounting		.18			.15			.20	
Economics and Finance		.03			.01			.04	
Engineering		.34			.25			.41	
Law		.12			.17			.07	
Public administration		.24			.36			.12	
Other		.10			.05			.15	

Notes: This table presents the socio-demographic characteristics for survey respondents for pooled sample as well as German and Finnish sample, respectively.

Table A.2: Workplace characteristics of respondents

	Pooled sample			Germany			Finland		
	n	Mean/Prop.	SD	n	Mean/Prop.	SD	n	Mean/Prop.	SD
Panel B: Workplace characteristics									
Type of office	933			530			403		
Municipality		.40			.43			.36	
State/regional government		.17			.21			.12	
Federal government		.14			.10			.19	
Publicly owned company		.18			.17			.19	
Other		.11			.09			.13	
Organization size	933			530			403		
Just me		.08			.05			.11	
2-4		.29			.24			.36	
5-10		.22			.21			.23	
11-19		.11			.12			.10	
20-99		.19			.22			.15	
More than 100		.11			.16			.05	
Type of position	933			530			403		
Civil servant		.31			.33			.28	
Temporary employment		.04			.02			.07	
Unlimited employment		.62			.62			.61	
Other or prefer not to tell		.03			.02			.04	
Supervisor	917	.40	.49	517	.45	.50	400	.33	.47
Experience in current position (in years)	932	8.49	7.97	529	8.90	8.34	403	7.95	7.44
Experience in procurement (in years)	930	11.58	9.08	529	12.23	9.87	401	10.72	7.83
I am satisfied in competence of my own department	933			530			403		
Strongly agree		.19			.10			.31	
Agree		.48			.59			.33	
Disagree		.28			.27			.29	
Strongly disagree		.06			.04			.08	
Typical procedure used	933			530			403		
Negotiated		.13			.20			.04	
Open		.66			.55			.81	
Restricted		.13			.15			.11	
Sole-source and other		.08			.11			.03	
Typical award mechanism used	933			530			403		
Best price/quality		.56			.57			.55	
Lowest price		.42			.42			.43	
Other		.02			.02			.02	
Share of contracts above EU threshold	923	.42	.49	525	.34	.47	398	.53	.50
Uses secondary objectives	933	.48	.50	530	.55	.50	403	.39	.49
Typical # of bidders	933	18.92	206.45	530	18.86	223.22	403	19.00	182.32
Typical # of bidders (censored at 500)	929	7.70	28.44	528	7.56	31.32	401	7.87	24.17
Litigations in previous year	933	1.27	4.88	530	1.76	4.71	403	.63	5.02
Most common reason for litigation	360			250			110		
A bid protest		.60			.62			.55	
A solicitation challenge		.34			.30			.42	
Other		.06			.08			.04	

Notes: This table presents the workplace characteristics for survey respondents for the pooled sample as well as the German and Finnish sample, respectively. We present a right censored statistic for typical number of bidders to tackle the possible issue of respondent's misunderstanding the question.

Table A.3: Main results

	Joint	Germany	Finland
much lower than I expected	0.0160 (0.0141)	0.0109 (0.0181)	0.0232 (0.0221)
a bit lower than I expected	0.0235* (0.0134)	0.0203 (0.0172)	0.0275 (0.0213)
what I expected	0 (.)	0 (.)	0 (.)
a bit higher than I expected	-0.0982*** (0.0145)	-0.126*** (0.0187)	-0.0630*** (0.0228)
much higher than I expected	-0.280*** (0.0147)	-0.291*** (0.0193)	-0.267*** (0.0224)
as I expected	0 (.)	0 (.)	0 (.)
a bit better than I expected	0.0529*** (0.0118)	0.0553*** (0.0155)	0.0475*** (0.0182)
much better than I expected	0.121*** (0.0117)	0.137*** (0.0154)	0.0951*** (0.0178)
1 bid	0 (.)	0 (.)	0 (.)
2 bids	0.0956*** (0.0127)	0.0965*** (0.0164)	0.0925*** (0.0199)
4 bids	0.166*** (0.0130)	0.151*** (0.0171)	0.188*** (0.0196)
8 bids	0.168*** (0.0137)	0.158*** (0.0178)	0.181*** (0.0212)
a firm I know and trust	0.0841*** (0.0124)	0.0558*** (0.0160)	0.120*** (0.0193)
a firm unknown to me	0 (.)	0 (.)	0 (.)
a firm that I had bad experiences with	-0.442*** (0.0120)	-0.468*** (0.0156)	-0.408*** (0.0188)
a local bidder from my region	0.0288*** (0.00795)	0.0147 (0.0104)	0.0471*** (0.0122)
a non-local bidder not from my region	0 (.)	0 (.)	0 (.)
no	0 (.)	0 (.)	0 (.)
yes	-0.124*** (0.0145)	-0.0944*** (0.0185)	-0.163*** (0.0229)
Constant	0.520*** (0.0162)	0.552*** (0.0205)	0.481*** (0.0256)
Observations	11196	6360	4836
R^2	0.29	0.31	0.27

Notes: This table presents effects of tender outcome attributes on the probability of deciding in favor of a tender outcome. Dependent variable is a binary variable with a value of 1 if profile was chosen and a value of 0 otherwise. Standard errors are in parenthesis. Effects are estimated using OLS and standard errors are clustered at the respondent level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Analysis of perceived discretion and threats

	Perceived discretion	Perceived threats					
		Mistakes by procurer	Budgetary constraints	Issues with winner	Loser challenged	Rigid regulation	Other
Age: 20-30	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (0.0353)
Age: 31-40	0.0782* (0.0123)	-0.0136 (0.106)	0.0417 (0.154)	-0.0917 (0.0872)	-0.0846 (0.121)	-0.175 (0.171)	0.0463 (0.0617)
Age: 41-50	0.0988 (0.0492)	-0.0217 (0.108)	0.00439 (0.0905)	-0.126 (0.0561)	-0.0524 (0.122)	-0.0833 (0.129)	0.0346 (0.0965)
Age: 51-60	0.0349 (0.0104)	-0.0342 (0.0877)	0.0283 (0.129)	-0.174 (0.0901)	-0.106 (0.101)	-0.120 (0.209)	0.0177 (0.136)
Age: above 60	0.0558 (0.0563)	0.00836 (0.148)	0.130 (0.200)	-0.160** (0.00663)	-0.0483 (0.0821)	-0.147 (0.136)	-0.0523 (0.127)
Age: I would prefer not to tell	0.0676 (0.0878)	-0.0239 (0.403)	-0.132 (0.232)	0.0799 (0.398)	0.343* (0.0312)	-0.178 (0.136)	0.383* (0.0549)
Female	-0.00835 (0.0605)	0.00801* (0.000989)	0.0394 (0.0456)	0.0145** (0.000510)	-0.0398 (0.0318)	-0.0688 (0.0448)	-0.0283 (0.0194)
Has master's degree or equivalent	-0.0224 (0.0140)	0.0240 (0.0122)	-0.00326 (0.0336)	-0.0179 (0.00725)	0.0176 (0.0103)	-0.00921 (0.0391)	0.0215 (0.00530)
Municipality	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (0.0353)
State government	0.00516 (0.00478)	0.0233 (0.0135)	-0.00516 (0.0732)	0.0352 (0.0118)	0.00792 (0.00351)	-0.0231 (0.0211)	-0.0415 (0.0286)
Federal government	0.0288 (0.0406)	0.0343 (0.0151)	-0.0295 (0.0358)	0.0339 (0.0503)	-0.0205** (0.000782)	-0.0687 (0.0266)	0.0674 (0.0603)
Public company	0.0369 (0.0403)	-0.0505 (0.0490)	-0.0771 (0.0564)	-0.0870 (0.0194)	-0.0145 (0.0347)	0.0973 (0.0226)	0.00126 (0.0473)
Other	-0.0406 (0.0723)	-0.0869 (0.120)	-0.131* (0.0170)	0.00404 (0.0718)	0.0169 (0.0107)	0.0152 (0.0440)	-0.0655 (0.0618)
Organization size: Just me	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Organization size: 2-4	-0.00125 (0.0137)	0.0149 (0.0481)	-0.0199 (0.0564)	0.0939 (0.0765)	0.00540 (0.0171)	-0.0574 (0.0503)	-0.0403 (0.0294)
Organization size: 5-10	-0.0183 (0.0849)	0.0485 (0.00895)	-0.0262 (0.00732)	0.0564 (0.0173)	0.0207 (0.0409)	-0.0722 (0.0900)	0.0520 (0.0151)
Organization size: 11-19	-0.0576 (0.00957)	0.138* (0.0160)	-0.00724 (0.0107)	0.146** (0.00840)	-0.00407 (0.0170)	-0.0613 (0.0274)	-0.0236 (0.0124)
Organization size: 20-99	-0.106 (0.0659)	0.0852 (0.0408)	-0.0569 (0.0478)	0.0972 (0.0481)	0.00724 (0.00155)	0.0150 (0.0992)	-0.0160 (0.0648)
Organization size: More than 100	-0.0240 (0.0113)	0.107 (0.0541)	-0.0467 (0.0530)	0.121 (0.0697)	0.00190 (0.0167)	0.0620 (0.0791)	0.0348 (0.0416)
Supervisor	0.0349 (0.0739)	0.00693 (0.0222)	-0.0151 (0.0501)	0.0346* (0.00381)	-0.0172 (0.0198)	-0.00280 (0.0421)	0.0194 (0.0859)
Experience in current position	-0.000855 (0.00228)	-0.00672 (0.00222)	0.00388 (0.00244)	0.000466** (0.0000289)	0.000700 (0.00286)	-0.00332 (0.00546)	0.00189 (0.00217)
Overall experience	0.00257 (0.000822)	0.00100 (0.00116)	-0.00261 (0.00340)	0.000630 (0.00298)	0.00188 (0.00237)	0.00167 (0.00247)	0.0000442 (0.00440)
Award mechanism: best price/quality	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Award mechanism: lowest price	-0.130 (0.0480)	0.0208 (0.0700)	0.0371** (0.00204)	0.0745 (0.0284)	-0.0183 (0.0160)	-0.0584 (0.0219)	-0.0478* (0.00548)
Award mechanism: other	0.0617 (0.117)	-0.190** (0.00858)	-0.0352 (0.0167)	-0.214 (0.190)	-0.0141 (0.0904)	-0.204 (0.197)	0.354* (0.0387)
Contract size above EU threshold	-0.00449 (0.0199)	-0.0291 (0.0590)	-0.0270 (0.0541)	0.0314 (0.0127)	0.0698 (0.0162)	-0.0811* (0.0119)	0.00438 (0.0154)
Uses secondary objectives	0.0452 (0.0121)	-0.00580 (0.00917)	0.00179 (0.00244)	0.00634 (0.00377)	0.0336 (0.0238)	0.0263 (0.00923)	-0.00465 (0.0659)
Typical # of bidders	0.000421* (0.0000538)	-0.000595 (0.000288)	0.000209 (0.000206)	0.000712 (0.000133)	0.000124 (0.000544)	-0.000519 (0.000156)	-0.0000246 (0.000263)
Litigated in previous year	0.0489 (0.0714)	0.0653 (0.0984)	-0.0289 (0.0321)	0.0624** (0.00313)	0.249 (0.0733)	-0.0420* (0.00616)	0.0300 (0.0365)
Country: Finland	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Country: Germany	-0.325** (0.00603)	-0.136* (0.0109)	-0.0436 (0.00767)	-0.113** (0.00856)	-0.133 (0.0246)	-0.0108 (0.00784)	0.0991** (0.00158)
Constant	0.787*** (0.00229)	0.346 (0.0923)	0.246*** (0.00187)	0.403 (0.117)	0.155 (0.100)	0.688 (0.0658)	0.203 (0.0790)
Observations	869	869	869	869	869	869	869
R ²	0.147	0.051	0.042	0.046	0.163	0.046	0.051

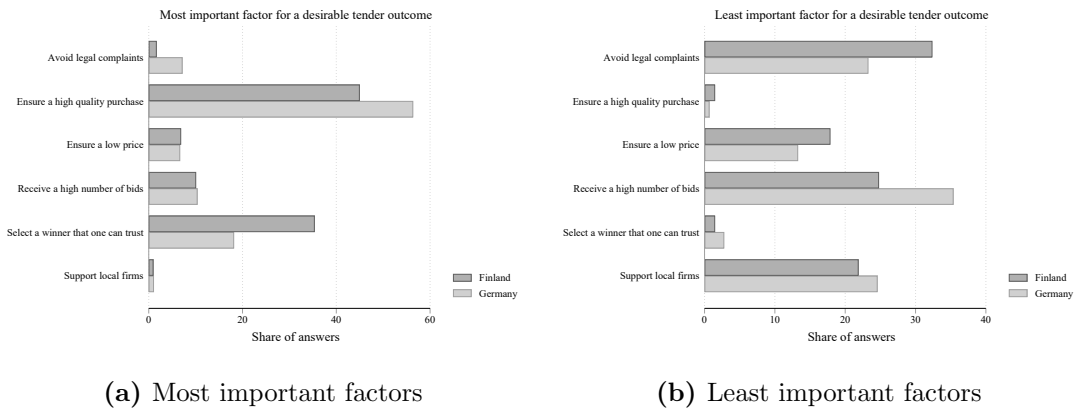
Notes: This table presents results from OLS regressions on the perceived discretion and perceived threats to the procurement process. Dependent variable is a binary variable with a value of 1 if a person perceived to have discretion or perceived a given threat to be an issue for successful procurement, and a value of 0 otherwise. Each respondent was able to acknowledge several threats. Heteroskedasticity-robust standard errors are in parenthesis. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A.5: Comparing personal and office-level interests

Attributes	Fraction of unaligned interests	
	Most important factor	Least important factor
Avoid legal complaints	17%	21%
Ensure a high quality purchase	33%	2%
Ensure a low price	9%	20%
Receive a high number of bids	11%	31%
Select a winner that one can trust	20%	2%
Support local firms	7%	21%

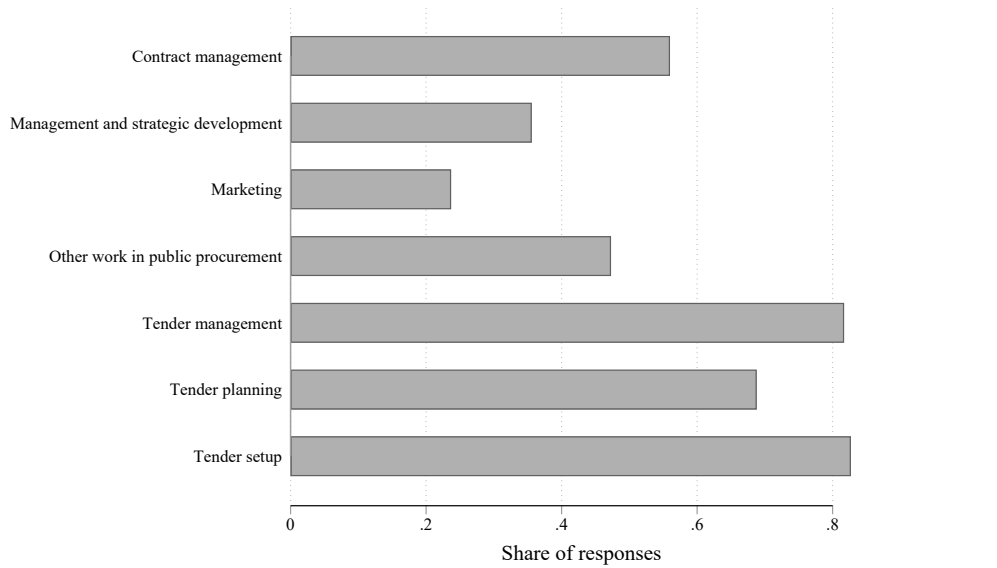
Notes: This table shows the attributes and the evaluation of respondents regarding their personal preferences (Q24) and their perceived importance for the office (Q27) using the follow-up questions. Specifically, the table shows the fraction of buyers for whom there is a discrepancy between the stated personal preference and the perceived importance for the office, separately for the most and least important factor. The comparisons rest on 241 observations.

Figure A.1: Bureaucratic preferences over factors of tender outcomes (Q24)



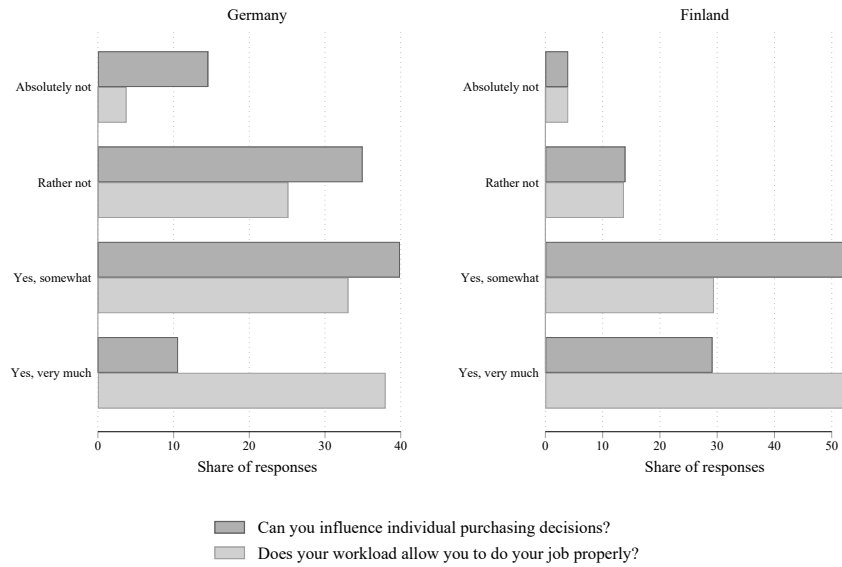
Notes: The figures present what respondents consider (a) most important factors and (b) least important factors for a desirable tender outcome. Respondents were forced to pick only one option in both questions.

Figure A.2: Which steps of the procurement process apply to your job? (Q6)



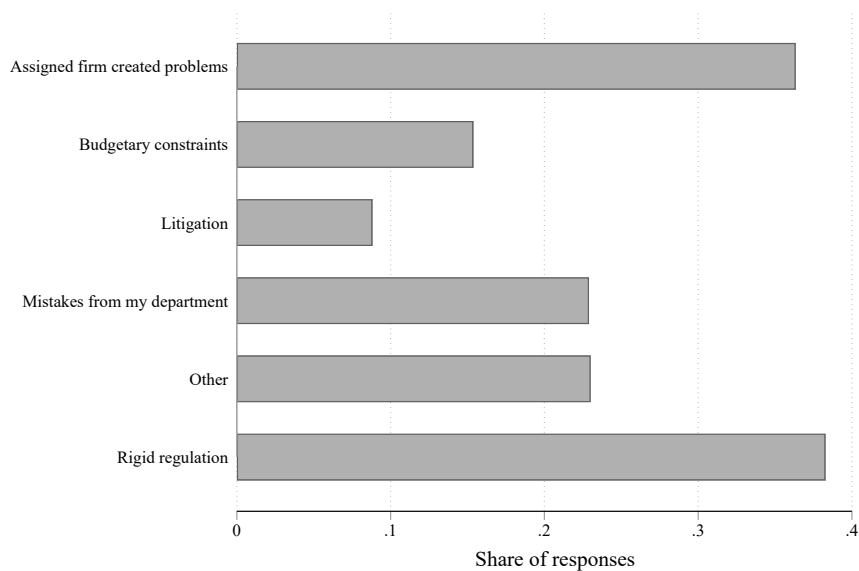
Notes: The figure presents which tasks respondents work with. Respondents were allowed to choose multiple options.

Figure A.3: Distribution of perceived discretion and workload (Q10, Q11)



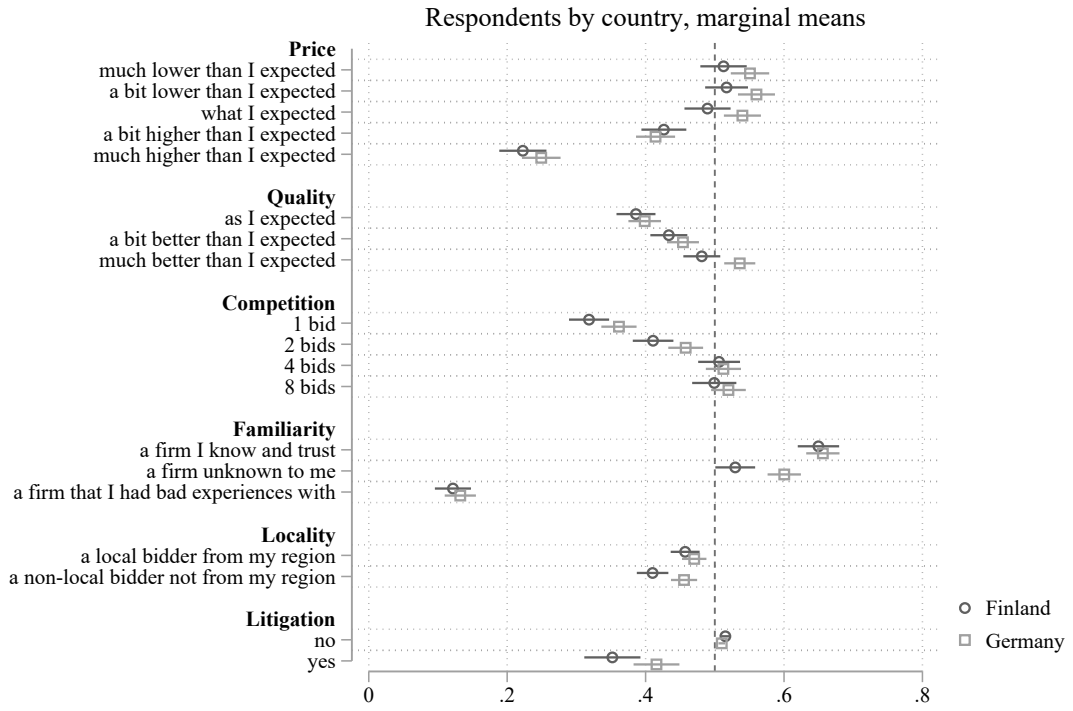
Notes: The figure presents how respondents feel about being able to influence their work as well as their workload. Respondents were forced to pick only one option to either questions.

Figure A.4: Distribution of perceived threats to the procurement process (Q22)



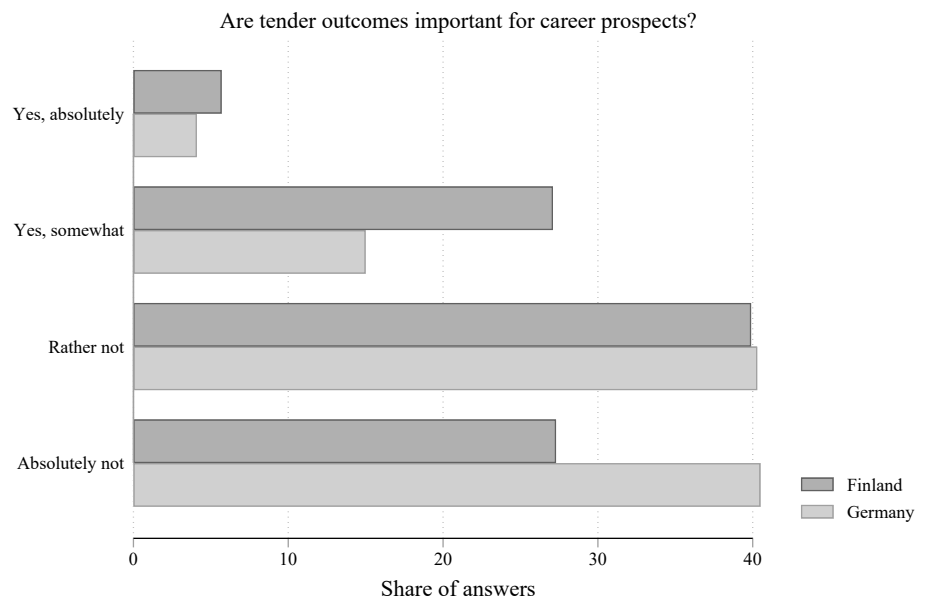
Notes: The figure presents what respondents perceive as threats to procurement process. Respondents were allowed to choose multiple options.

Figure A.5: Marginal means of conjoint experiments per country



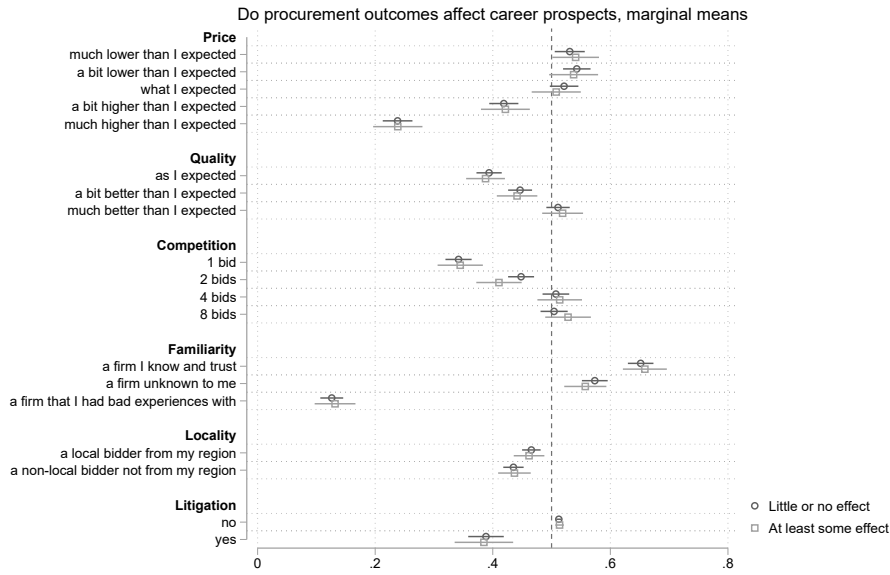
Notes: Marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure A.6: Distribution of perceived career incentives (Q26)



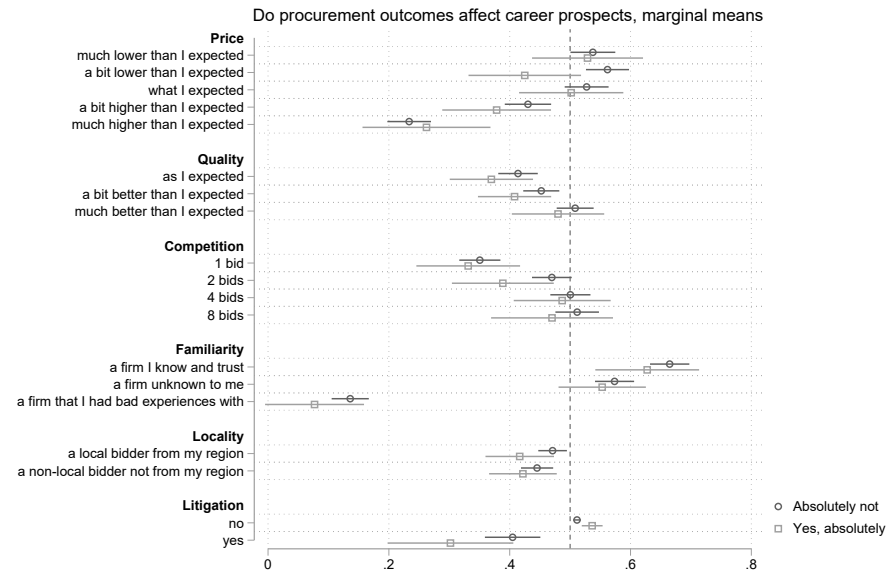
Notes: The figure presents how respondents feel about career prospects. Respondents were forced to pick only one option.

Figure A.7: Do tender outcomes matter for career prospects and what role does it play for buyer priorities?



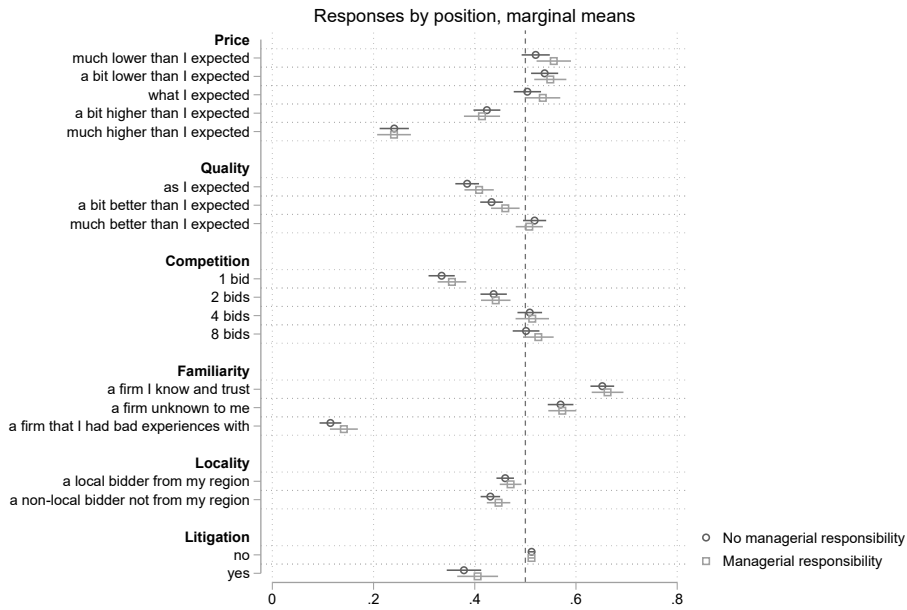
Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome conditional on respondent perception of procurement outcomes' effect on career prospects. The horizontal lines indicate 95% cluster-robust confidence intervals..

Figure A.8: Do tender outcomes matter for career prospects and what role does it play for buyer priorities? Endpoint comparisons of career concerns



Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome conditional on respondent perception of procurement outcomes' effect on career prospects. Only extreme responses ("Absolutely not" and "Yes, absolutely") are included. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure A.9: What role does personnel responsibility play?



Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome conditional on respondent having managerial responsibility. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure A.10: Does office size matter for buyer priorities?



Notes: The figure presents effects of tender outcome attributes on the probability of deciding in favor of a tender outcome conditional on reported levels of office size of the workplace. The horizontal lines indicate 95% cluster-robust confidence intervals. Points without these lines indicate the respective reference categories for the effects of the attributes.

B Appendix: Representativeness of the Finnish Sample

To assess the representativeness of our survey Finnish sample of POs and offices, we draw on detailed administrative data on the universe of Finnish procurement contracts and compare this to key office-level information.³⁰ Our survey execution through a public agency (the FCCA) also allows linking the response margin (i.e., whether or not someone answered the survey) to further administrative data. This provides the basis for a detailed non-response analysis, showing that participation in our survey is not systematically correlated to various office-level characteristics from register data.

We rely on two sources of real-world procurement data for Finland. The publicly administered “Hilma” dataset contains information on the universe of competitively procured public procurement contracts above certain national thresholds. We supplement this administrative data with information from the private electronic procurement platform “Cloudia Oy”, which provides additional details for about 20% of all public procurement contracts.³¹

We begin assessing the representativeness of the Finnish survey by comparing observed office-level characteristics of our sample to information on the whole Finnish public procurement market stemming from the Hilma and Cloudia Oy databases. Table B.1 compares mean values in the survey with those of the real world along key characteristics. The overall picture is one of strong similarity between our survey and the administrative data. In both samples, most contracts (82 and 79%) are allocated according to “open” procedures, and the most frequent type of institution is a municipality (37 and 33%). Just below half of all contracts are below EU thresholds (48 and 45%), the typical average number of bidders is quite close (4.7 to 3.9), and the most frequent awarding mechanism is based on best price/quality considerations (55 and 54%).

While these clear similarities between our survey and further aggregated data sources speak strongly in favor of representativeness, there could still be remaining issues of selective response behavior to our survey invitation. To address these, we use a separate dataset on the response margin (whether or not someone responded to the survey) provided to us by the FCCA, which sent the survey invitations per email. This separate dataset contains an indicator for whether or not a given (anonymized) recipient completed the survey, their municipality, the office name, and their respective email domain. While it is impossible to link this data on the response margin to the actual survey

³⁰We stress again that for Germany no such data is available.

³¹The thresholds for entering the Hilma database are relatively low (e.g., 60,000 euro for goods and most services; 150,000 euro for construction works) and thereby well below the given EU thresholds for public tenders. The Hilma database is publicly available in Finland. The Cloudia Oy data used in the analysis contains detailed information on tender notices for about 20% of all procurement notices in the year 2016, and is provided by the privately administered Cloudia Oy procurement platform provider.

responses, we can link it to the Hilma and Cludia Oy databases via the municipality and office name. This provides a unique opportunity to conduct a detailed non-response analysis with administrative data while guaranteeing full data security of respondents.³²

For our non-response analysis, we regress an indicator for survey completion on the observable office-level administrative information. Table B.2 provides regression results for both OLS and logit specifications. We observe office type and region for the full sample of survey recipients (N=1,353). We are able to identify 61% of these offices in the Hilma database, providing us with the average contract size, most commonly procured item/service category, and typical award procedure. After additionally merging our observations to the Cludia Oy database, we also observe the average number of bidders and the award mechanism used, albeit for a relatively small (13%) subset of the survey recipients.

Overall, we observe very little correlation between individual survey participation and the observable office-level characteristics. The type of office, the contract size, the most commonly procured industry, and the typical award procedure are all uncorrelated with the participation margin—the few significant coefficients are all within the smallest 16% subsample merged to the Cludia Oy dataset. However, a small number of significant regional estimators and a weakly (i.e., at the 0.1 level) significant F-test statistic for joint regional effects warrant some concern. However, we find no clearly discernible pattern in the type of region that might have an effect on the response margin, and many of the regions have a very low number of respondents. All in all, the results of this detailed non-response analysis show that participation in our survey does not seem to be correlated to observable features of a given PO bureau in a meaningful way. Taken together with the previously presented similarities of our survey to aggregate statistics, we conclude that our survey in Finland is broadly representative of Finnish public procurement activity. However, we want to caution that respondents may have selected themselves into participation based on unobservable factors like pro-social motives.

³²The FCCA sent each survey recipient an individualized link to trace whether or not they actually responded. This information on the response margin was, however, captured separately from the actual survey responses, making a linking of the survey responses to participants technically impossible. Moreover, the data on the participation decision provided to the research team for the non-response analysis, while at the individual level, did not contain any names or individual identifiers transferable to other data sources. The only identifiers used to link this data were at the office level.

Table B.1: Comparing Survey Responses to real-world Data

	Survey	Real world
<i>Contract procedure</i>		
Open	81.6	79.3
Restricted	10.4	3.9
Negotiated	4.3	4.7
Sole-source, other or unknown	3.4	12.1
<i>Institution type</i>		
Municipality	37.0	33.0
State Government	12.1	8.9
Consortium of municipalities	18.4	23.4
Other	32.6	34.5
<i>Contract size</i>		
Below EU threshold	47.6	45.25
Above EU threshold	51.2	29.64
Unknown / not posted	1.2	25.11
<i>Awarding mechanism</i>		
Lowest price	42.3	46.5
Best price/quality	55.3	53.5
Other	2.4	
<i>Typical number of bidders</i>		
Mean	4.73	3.95
Median	4	3
SD	2.71	3.72

Notes: This table presents a comparison between survey responses and real world data. Real world information on contract procedure, institution type and contract sizes are obtained from Hilma database, where public buyers are obligated by law to post all procurement notices above certain national thresholds. Buyers are not obligated to post the actual value of the contract size, which explains why it is missing for one fourth of the notices. This also partially explains the large difference in the share of contracts that are above EU threshold. The most recent available year, 2019, is used. Information on awarding mechanism and typical number of bidders are obtained from the Cludia Oy database, which is a more detailed and comprehensive database, containing roughly 20% of procurement notices posted in 2016. When comparing the typical number of bidders, we trim the top 1% of bid amounts to combat the issue of outliers.

Table B.2: Non-response analysis, Finnish sample

	(1) OLS	(2) OLS	(3) OLS	(4) Logit	(5) Logit	(6) Logit
Federal government	0.0217 (0.0476)	-0.0733 (0.0625)	-0.150 (0.151)	0.106 (0.210)	-0.308 (0.267)	-0.842 (0.731)
Government owned company	0.0302 (0.0353)	0.00305 (0.0507)	0.0253 (0.231)	0.146 (0.164)	0.0166 (0.219)	0.0153 (1.289)
Municipality	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Other	0.0278 (0.0458)	0.0196 (0.0695)	-0.0799 (0.250)	0.135 (0.208)	0.0851 (0.290)	-0.683 (1.172)
State government	0.0744 (0.0430)	-0.00404 (0.0657)	-0.650*** (0.193)	0.356* (0.198)	-0.00495 (0.288)	-4.189** (1.850)
Central Finland	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
Central Ostrobothnia	0.0681 (0.117)	-0.0671 (0.185)		0.339 (0.562)	-0.306 (0.838)	
Häme	0.102 (0.0866)	0.0152 (0.137)	-0.0840 (0.277)	0.493 (0.423)	0.0647 (0.575)	-0.0349 (1.321)
Karelia and Kainuu	0.0303 (0.0827)	-0.0419 (0.141)	-0.367 (0.282)	0.158 (0.427)	-0.188 (0.606)	-1.661 (1.382)
Kymenlaakso	0.0571 (0.105)	-0.00765 (0.163)	0.509* (0.300)	0.288 (0.511)	-0.0524 (0.719)	0 (.)
Lapland	-0.0667 (0.0847)	-0.158 (0.147)	0.117 (0.324)	-0.390 (0.509)	-0.771 (0.729)	0 (.)
North Ostrobothnia	-0.0484 (0.0748)	-0.0984 (0.151)	-0.0634 (0.286)	-0.278 (0.426)	-0.446 (0.669)	0.711 (1.965)
North Savo	0.0620 (0.0761)	-0.0426 (0.125)	0.0949 (0.256)	0.312 (0.388)	-0.187 (0.531)	0.804 (1.597)
Ostrobothnia	-0.0435 (0.0892)	-0.178 (0.144)		-0.260 (0.529)	-0.893 (0.733)	
Pirkanmaa	0.120 (0.0750)	0.0169 (0.126)	0.0570 (0.268)	0.573 (0.376)	0.0705 (0.527)	1.131 (1.635)
Satakunta	-0.0372 (0.108)	-0.0264 (0.195)		-0.222 (0.643)	-0.114 (0.849)	
South Ostrobothnia	-0.0170 (0.0831)	-0.171 (0.130)	-0.0319 (0.270)	-0.0868 (0.450)	-0.867 (0.630)	0.880 (2.403)
South Savo	0.0313 (0.0993)	-0.0379 (0.157)	-0.546 (0.455)	0.163 (0.506)	-0.167 (0.672)	-2.365 (1.836)
Southwest Finland & Åland	0.00622 (0.0752)	-0.00656 (0.128)	-0.345 (0.237)	0.0358 (0.400)	-0.0302 (0.540)	-1.388 (1.079)
Unknown	0.223*** (0.0860)	0.0222 (0.147)	-0.789*** (0.252)	0.992** (0.407)	0.0950 (0.615)	-5.203*** (1.724)
Uusimaa	0.115* (0.0653)	0.0313 (0.114)	-0.267 (0.191)	0.546 (0.341)	0.129 (0.476)	-1.050 (0.885)
Avg. contract size		0.00000713 (0.00000441)	0.0000173** (0.00000749)		0.0000302 (0.0000213)	0.000109** (0.0000494)
Most common industry: goods		0 (.)	0 (.)		0 (.)	0 (.)
Most common industry: services		-0.0379 (0.0560)	0.0519 (0.159)		-0.163 (0.239)	0.409 (0.878)
Most common industry: construction		-0.0723 (0.0608)	0.154 (0.209)		-0.317 (0.263)	0.921 (1.109)
Award procedure: open		-0.0995* (0.0558)	0.0567 (0.236)		-0.416* (0.229)	0.339 (1.224)
Award procedure: restricted, negotiated, other		0 (.)	0 (.)		0 (.)	0 (.)
Avg. number of bidders			-0.0570** (0.0274)			-0.350** (0.143)
Share of best price/quality award mechanisms			0.399 (0.319)			1.632 (1.577)
Constant	0.345*** (0.0307)	0.533*** (0.0895)	0.367 (0.264)	-0.653*** (0.138)	0.152 (0.376)	-0.391 (1.324)
Merged to Hilma	No	Yes	Yes	No	Yes	Yes
Merged to Cludia Oy	No	No	Yes	No	No	Yes
Observations	1353	789	174	1353	789	171
R^2	0.027	0.027	0.188			
Pseudo R^2				0.022	0.021	0.152

Notes: This table presents results from OLS and logit regressions on the probability of completing the survey. Dependent variable is a binary variable that gets a value of 1 if a respondent completed the survey and a value of 0 otherwise. Partial, unfinished responses are treated as zeroes. Independent variables are office type, respondent's region, average contract size, most commonly procured industry, award procedure, average number of bidders and award mechanism. The number of observations in column (1) and (4) correspond to the number of individuals who received the survey. Regressions (1) and (4) include the whole survey sample, regressions (2) and (5) include are based on a merge with the Hilma database, and columns (3) and (6) use a sample that is additionally merged to the Cludia database. Some of the regions were combined to further guarantee anonymity. Heteroskedasticity-robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Appendix: Cross-validating Results for Italy

At the outset of this project, we intended to expand the sample of POs to Italy. According to the national regulation (“Codice degli Appalti”), a PO can be embodied by the “Responsabile Unico del Procedimento”, the officer in charge of the tender process, plus other lower-level contracting employees that can provide support. Both professional figures are involved with different responsibilities and frequency in the public contracting process but are subject to the same regulation. Thus, both were eligible POs for this study. We contacted DBInformation S.p.A., a large private company that provides multimedia services to Italian companies to support their development. One of the core business activities provided is consultancy in the realm of public procurement. The dedicated corporate division is Telemat.³³ Since the 1980s, Telemat has been providing its clients daily information on public tenders through the constant monitoring of official sources, support to the understanding of regulations, plus professional training. Thanks to its business activity, which is unique in Italy, Telemat stores the email addresses of a vast number of active Italian POs. Thanks to the collaboration with Telemat, around 59,624 POs were contacted via email (plus two reminders) between September and October 2020 to participate in our survey.

The overall open rate of the invitation email was 15.6%. With the email came a presentation of the survey, and a link was provided to access it. Unfortunately, only about 1% of the email recipients clicked on the survey link. According to the survey structure, those who accessed the page had to agree on proceeding with the survey or not with 90 “clickers” decided to start the survey. Seventy-two of them completed all the survey blocks and can be included for a meaningful comparison of the baseline results with the Finnish and German counterparts. Please see Table C.1 for detailed summary statistics of the Italian sample.

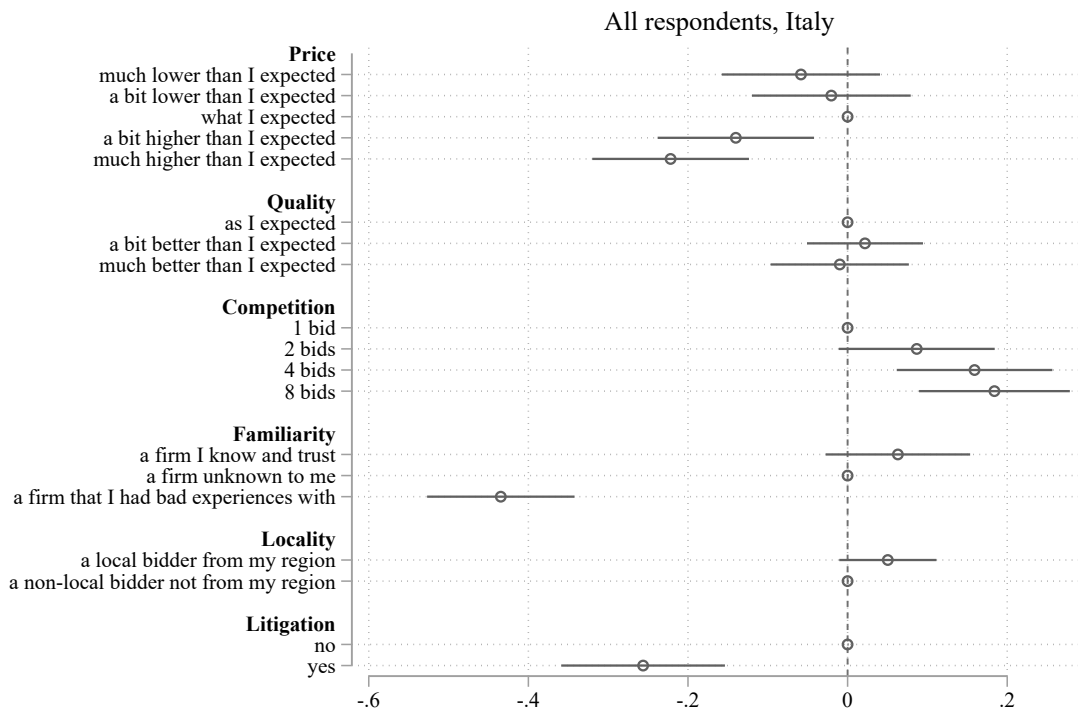
³³<https://www.telemat.it/>.

Table C.1: Respondent characteristics Italy sample

	Italy		
	n	Mean/Prop.	SD
Panel A: Socio-demographic characteristics			
Age group	72		
20-30 years		.06	
31-40 years		.25	
41-50 years		.50	
51-60 years		.19	
above 60 years		.00	
Female		.47	.50
Master's degree		.74	.44
Education background	.56		
Accounting		.04	
Business administration		.05	
Economics		.25	
Engineering		.23	
Finance		.18	
Panel B: Workplace characteristics			
Type of office	72		
Municipality		.51	
State government		.10	
Federal government		.11	
Public company		.08	
Other		.19	
Organization size			
Just me		.21	
2-4		.40	
5-10		.25	
11-19		.06	
20-99		.07	
More than 100		.01	
Type of position	72		
Civil servant		1.00	
Supervisor		.40	.49
<i>Experience in current position (in years)</i>	72	14.50	9.62
Experience in procurement (in years)		16.89	10.01
I am satisfied in competence of my own department	72		
Strongly agree		.13	
Agree		.42	
Disagree		.36	
Strongly disagree		.10	
Type of procedure used	72		
Negotiated		.36	
Open		.26	
Restricted		.10	
Sole-source and other		.28	
Type of award mechanism used	72		
Best price/quality		.50	
Lowest price		.46	
Other		.04	
Share of contracts above EU thresholds		.11	.32
Has secondary objectives		.38	.49
Typical number of bidders		7.43	6.87
Typical number of bidders (censored at 500)		7.43	6.87
Litigated in previous year	72	.38	.76
Most common reason for litigation	17		
A bid protest		.59	
A solicitation challenge		.18	
Other		.24	

Notes: The table presents the socio-demographic and workplace characteristics for survey respondents for the Italian sample. We present a right censored statistic for typical number of bidders to tackle the possible issue of respondent's misunderstanding the question.

Figure C.1: Baseline results, Italy



Notes: Effects of tender outcome attributes on the probability of deciding in favor of a tender outcome. The horizontal lines indicate 95% cluster-robust confidence intervals. Points without these lines indicate the respective reference categories for the effects of the attributes.

D Appendix: Details on Sampling Procedure

Although the samples in both countries cover a heterogeneous set of POs (see Section 2.4 for sample characteristics), the sampling of eligible participants differs. 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), along with a special endorsement letter from the Minister of Local Government. We are confident that such a partnership does not undermine the validity of our conjoint experiment through experimenter bias. Such concerns—possibly affecting revealed preferences—are alleviated by the anonymized nature of the survey, which is made clear to respondents upfront and the fact that the hypothetical scenarios comprised several randomized attributes which made it hard for respondents to act on potential experimenter demand for specific tender attributes. Also, POs were especially incentivized to respond truthfully since we motivated our study (in both countries) for POs to make their voice visible in public discussion. In line with this argument, our conjoint results in Finland, presented in Section 3.2, are similar to those in Germany, where no regulatory partnership was involved whatsoever. 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 Vergabernetzwerk* (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. Although the network includes a wider group of professionals, including private suppliers, government and judicial experts, we only invited those platform members to our survey that distinctly belong to the group of public procurement officers. To filter out any remaining respondents not employed as POs, we ask detailed screening questions about the task structure (see Section 2.2). Overall, we invited 7,247 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, we cannot claim that the German sample is representative of all German public buyers as, to the best of our knowledge, there is no data available on public contracting personnel to benchmark the background answers. The latter, in turn, are valuable for a pioneering overview of the German bureaucratic context. However, in both countries, we receive a diverse set of respondents with respect to demographics, task structures, and workplace characteristics. Please refer to Section 2.4 for a detailed description of our sample. Finally, considering both settings, the attrition rate was only 20%—despite the relative comprehensiveness of the survey—, with most of the drop-outs (15%) quitting the survey during the opening background questions and the remaining 5% during the conjoint experiments.

E Appendix: Detailed Survey Structure

Bloc A: Background characteristics. The bloc is divided into three distinct parts labeled as A.1–A.3, which are described in more detail in what follows.

Bloc A.1 - Socio-demographics. Specifically, *Bloc A.1* asks bureaucrats about their age, gender, education level, and field of higher-education study (if applicable), the government level they are working at, and the number of procurement officers working at their bureau. The corresponding questions are displayed as Q1–Q5 in Appendix F.

Bloc A.2 - Work environment and task structure. *Bloc A.2* elicits specifics about individual tasks and work features of POs (Q6–Q13). In particular, we initially ask which step(s) of the procurement process apply to a respondent’s job (Q6). If participants state that their work is not directly related to public procurement, they are screened out and the survey ends. Further questions survey whether respondents have personnel responsibilities (Q7), what purchase categories they regularly perform (Q8), the stability of their work contract (Q9), and their job experience in public procurement at their current workplace (Q12) as well as in public procurement in general (Q13). Moreover, we ask to what degree they perceive to be able to actively influence purchasing decisions (i.e., their perceived discretion) and whether their workload (prior to the COVID-19 pandemic) made it possible for them to carry out their required work activities (Q10 and Q11).³⁴

Bloc A.3 - Procurement practices. *Bloc A.3* surveys detailed procurement practices of the individual procurement officers and their respective offices. Specifically, we ask what tendering procedure our respondents are typically working on (Q14), which awarding mechanism they regularly use (Q15), whether typical tender values are above or below EU regulatory thresholds (Q16),³⁵ and whether they think their bureau is equipped with the appropriate managerial capability, relevant knowledge, and skills necessary to pursue its goals (Q17).³⁶ Moreover, we ask about the use of secondary objectives, which give room to additional considerations in the evaluation process, such as

³⁴Since we fielded the survey in the second half of 2020, the work environment and tasks of POs were most likely affected by the COVID-19 pandemic. Therefore, we phrased all questions to situations before the onset of the pandemic by wording such as “in the previous year”. However, Q11 on workload might be particularly affected by the the pandemic, and therefore we explicitly asked respondents to refer to times “before the onset of the COVID-19 pandemic”.

³⁵EU law sets out harmonized public procurement rules that govern the way public authorities purchase goods, works, and services. These rules apply to tenders whose monetary value exceeds certain industry-specific thresholds. For tenders of lower value, national rules apply.

³⁶Measuring the perceived quality of the bureau is important since recent work has established a link between bureaucratic competence and procurement outcomes (Decarolis et al., 2020a).

environmental standards (Q18a and b). Q19 then asks whether they perceive secondary objectives as an appropriate tool to improve public procurement outcomes. We also ask how many bids (on average per tender) respondents received in the previous year, how many tenders were challenged by litigations, what caused these litigations most often, and POs' views on the biggest threat to the procurement process (Q20, Q21a, and Q21b and Q22).

Bloc B: Conjoint experiments. Bloc B contains our central choice elicitation procedure, the conjoint experiments. To guarantee the full attention of our respondents, we precede the experiments by a short attention flag asking honest feedback on whether or not our respondents have given the previous questions their full attention (Q23, see Figure F.2 in Appendix F for a screenshot). This self-reported attention flag is based on the recommendation of Meade and Craig (2012) since it is easy to implement and should suffice to screen out at least some careless respondents.³⁷ Only 2% self-report not having paid attention and are thereby disqualified from the rest of the survey. The remainder continues with the conjoint experiments, which are explained in detail in Section 2.3. Furthermore, we screen the answers for outlier response times such as abnormally quick responses, where respondent is unlikely to have given full attention to the experiment, as well as abnormally slow responses, where respondent might have taken a break mid-survey and thus jeopardized the conjoint survey design. The median response time was 213 seconds with 90% of respondents finishing the conjoint experiment within two to ten minutes. We find that our results are robust to excluding answers faster than 2 minutes and slower than 10 minutes. The same holds for the excluding the fast and slowest 1 and 5% of respondents, respectively. Results are available upon request.

Bloc C: Follow-up questions. After presenting the randomized choice scenarios, we ask several follow-up questions aiming to uncover associations between our experimental findings and reported behavior, stated preferences, and perceived career incentives. Specifically, we ask respondents to choose which of the previously experimentally varied attributes in the conjoint (e.g., price, quality etc.) are the *most* and *least* important in terms of three distinct objectives: regarding their perceptions of the general desirability of a tender outcome (Q24, Figure F.4 in Appendix F), regarding the aims pursued in their daily work (Q25, Figure F.5), and regarding their effects on individual career prospects (Q27, Figure F.6) indicating their perceived organization's interests. The last elicitation regarding career prospects is preceded by a short question on whether or not

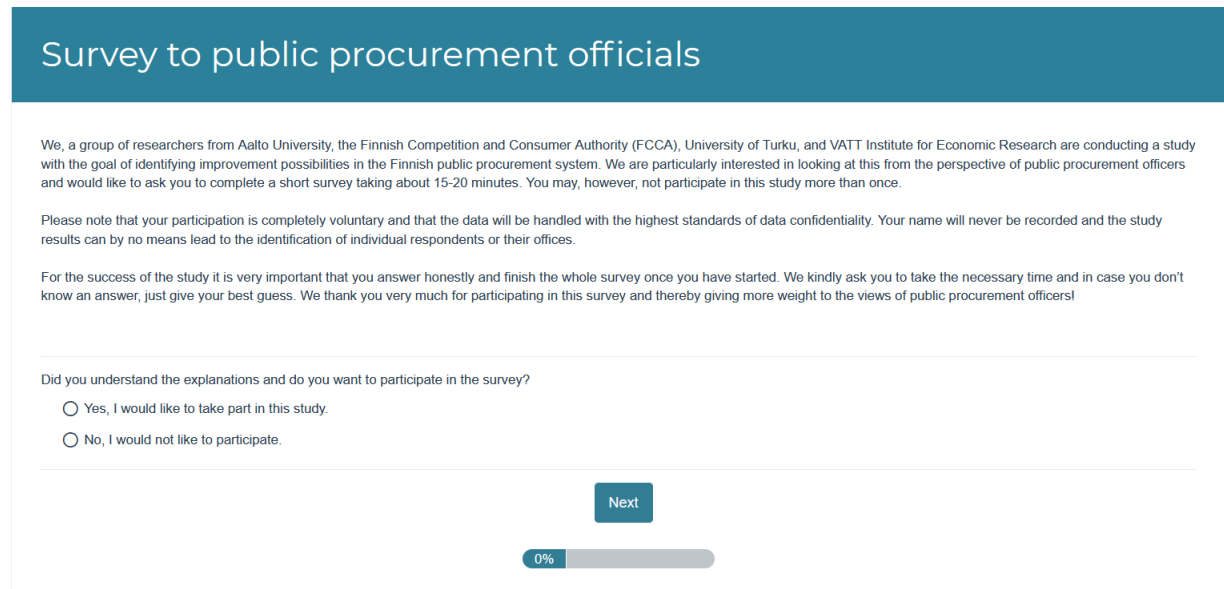
³⁷We abstained from implementing another recommended attention flag (also labelled as "screeners" in Haaland et al. (2021) and Berinsky et al. (2014)) where one would ask respondents to agree with a bogus answer since we were afraid of attrition in response to this question in our special sample of real-life bureaucrats. For the very same reason, we also abstained using multiple attention flag items.

tender outcomes are in general important for career outcomes (Q26), and in case respondents wholly negate, they do not need to respond to the type of objective in Q27. The ordering of these aspects is randomized to guarantee unprimed and unconfounded responses on average. Section 4 analyzes and compares personal and (perceived) office-level preferences regarding key attributes of the public procurement process in more detail.

F Appendix: Detailed Questionnaire

This appendix presents the detailed wording of the questions in the survey module which is translated into English. The actual surveys were fielded in the respective native language (GER: German, FI: Finnish and Swedish, IT: Italian).

Figure F.1: Opener



The screenshot shows the beginning of a survey. At the top, there is a teal header with the title "Survey to public procurement officials". Below this, there are three paragraphs of introductory text. The first paragraph explains that researchers from Aalto University, FCCA, University of Turku, and VATT are conducting a study on the Finnish public procurement system. The second paragraph states that participation is voluntary and data confidentiality is guaranteed. The third paragraph asks for honest answers and thanks the respondent. Below the text, there is a question: "Did you understand the explanations and do you want to participate in the survey?". There are two radio button options: "Yes, I would like to take part in this study." and "No, I would not like to participate.". At the bottom of the form, there is a blue "Next" button and a progress bar showing 0% completion.

Bloc A: Background characteristics.

- **Q1:** How old are you?
under 20 years; 20-30 years; 31-40 years; 41-50 years; 51-60 years; above 60 years; I would prefer not to tell
- **Q2:** What is your gender?
Male; Female; Other
- **Q3a:** Which category best describes your highest level of education? [if the highest level of education is superior to "High School" then move to Q3b]
Primary education or less; High School or Vocational School Diploma; Bachelor or Vocational College; Master Degree; Professional Degree (J.D., M.D., M.B.A., etc); Doctorate; Other. Please specify (optional): [insert text]; I would prefer not to tell
- **Q3b:** Which of the following options best describe the most relevant field of study of your highest degree?
Accounting; Public administration; Business administration; Economics; Finance; Engineering; Law; Other. Please specify (optional): [insert text]
- **Q4:** Which of the following categories best describes the institution you are working for?
Municipality; State government; Federal government; Public company; Other. Please specify (optional): [insert text]

- **Q5:** How many contracting employees does your procurement office currently employ?
Just me; 2-4; 5-10; 11-19; 20-99; More than 100

- **Q6:** Which of these general steps of the procurement process apply to your job? Pick all that are relevant. [if the last option selected then survey ends]
Management and strategic development (supervision of the procurement process and strategic planning); Tender planning (identifying the specific purchase; determining the estimated cost of purchase; budgeting the purchase); Tender set-up (planning the tender timeline; determining tendering procedure and awarding mechanism; preparing the tender documents); Marketing (doing market research and publicizing tender.); Tender management (Bids' evaluation; after-bidding dialog; selecting winner; contract assignment); Contract management (dialogue with contractors; renegotiation of contract terms; finalizing the contract; final assessment/audit); Other work concerning public procurement (administration, research, legal work, education, training, etc.); My work has nothing to do with public procurement

- **Q7:** Do you have managerial responsibility for employees?
Yes; No; I would rather not tell

- **Q8:** Which purchase category best describes the typical procurement process you are involved in?
Construction; Health care services or social services; Health care goods or social goods; Other services: Please specify (optional): [insert text]; Other goods: Please specify (optional): [insert text]; Other services. Please specify (optional): [insert text]

- **Q9:** Which of the following descriptions best fits your current job?
Temporary employment; Unlimited employment; Civil servant; Other. Please specify (optional): [insert text]; I would prefer not to tell

- **Q10:** How much do you think you can influence individual purchasing decisions in your work?
Very much; Somewhat; Rather not; Absolutely not

- **Q11:** If you think of your daily work schedule (before the onset of the COVID-19 pandemic), do you think your workload allowed you to do your job properly and as intended?
Yes, absolutely; Yes, somewhat; No, rather not; No, absolutely not

Please share your view on the following statements.

- **Q12:** I have worked in public-procurement-related tasks for my current employer for
less than or equal to 1 year; more than a year, specifically: [insert number]

- **Q13:** I have worked in public-procurement-related tasks in my career in total for
less than or equal to 1 year; more than a year, specifically: [insert number]

- **Q14:** What kind of tendering procedure is usually employed in the tenders you work on?
Open (tendering is public, no preselection, everyone can submit a bid); Restricted (tendering is public, only those who are preselected can submit a bid); Negotiated (tendering is not public, potential contractors are invited to submit a bid); Sole-source (i.e., direct negotiation with one supplier

only); *Other. Please specify (optional): [insert text]*

- **Q15:** What kind of awarding mechanism is usually employed in the tenders you work on?
Lowest price; Most economically advantageous tender (i.e., best price/quality); Other. Please specify (optional): [insert text]
- **Q16:** Consider the typical procurement process you work on. Is the respective tender value above or below the European regulatory threshold for public procurement purchases?
Above the threshold; Below the threshold; I do not know
- **Q17:** Do you agree with the following statement? My office has the appropriate managerial capability, relevant knowledge, and the skills necessary to accomplish its goals as required.
Strongly disagree; Disagree; Agree; Strongly agree
- **Q18a:** Do your tenders usually involve secondary objectives (for example, minimum requirements for environmental considerations or for social employment)? [if answer is *No*, then skip directly to Q19]
Yes; No
- **Q18b:** Which is the most recurrent type of secondary objectives which you encounter in your work?
Environmental sustainability considerations; SMEs (small and medium-sized enterprise) considerations; Social considerations or social employment targets; Innovation concerns; Other. Please specify (optional): [insert text]
- **Q19b:** Do you think that secondary objectives are an appropriate tool to improve public procurement outcomes?
Yes, very much; Yes, somewhat; No, rather not; No, absolutely not; I do not know
- **Q20:** How many bids (per tender) did you receive on average in all tenders that you were involved in during the previous year?
About bids
- **Q21a:** Of the tenders that I was personally involved in during the course of the last year, how many were challenged by a judicial complaint? [if zero, then skip to Q22]
About [insert text]
- **Q21b:** What was the most recurrent cause of those litigations?
A bid protest (i.e., a challenge to contract award); A solicitation challenge (i.e., a challenge to the tender documents or the decision to exclude firms from bidding); Other. Please specify (optional): [insert text]
- **Q22:** According to your personal experience when working on public procurement purchases, which of the following aspects typically create problems with respect to the procurement process? Pick all that are relevant.
Mistakes from my department in conducting the procurement process; Budgetary constraints for

me and my department from superior offices; The assigned firm did not work properly or created problems; The losing or potential bidder(s) challenged the tender; Rigid regulation; Something else. Please specify (optional): [insert text]; I do not know

Bloc B: Conjoint experiments. The conjoint experiments are comprised of 6 repeated choice scenarios (for an example, see Figure 1 and for a detailed exposition of attribute levels, see Table 1 in the main text). The conjoint is preceded by the attention check (Q23).

Figure F.2: Attention check question (Q23)

Before we proceed with the next questions, we would like to ask you for honest feedback on the answers you have given so far. For our study, it is important that we only include responses from people who have given their full attention to this study.

Should we use your answers for our study or should we not consider your answers further because you have not yet given your full attention to the questions?

- Yes, I have given my full attention to the questions so far and think you should use my answers for your study.
- No, I have not given my full attention to previous questions and I think you should not use my answers for your study.

Next

72%

Note: If second option is chosen, then the survey ends for the respondent.

Figure F.3: Short explanation of choice scenarios

Next, we present to you 6 pairs of hypothetical tender outcome scenarios. In these scenarios, think of the type of purchases that you are usually involved with in your work as a public procurement employee. Every time, for the same purchase, two award outcome scenarios are presented. Consider that each new comparison pair deals with a new purchase. Please look at each pair of tender outcomes carefully. Then make a decision according to the option you would personally prefer, i.e., what scenario you prefer as a procurement employee.

The exact situation is as follows.

A procurement contract resulting from a tender has just been awarded. This means that all bids are submitted and you or your office have already chosen the winner. Losing bidders have had time to challenge the contract award decision if they wanted to. However, the object of the contract is not yet delivered. This means that information on the price and quality in these scenarios is only based on the bids.

Next

75%

End of Bloc B: Thank you very much for your answers so far! There are only a few questions remaining.

Bloc C: Follow-up questions. These are presented after completing the conjoint experiments.

Figure F.4: Role of experimental attributes for desirable tender outcomes (Q24)

23. Please evaluate which of the following factors is important for a **desirable tender outcome**.

Note that you can only pick one answer category to be either the most and the least important factor for determining a desirable tender outcome, respectively. This means that for each row (most important/least important), you can only select one answer.

	Ensure a high quality purchase	Support local firms	Ensure a low price	Avoid legal complaints	Receive a high number of bids	Select a winner that one can trust
Most important	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Least important	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Next](#)

84%

Figure F.5: Role of experimental attributes for stated behavior related to tender outcomes (Q25)

24. Which of the following factors best describes the goals you do or do not pursue in your **daily work**?

Note that you can only pick one answer category for the goals that you actively do or do not pursue in your own work, respectively. This means that for each row (I actively try to.../I do not actively try to...), you can only select one answer.

	Ensure a high quality purchase	Support local firms	Ensure a low price	Avoid legal complaints	Receive a high number of bids	Select a winner that one can trust
I actively try to...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I do not actively try to...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Next](#)

88%

- **Q26:** Do you consider tender outcomes important for your career prospects? [if the last two answers are chosen, the survey ends for the respondent]
Yes, absolutely; Yes, somewhat; Rather not; No, absolutely not

Figure F.6: Role of experimental attributes for career prospects (Q27)

26. Please evaluate which of the following factors is important for your **career prospects**.

Pick only one answer category to be either the most and the least important factor for your career prospects. Note that you can only pick one answer category to be either the most or the least important factor for your career prospects, respectively. This means that for each row (Most important/Least important) you can only select one answer.

	Ensure a high quality purchase	Support local firms	Ensure a low price	Avoid legal complaints	Receive a high number of bids	Select a winner that one can trust
Most important	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Least important	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Submit](#)

97%

Thank you very much for answering the survey!

G Appendix: Experiment Timing and Procurement Flowchart

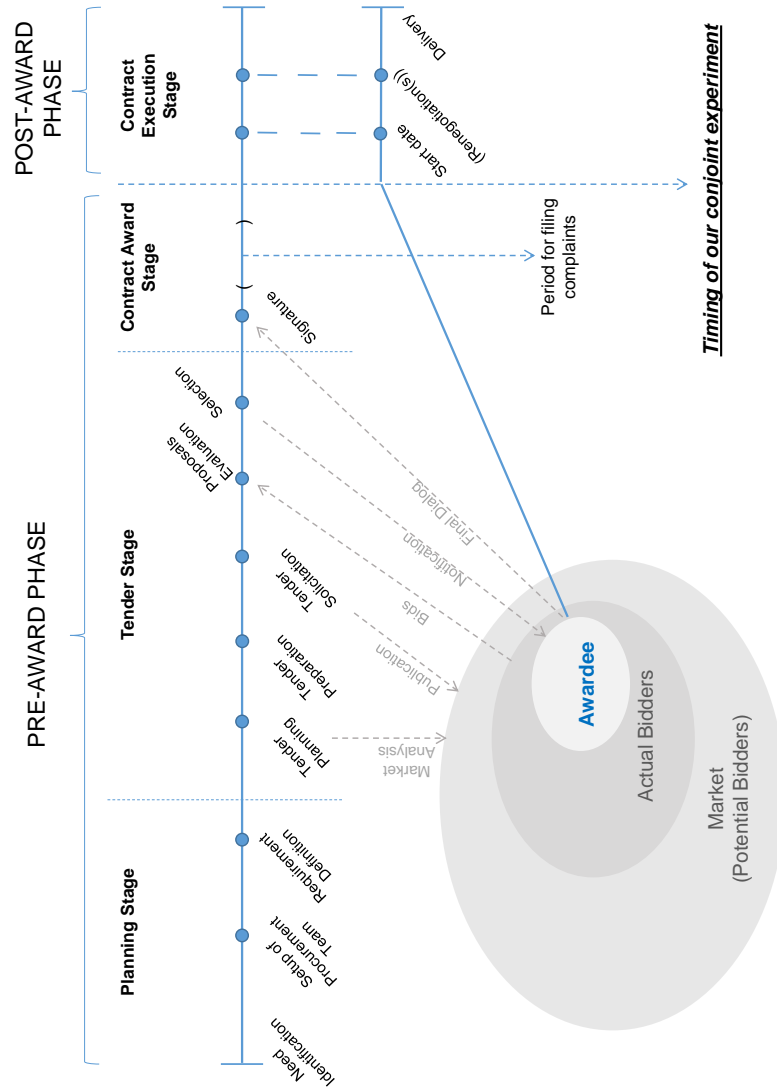


Figure G.1: Our Experiment Timing Along the Public Procurement Process

H Appendix: Robustness Checks

No carryover effects. An important identification assumption is the stability or the avoidance of carryover effects across different hypothetical choices in our experimental design. Non-stable effects across choice tasks would suggest that respondents value a particular degree of, say, competition in tender outcome scenarios in a certain task more or less depending on what type of profiles they have seen before or after that particular profile. While such effects are unlikely given the randomization of profiles across experimental rounds, we also test for these types of spillovers in Figure H.1 by executing distinct regressions of respondent choices per individual choice task among tender outcomes. Our results are very similar across all rounds of choice tasks suggesting the absence of carryover effects across rounds.

No attribute and profile order effects. Another concern for identification could be that respondents are affected by the order in which attributes are presented to them. This could be due to a priming (or primacy) effect in which a larger share of the attention goes to attributes placed near the top of the choice task. Our survey design, in which we randomize the attribute order (but keep it constant for a given respondent), attenuates this concern considerably. Nevertheless, we test for such attribute order effects with a set of regressions for each attribute (e.g., price), in which we regress the binary choice on a set of dummies for the attribute realizations, a dummy for the position of the attribute, and the interactions of these variables. Figure H.2 shows that the order in which the attributes are presented does not seem to have an effect on choices. A related concern could be that respondents favor different profiles due to the order in which they are presented within a particular task, that is, whether a given profile is shown as card A or B. We test this by splitting our sample depending on whether a given profile was presented on the left or right side of the choice task. Figure H.3 finds no discernible differences in individual estimates between these two samples.

Specification checks. We also show the robustness of our results with a number of sensitivity checks varying our econometric specification. Table H.2 includes various combinations of specifications with controls at the respondent level (socio-demographics and/or workplace characteristics), different sets of fixed effects (no fixed effects, set and/or card-level fixed effects), as well as standard errors without clustering. Moreover, our results are robust to using a logit model instead of a linear probability model (see Figure H.4).

Plausibility checks. We also perform several plausibility checks of our experimental results by drawing on additional survey information. As a follow-up to the conjoint experiments, we directly ask respondents to identify the most and the least advantageous

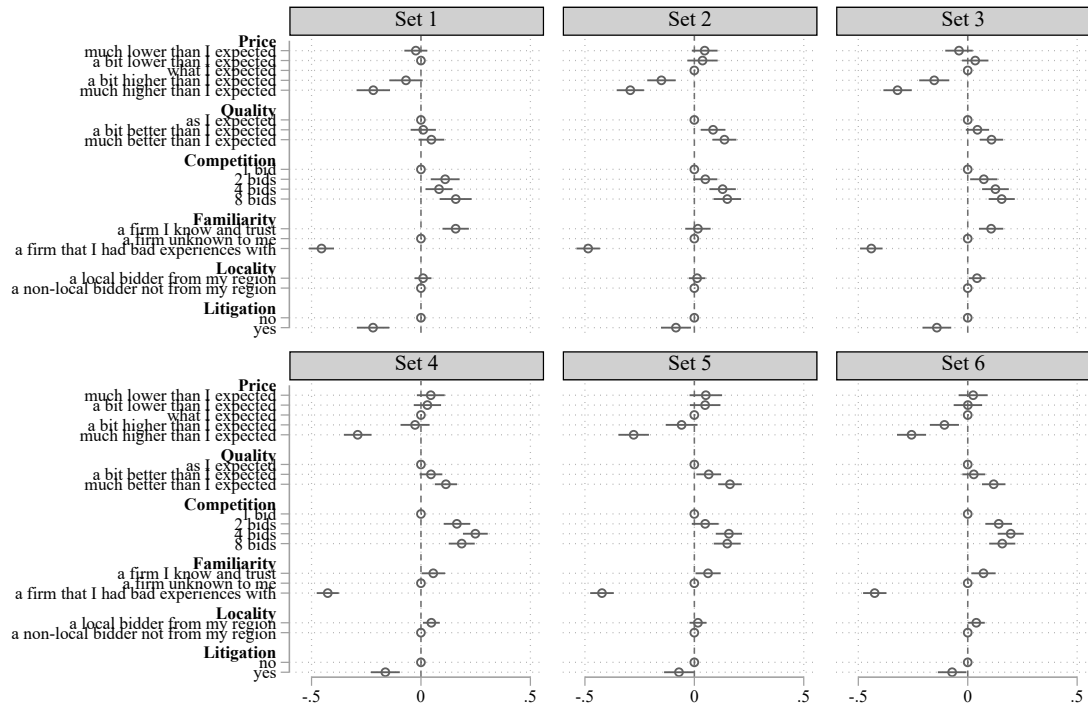
factor for a desirable tender outcome. We then split the sample into subgroups based on which particular factor was deemed to be most (Figure H.5) and least (Figure H.6) desirable and compare the experimental results.³⁸ It appears that stated considerations tend to coincide with the experimentally elicited priorities. For example, respondents that report a high number of bids as the most important factor towards a desirable tender outcome exhibit marginal means significantly smaller (in case of the “1 bid” realization) and larger (in case of the “4 bid” or “8 bid” realizations) than respondents that report other issues as more desirable. We find similarly plausible subgroup effects when differentiating respondents according to their predominantly used awarding mechanism.³⁹ Figure H.7 shows that when bureaucrats typically select bids using a “lowest price” allocation rule, they are particularly worried about paying prices that are too high. However, the differences for both of these subgroup divisions (stated preferences and type of awarding procedure) are small in economic terms and frequently statistically insignificant. This suggests that while individual tastes and commonly used procurement practices have an effect in their expected directions, our choice experiments identify a more general pattern of bureaucratic preferences, reconfirming the “similarity of findings” discussed in Section 4.

Multiple-hypothesis testing. We check whether our standard errors are robust to multiple-hypothesis testing when accounting for the multitude of treatments in our survey design. We follow the procedure proposed by Westfall and Young (1993). We treat each attribute realization (e.g., “much lower than I expected” and “a bit lower than I expected” for price) as separate treatments, resulting in one outcome and 13 potential treatments after accounting for baselines. The Westfall-Young procedure confirms that all statistically significant treatments retain their significance at conventional levels even when adjusting for multiple comparisons (please see Table H.1).

³⁸When analyzing subgroups, we estimate marginal means instead of conditional AMCEs to account for potential differences in the omitted categories between subgroups (Leeper et al., 2020). We discuss this approach further in Section 4. The subgroups are based on Q24 in the questionnaire, and for a descriptive overview of which factors respondents considered desirable, please refer to Figure A.1 in Appendix A.

³⁹This measure for the predominantly used awarding mechanism rests on asking POs about their “typical” awarding procedure, and is effectively a subjectively weighted average between the two most commonly used procedures. About 40% of respondents typically use “lowest price”, while the remainder mostly uses scoring rules base on the best price-quality ratio (details in Table A.2).

Figure H.1: Testing for carryover effects across profiles



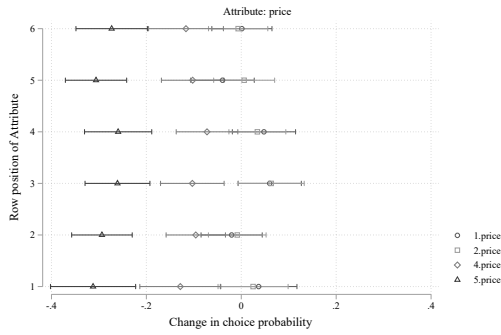
Notes: This figure presents the results from running the main regression separately for each choice set offered to the respondents. Set number refers to the ordinal number of choice set offered. The horizontal lines indicate 95% cluster-robust confidence intervals. Points without these lines indicate the respective reference categories for the effects of the attributes. Please note that there was a bug in the Alchemer survey software that caused price attribute "what I expected" to not be offered in set 1, hence a different baseline.

Table H.1: Westfall-Young corrected p-values

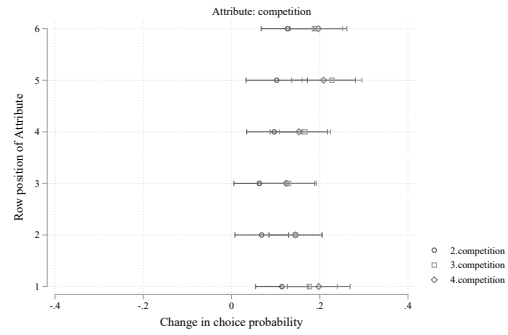
	Full sample
much lower than I expected	0.0160 (0.2546) [0.2551]
a bit lower than I expected	0.0235 (0.0806) [0.1411]
what I expected	0 (.) [.]
a bit higher than I expected	-0.0982*** (0.0000) [0.0000]
much higher than I expected	-0.280*** (0.0000) [0.0000]
as I expected	0 (.) [.]
a bit better than I expected	0.0529*** (0.0000) [0.0001]
much better than I expected	0.121*** (0.0000) [0.0000]
1 bid	0 (.) [.]
2 bids	0.0956*** (0.0000) [0.0000]
4 bids	0.166*** (0.0000) [0.0000]
8 bids	0.168*** (0.0000) [0.0000]
a firm I know and trust	0.0841*** (0.0000) [0.0000]
a firm unknown to me	0 (.) [.]
a firm that I had bad experiences with	-0.442*** (0.0000) [0.0000]
a local bidder from my region	0.0288*** (0.0003) [0.0010]
a non-local bidder not from my region	0 (.) [.]
no	0 (.) [.]
yes	-0.124*** (0.0000) [0.0000]
Constant	0.520*** (0.0000)
Observations	11196

Notes: Dependent variable is a binary variable with a value of 1 if profile was chosen and a value of 0 otherwise. Effects are estimated using OLS and standard errors are clustered at the respondent level. P-values are in parenthesis and Westfall-Young corrected p-values are in brackets. Stars refer to * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$

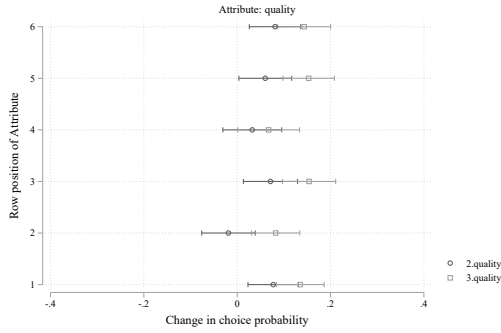
Figure H.2: Testing for attribute order effects



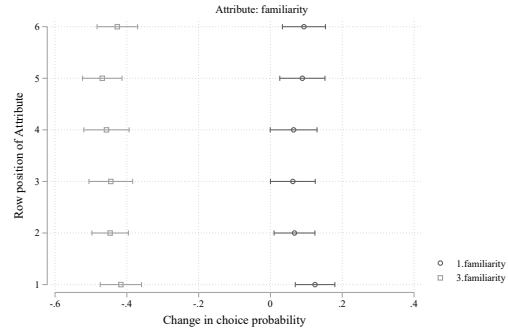
(a) Price



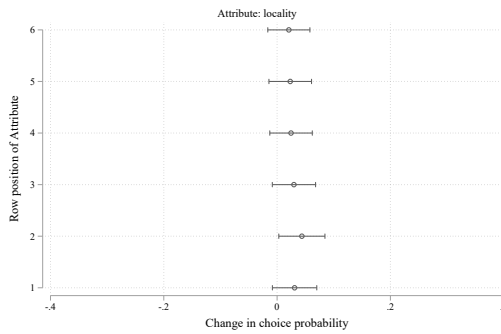
(b) Competition



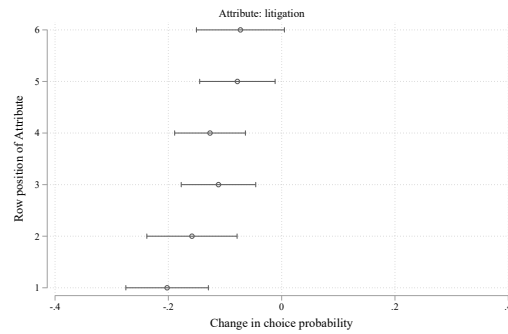
(c) Quality



(d) Familiarity



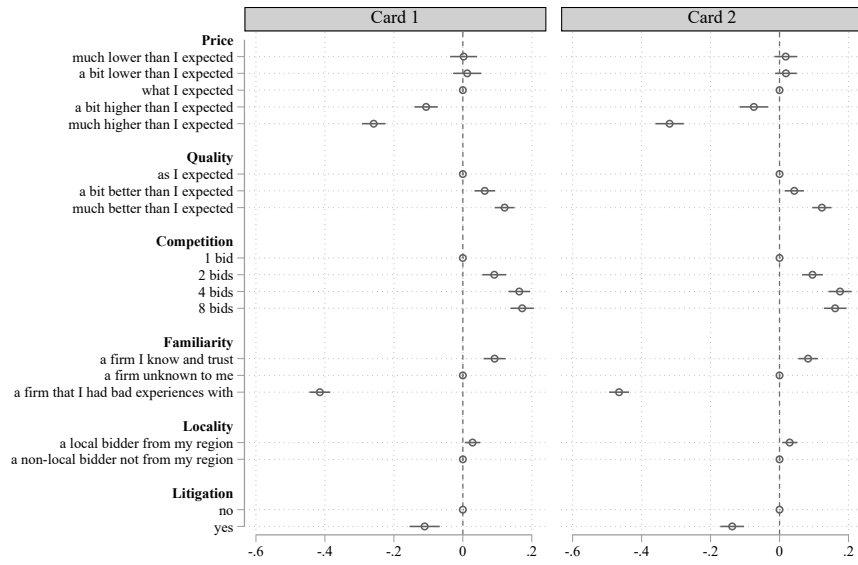
(e) Regionalism



(f) Litigation risk

Notes: Figures present regression results for attribute order tests. The estimates are for the interaction coefficient of attribute and a dummy depicting its row position in the survey when regressed on the probability of choosing an outcome. The attribute row positions were randomized between respondents, but constant for each respondent. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure H.3: Testing for profile order effects



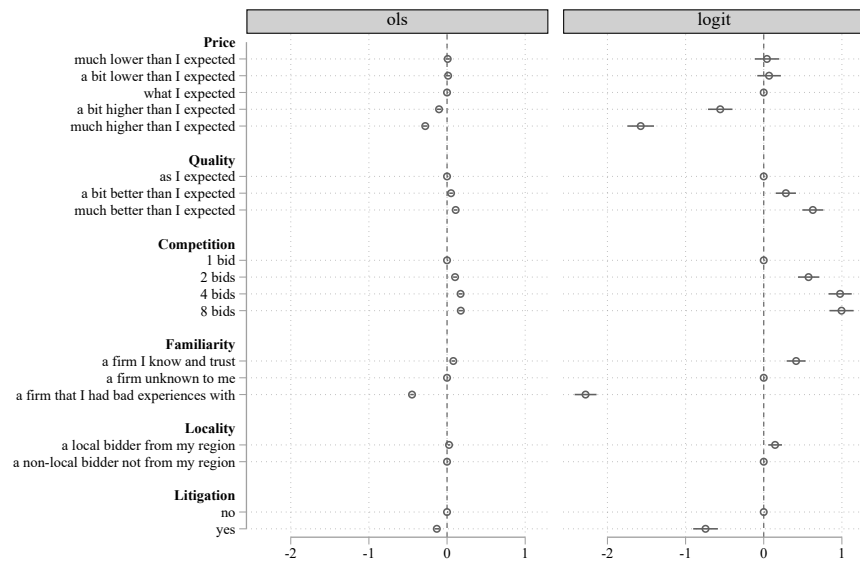
Notes: The figure presents main regression results for each card (i.e., option) separately. The reason is to test whether the attributes would have different results when presented in the left as opposed to the right card. The horizontal lines indicate 95% cluster-robust confidence intervals.

Table H.2: Robustness of estimated coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
much lower than I expected	0.0160 (0.0141)	0.0153 (0.0142)	0.0136 (0.0142)	0.0127 (0.0144)	0.0127 (0.0130)	0.0165 (0.0142)	0.0122 (0.0175)	0.00994 (0.0130)
a bit lower than I expected	0.0235* (0.0134)	0.0196 (0.0137)	0.0212 (0.0136)	0.0172 (0.0139)	0.0172 (0.0129)	0.0247* (0.0136)	0.0184* (0.00858)	0.0123 (0.0132)
what I expected	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
a bit higher than I expected	-0.0982*** (0.0145)	-0.0964*** (0.0147)	-0.0994*** (0.0147)	-0.0977*** (0.0149)	-0.0977*** (0.0133)	-0.0995*** (0.0146)	-0.0994*** (0.0219)	-0.0947*** (0.0131)
much higher than I expected	-0.280*** (0.0147)	-0.279*** (0.0150)	-0.282*** (0.0148)	-0.281*** (0.0152)	-0.281*** (0.0131)	-0.283*** (0.0149)	-0.283*** (0.0155)	-0.278*** (0.0132)
as I expected	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
a bit better than I expected	0.0529*** (0.0118)	0.0549*** (0.0119)	0.0516*** (0.0119)	0.0536*** (0.0121)	0.0536*** (0.0103)	0.0554*** (0.0120)	0.0528*** (0.00733)	0.0546*** (0.0101)
much better than I expected	0.121*** (0.0117)	0.122*** (0.0117)	0.119*** (0.0118)	0.120*** (0.0119)	0.120*** (0.0101)	0.124*** (0.0119)	0.120*** (0.0110)	0.121*** (0.0101)
1 bid	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
2 bids	0.0956*** (0.0127)	0.101*** (0.0129)	0.0974*** (0.0128)	0.103*** (0.0131)	0.103*** (0.0118)	0.0966*** (0.0128)	0.102*** (0.0199)	0.102*** (0.0116)
4 bids	0.166*** (0.0130)	0.173*** (0.0132)	0.168*** (0.0130)	0.175*** (0.0132)	0.175*** (0.0116)	0.168*** (0.0131)	0.178*** (0.0183)	0.178*** (0.0117)
8 bids	0.168*** (0.0137)	0.175*** (0.0139)	0.167*** (0.0138)	0.174*** (0.0140)	0.174*** (0.0116)	0.170*** (0.0138)	0.177*** (0.0126)	0.176*** (0.0116)
a firm I know and trust	0.0841*** (0.0124)	0.0806*** (0.0127)	0.0835*** (0.0125)	0.0800*** (0.0128)	0.0800*** (0.0106)	0.0849*** (0.0126)	0.0797** (0.0206)	0.0816*** (0.0101)
a firm unknown to me	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
a firm that I had bad experiences with	-0.442*** (0.0120)	-0.449*** (0.0120)	-0.442*** (0.0122)	-0.450*** (0.0121)	-0.450*** (0.0100)	-0.450*** (0.0123)	-0.450*** (0.00785)	-0.448*** (0.0101)
a non-local bidder not from my region	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
a local bidder from my region	0.0288*** (0.00795)	0.0242*** (0.00807)	0.0284*** (0.00807)	0.0236*** (0.00819)	0.0236*** (0.00819)	0.0336*** (0.00867)	0.0235** (0.00724)	0.0232*** (0.00818)
no litigation	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)
litigation	-0.124*** (0.0145)	-0.122*** (0.0148)	-0.125*** (0.0147)	-0.123*** (0.0150)	-0.123*** (0.0141)	-0.128*** (0.0149)	-0.123*** (0.0215)	-0.126*** (0.0137)
constant	0.520*** (0.0162)	0.518*** (0.0183)	0.519*** (0.0178)	0.521*** (0.0197)	0.521*** (0.0325)	0.518*** (0.0165)	0.521*** (0.0238)	0.519*** (0.0319)
Observations	11196	10740	10392	10392	10392	11196	10392	10392
R ²	0.293	0.298	0.293	0.298	0.298	0.298	0.299	0.295
Controls	None	Socio-demographic	Workplace	All	All	None	All	All
Fixed Effects	No	No	No	No	No	Respondent	Set	Card
Standard errors	Clustered	Clustered	Clustered	Clustered	Robust	Robust	Robust	Robust

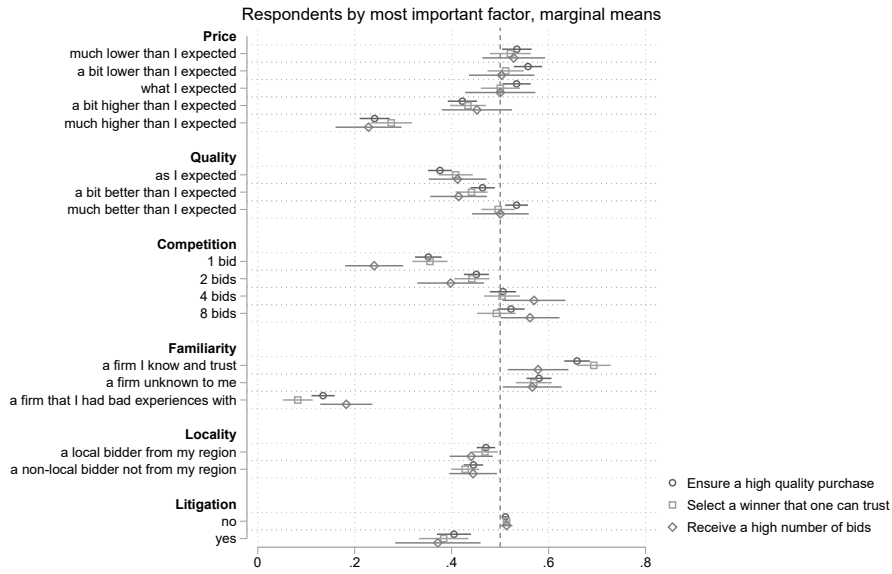
Notes: Dependent variable in each regression is a binary variable with a value of 1 if profile was chosen and a value of 0 otherwise. Standard errors are in parenthesis. Effects are estimated using OLS. Clustered standard errors are clustered at the respondent level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure H.4: Alternative specifications



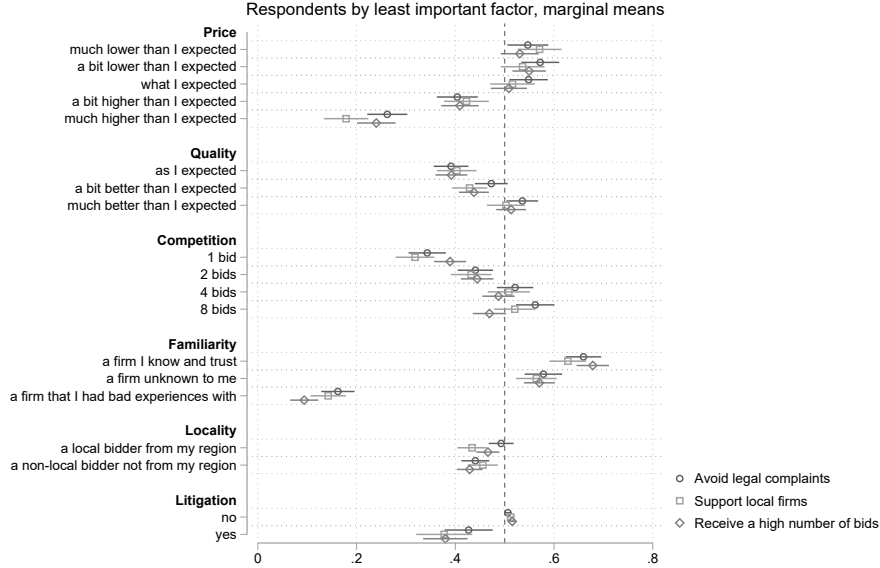
Notes: The figure presents effects of tender outcome attributes on the probability of deciding in favor of a tender outcome using both OLS and logit specification. The horizontal lines indicate 95% cluster-robust confidence intervals. Points without these lines indicate the respective reference categories for the effects of the attributes.

Figure H.5: Role of most important factor for public buyer priorities



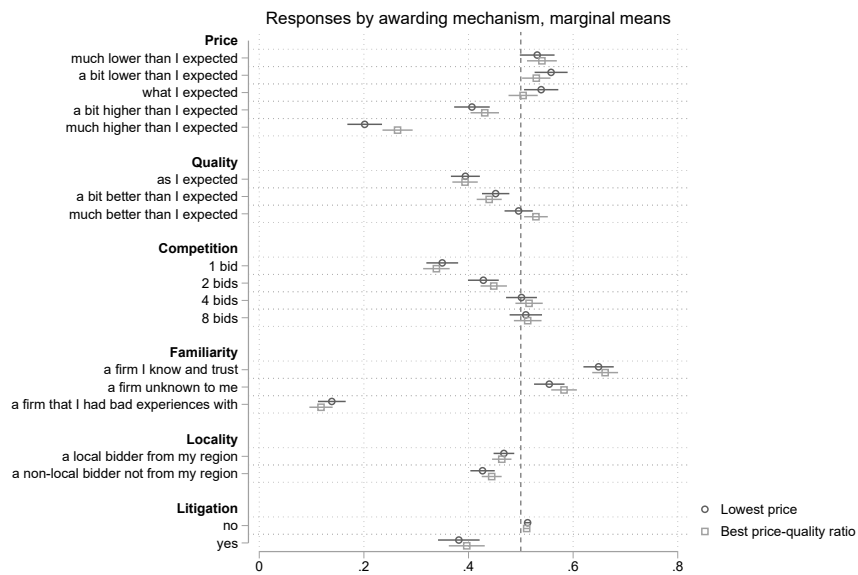
Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome, conditional on which factor a respondent considers to be most important for a desirable tender outcome. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure H.6: Role of least important factor for public buyer priorities



Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome, conditional on which factor a respondent considers to be least important for a desirable tender outcome. The horizontal lines indicate 95% cluster-robust confidence intervals.

Figure H.7: Role of award mechanisms for public buyer priorities



Notes: The figure presents marginal means of tender outcome attributes on the probability of deciding in favor of a tender outcome conditional on what award mechanics respondent typically uses. The horizontal lines indicate 95% cluster-robust confidence intervals.

I Appendix: Testing for Effect Heterogeneities

In our analysis, we assess the heterogeneity of our effects along a range of observable characteristics with respect to socio-demographics, task structures, and office-level features. Our approach consists in splitting the sample into two or more subgroups depending on the realization of the potential heterogeneity under study. We test the following sub-groups:

- *Socio-demographic characteristics*—Age groups (i.e., age 21–40 vs. > 41), gender (i.e., female yes vs. no), education level (i.e., bachelor vs. master), education type (i.e., business vs. engineering vs. other) as well as work experience in procurement (i.e., 10 or less years vs. above 10 years)
- *Task structure and features*—Typical type of purchases (i.e., construction vs. goods vs. services), prior litigation experienced (i.e., yes vs. no), typical awarding mechanism used (i.e., lowest price vs. best price-quality ratio), typical use of secondary objectives (i.e., yes vs. no), type of secondary objectives (i.e., environmental vs. SME support), typical size of a procurement contract (i.e., below vs. above EU threshold), typical number of bidders per contract in last year (i.e., less than 5 vs. 5 or more bidders) as well as perceived workload (i.e., too much vs. sufficient workload), type of job/phase of procurement process (i.e., tender set-up, planning and management vs. marketing vs. strategic development vs. contract management), perceived factors aspects which create problems in tendering process (i.e., mistakes by my department vs. budgetary constraints vs. litigation vs. assigned firm created problems vs. rigid regulation), perceived discretion: own influence on purchasing decisions (i.e., has influence vs. does not have influence) as well as typical tender procedures (i.e., open vs. negotiated vs. restricted vs. other)
- *Office-level characteristics*—Office competence (i.e., satisfactory vs. non-satisfactory) and organization size (i.e., less than 5 vs. 5 or more employees in procurement as well as fewer than 5 vs 11 or more contracting employees), government tier of office (i.e., federal vs. municipal vs. state-level vs. public company)

In our analysis, we estimated all specifications for all of these different subgroups and found no systematic differences, with the differences being insignificant in almost all instances. The detailed results are available from the authors upon request. We conclude from this heterogeneity analysis that the effects are extremely stable across observables.