This is a report published by *Umeå University*

TENDERING DESIGN WHEN PRICE AND QUALITY IS UNCERTAIN

THEORY AND EVIDENCE FROM PUBLIC PROCUREMENT[*]

Mats A. Bergman, Södertörn University[♣]
Sofia Lundberg, Umeå University[♦]

September 17, 2014

Abstract

Departing from a simple normative theory for the choice between lowest price, highest quality (beauty contest) and more complex scoring rules, we empirically investigate the behavior of local and central authorities. We survey a gross sample of 40 contracting entities about perceived key characteristics of products bought in 651 public procurements and collect data on supplier selection methods for these procurements. We compare actual scoring rules with theoretical norms and analyze what product characteristics make deviation from the norm more or less likely. In addition, a control group of 275 authorities was surveyed about similar but hypothetical procurements. We find that more complex scoring rules are used more often when the authority is uncertain about costs and about delivered quality, in accordance with our hypotheses. However, authority effects are also found to directly and indirectly influence the choice of supplier-selection method, suggesting that tendering design is partially driven by local habits or institutional inertia.

JEL code: H57, D44, P16.

Key-words: Auctions, Contracting, Habit behavior, Moral hazard, Scoring rules, Supplier selection

1. INTRODUCTION

According to the European public procurement directives[1], contracting authorities can award contracts either on the basis of lowest price or to the economically most advantageous tender (EMAT). A naïve interpretation would be that when quality is of little importance the authority should chose lowest price while an authority that cares a lot for quality should rely on EMAT – and should give a large weight to quality and a small weight to price (cost). Since quality can be safeguarded by minimum quality standards, however, the principles that should guide the choice between lowest price and EMAT – and that should be decisive for the weight given to quality – are more complex[2] We seek to contrast these normatively deduced principles against the practices actually used by contracting authorities, while also testing the alternative hypothesis that procurement design is mainly the outcome of ingrained organizational traditions.

Building on previous research based on part of the data used in the current paper (Bergman and Lundberg, 2012) we focus on two fundamental problems that a contracting authority has to face. First, when quality is non-verifiable there is often an incentive for the supplier to shirk on quality after the contract has been secured – i.e., the procurer has to deal with moral hazard. We will refer to this as quality uncertainty. Second, the procurer may face substantial uncertainty concerning the cost of meeting alternative quality levels. If quality standards are set too high, the price may turn out to be higher than expected and quality may on the margin cost more than it is worth.

We argue that EMAT should be used when the degree of cost uncertainty is high or when the delivered quality is uncertain and that lowest price should be used when the degree of cost uncertainty is low or when it is critical to reach a threshold level of (measurable) quality. Our empirical findings, based on a sample of 651 actual procurements as well as on two surveys directed to procuring authorities, largely conform to these predictions, but we also find indications that local traditions are important determinants of tendering design. In line with the neo-institutional tradition or the "organizational-process model" (Alison, 1971), as well as transaction-cost theory, a given procuring authority has a tendency to make similar choices when facing different circumstances, while making different choices than other authorities that face similar circumstances.

---

[1] Directive 2004/17/EC and Directive 2004/18/EC.
[2] See Bergman and Lundberg (2013) for references and a discussion.

Securing quality in procurement is complex, especially for non-verifiable quality characteristics. There is a trade-off between detailed rules and discretion; i.e., between rules that result in a high degree of predictability and that curb corruption – and rules that allow contracting authorities to base the allocation of contracts on reputation or to effectively blacklist providers known to perform poorly in non-verifiable quality dimensions (Tadelis, 2012). Due to stricter rules in public than in private procurement, the risk of quality degradation is higher in the former but, presumably, so is the risk of corruption and this is what justifies the emphasis on predictability and transparency. Also, the risk of degradation of non-contractible quality is perceived to be higher with outsourcing than with in-house production, since private providers have stronger cost-savings incentives (Hart, Schleifer, and Vishny, 1997).

Different jurisdictions handle this trade-off differently. In particular, recent US and EU reforms have adjusted policies in the opposite directions. The US has sought to strengthen executive leadership, shifting focus from detailed rules and regulations to a more flexible, dynamic and entrepreneurial approach in its government management – i.e., officials have been allowed more discretion.[3] See, e.g., Kelman, 2002; Potoski, 2008; Schooner, 2011. In contrast, the EU procurement directives have become stricter in stipulating that the call for tender specify how bids will be evaluated in terms of a scoring rule, giving less room for discretion while seeking to reduce the risk for undue discrimination. See Arrowsmith (2006) and Sánches Graells (2013).[4]

Although the procurement auction and the principles for supplier selection are regulated by EU directives and national laws within Europe, they give the contracting authority substantial freedom in designing the bid evaluation process and in choosing what qualification and award criteria to use. The tension between different objectives, diverging policies across the Atlantic and the more general question whether the choices made by authorities reflect rational decision making or institutional inertia together motivate the current study. We seek an understanding of the contracting authorities' perception of the relative importance of price and quality and of the degree of uncertainty, of how this relates to revealed behavior in the choice of supplier selection methods and scoring rules and of the importance of the organization's habits in shaping its choices in specific instances of design choice.

---

[3] This was the result of the so-called Winter Commission in 1993 (Kelman, 2002). See e.g. Coggburn (2003), Kelman (2007), Thompson (2008) and Orszag (2009) for discussions of the US reform.
[4] An evolution in this direction has been endorsed by the academic community, see e.g., Chen (2008), Mateus et al.(2010), Telgen and Schotanus, (2010).

We have detailed information on 651 public procurements of a wide range of products. Our data includes type of product, type of contracting authority, supplier selection methodology, relative weights for price and quality (if applicable), duration of contract and other contract characteristics. The sample is representative for virtually all purchases made by local, regional and central government authorities in Sweden during 2007 and 2008. A survey was sent to the responsible authorities, asking for their opinion on the severity of the moral hazard problem and the cost uncertainty problem, as well as on the relative importance of meeting budget and quality targets.

To validate the responses and to explore alternative explanations for observed behaviour, we surveyed other local authorities (municipalities) about the same products. About 100 municipalities in this group responded, giving us more than 900 useable answers on the 216 products procured by municipalities among the 651 in our total net sample.

Overall, lowest price is used in close to 40 percent of the procurements, bid prices are transformed to scores in more than half of all procurements, while values are assigned to quality characteristics in about one tenth of the procurements. We use three different empirical techniques to investigate the relation between product and procurer characteristics – and the supplier selection method. First, simple empirical discrete-choice frameworks (e.g., logistic regressions). Second, seemingly unrelated regressions (SUR) and, finally, partial R^2 measures.

We find that more complex scoring rules are used more often – and that price is given a lower weight – when the authority is uncertain about price and about delivered quality, in accordance with our hypotheses. Furthermore, we find that extension periods are positively associated with quality uncertainty, that EMAT is used more often for high-value (above-threshold) procurements and that there seems to be important product-type effects. Other findings, that are more indicative of local habits playing a role, include strong authority effects and the tendencies that central-government authorities and large authorities favor lowest-price procurements.

The outline of the paper is as follows. A theoretical background on procurement and the quality – price trade-off is provided in Section 2, where we also formulate the hypotheses we will test. The survey and the data are presented in Section 3. The empirical approach is explained in Section 4, results are reported in Section 5 and Section 6 concludes.

2. THEORY AND HYPOTHESES

Auction theorists and empirical researchers have analyzed various design aspects relevant for selling auctions as well as for buying auctions, including public procurement. Prominent examples include the choice between first and second-price auctions or other auction formats, such as Dutch (or reversed) auctions; whether each bidder should pay (or be paid) its own price or whether all should pay (be paid) the market-clearing price; bidding strategies in auctions with multiple object etc. (Klemperer, 2004; Milgrom, 2004; Krishna, 2002.) Typically, the quality of the object is taken as given. While this is a natural assumption in selling auctions, in buying auctions it is often not.

In public procurements, two highly significant concerns are, first, that the contract emanating from the tender process strikes the right balance between cost and different quality attributes and, second, that the contracted quality is actually delivered. The first is a problem of optimization; the second a problem of managing asymmetric information. Both problems can at least to some extent be managed through the choice of bid-evaluation mechanism – lowest price or a scoring rule that is responsive to both quality and price, also known as EMAT, the Economically Most Advantageous Tender.[5] While bid evaluation and the design of the bid-evaluation mechanism are important considerations in procurement practice, they have attracted relatively little attention from the academic community.[6]

Departing from a rational-actor perspective, it can be presumed that procuring authorities will make design choices in accordance with normative theory. For example – and as will be discussed below – scoring rules that put a larger weight on quality can be expected in procurements where it is difficult to optimize quality ex ante and where ex-post delivery of quality may be in doubt because of asymmetric information.

An alternative perspective, however, is that the contracting authorities' choice of mechanism is to a large extent driven by habits. According to what is sometimes called the Organizational Process Modell (Allison, 1971; Allison and Zelikow, 1999), organizational behaviour is persistent, sometimes to the point where ingrained behaviour defies the organization's own

---

[5] EMAT has been the most common principle in most EU member countries (Verdeaux, 2003; Strand et al, 2011). The principle has been perceived as allowing wide discretion to the contracting authorities to select the winner, but rules are stricter today. Results based on Swedish data from the period with more lax regulation suggest that the authorities used the freedom the law gave them to pick the winner (Hyytinen et al., 2007).
[6] Exceptions include Cripps and Ireland (1994), de Boer at al (1998), Dimitri et al (2006) and Bergman and Lundberg (2013)

objectives.[7] The model has, e.g., successfully been applied to budgeting decisions. Past budget decisions are good predictors of future budgets (Green and Thompson, 2001).

More generally, research within the neo-institutional tradition (March and Olson, 1989; Powell and DiMaggio, 1991) emphasizes the importance of institutional inertia. For this reason, it is interesting to contrast the predictions of rational utility maximization against predictions based on institutional habits.

A third perspective of relevance is transactions-cost theory (e.g., Williamson, 1981). A simple prediction suggests that since transactions costs are likely to be higher when EMAT is used, there should be a tendency that the winning bid is based on lowest price more often in low-value procurements and vice versa. A version of this argument would suggest that lowest price is a more common criterion for authorities that are less accountable to those that ultimately benefit from the procurement and bear its costs, i.e., citizens and taxpayers, and more common among authorities that are more accountable to their nominal superiors, such as ministries of government – that may be more focused on budget targets. In practice this would suggest that municipalities, being more accountable to citizens than central-government authorities, favour EMAT, while the latter category would tend to rely more on lowest price.

We turn now to the predictions of a rational-actor perspective. When discussing quality optimization, we ignore the problem that delivered quality may deviate from contracted quality; this problem is instead the focus when we discuss moral hazard (asymmetric information).

Quality optimization

Assume initially that the buyer knows with certainty the sellers' costs of producing at different quality standards. The contracting authority may then optimize quality first, specify the required quality accordingly in the call for tender and contract with the bidder that is willing to meet these requirements at the lowest cost. Alternatively, the procurer may specify a scoring rule that rewards quality in excess of some point of reference, such that the rewards correspond to the buyer's valuation of the additional quality. Under this alternative mechanism, quality optimization will be "built-in" into the bid-evaluation process. However, if the cost of quality is known, the outcome should be the same and, in order to minimize transaction costs, lowest-price procurement would appear to be the preferred method.

---

[7] Simon (1947) holds the more positive view that organizational routines, or habits, are efficient responses to complex tasks.

Asker and Cantillon (2008, 2010) demonstrate that, under cost uncertainty and in principle, scoring rules dominate beauty contests, price-only auctions as well as menu auctions and that, when bidders have private information of a multi-dimensional nature and quality is contractible, scoring rules outperform bargaining models. Furthermore, scoring rules should be designed so as to correspond to the utility function of the contracting authority.[8]

Bergman and Lundberg (2013) discuss how, in practice, a scoring rules can be tailored to the contracting authority's utility function as well as the trade-off between simple and hence low-transaction cost but less precisely targeted selection criteria – and more complex, high-transaction cost but better targeted criteria. In analogy with Weitzman's (1974) analysis of regulation versus taxation, lowest-price procurement is likely to be close to optimal when the marginal cost of quality does not rise particularly fast (or not at all) and when the marginal value of additional quality falls rapidly above a critical quality level. A possible example is staffing-as-quality: the marginal cost of additional staff is approximately constant, while the benefit of additional staff beyond the required level can be expected to fall sharply. Here, the solution would be a lowest-price procurement for a contract with a fixed staffing requirement. Another example would be that a machine is required to meet legal safety and performance standards, but where the additional value of quality in excess of those standards is small.

Using the same logic, a scoring function that is linear in a quality measure is likely to be close to optimal when the marginal *value* (or utility) of that quality aspect is likely to be relatively constant and when the marginal cost of increasing quality is likely to rise steeply. A possible example is energy-efficiency-as-quality. In most situations, each unit of energy that is saved is equally valuable, while the cost of achieving additional energy efficiency is likely to increase rapidly once the lowest-hanging fruits have been picked. A practical way to implement this is to rank bids according to life-cycle costs.

It follows that there is no simple normative relation between how important quality is perceived to be and the optimal bid selection mechanism. What is optimal depends on the steepness of the marginal-cost curve relative to that of the marginal-benefit curve; a relation that, unfortunately, is difficult to assess for a large empirical sample of procurements.

---

[8] If competition is less than perfect and if the buyer can commit not to renegotiate the contract with the winner – a commitment that may in fact be the consequence of following the public procurement rules – there may be strategic reasons to slightly distort the utility of quality downward as this will intensify competition; see Laffont and Tirole (1993), and Che (1993).

Summing up, EMAT should be more likely when the cost of quality is uncertain, but not necessarily when quality is high on the buyer's priority ranking.

Moral hazard – uncertainty about delivered quality

If important quality aspects are non-verifiable (i.e., if there is moral hazard or adverse selection), the contracting authority may *not* want to specify the scoring function in too much detail, in order to be able to use reputational mechanisms to support a high-quality equilibrium; such mechanisms require that the buyer retains a degree of discretion. In private transactions, where buyers have substantial discretion and can react to non-verifiable quality signals, reputation, brand names and long-term informal relations are used to sustain high-quality equilibria through a link between current performance and future sales (Klein and Leffler, 1981; Macaulay, 1963; MacLeod, 2007). There exists an implicit contract or a market mechanism that ensures that high quality will be rewarded by large sales in the future. In plain English: If the supplier degrades quality without formally violating any clause of the contract or breaching any law, the buyer will use its discretion to punish the supplier in the future.

In contrast, public procurement legislation requires procedures to be objective and transparent for accountability reasons, limiting discretion and thereby the scope for such mechanisms (Banfield, 1975; Kelman 1990; Tadelis, 2012). In most countries a public procurer is in principle not allowed to discriminate in favor of strong brand names, nor in favor of providers that performed well in the past in non-verifiable dimensions. Similarly, while a public-procurement contract can give the buyer an option to extend the duration of the supply contract, the exact length of the extension must typically be specified in the original contract. Under many public procurement legislations – although not the Swedish – the criteria driving the decision to use the extension option must be 'objective', that is, verifiable. Even where a contracting authority has the possibility of linking future sales to provided quality, e.g., via vendor-rating and contract-renewal schemes, existing regulations make this link very tenuous for non-contractible dimensions that cannot be audited by third parties and therefore generate accountability concerns.

A possible way to reward non-verifiable quality – or, alternatively, to give the contracting authority more discretion – that has not received much attention in the academic literature is intentional vagueness as to the definition and measurement of quality. If the contracting authority retains the right to award quality scores on the basis of criteria that are not objectively verifiable it can, in fact, screen for low-quality providers and implement a

mechanism that gives the winning bidder incentives to maintain quality throughout the contract period. (Assuming that the bidders perceive the buyer's commitment to this mechanism as credible.)

Other mechanisms that have been proposed in the literature include "relational" procurement, i.e., procurement with explicit discrimination in favor of bidders with a good performance track record (Albano, Beradino and Alberto, 2011) and "reputational" procurement, i.e., procurement from a pool of eligible bidders under the threat of exclusion from the pool in response to the provision of deficient quality (Calzolari and Spagnolo, 2009; Spagnolo, 2012).

Supplier selection, testable hypotheses and research design

We summarize the above normative discussion into two testable hypotheses. If costs are uncertain, a scoring rule that represents the authority's utility function, i.e., EMAT, can be used to optimize quality.

> *Hypothesis 1 (H1):* A high level of uncertainty concerning the cost of different levels of quality makes the use of the lowest-price supplier selection method less likely.

In the face of moral-hazard concerns, the authority needs a mechanism that can sustain high-quality equilibrium. It must, therefore, give itself sufficient discretion in choosing supplier, making lowest price a poor choice, which leads us to the second hypothesis.

> *Hypothesis 2 (H2):* When quality is highly non-verifiable the contracting authority is more likely to use high quality weights, including Quality-only scoring, and/or to use qualitative methods to assess quality (in combination with P2S or Q2P).

These hypotheses are founded on a rational-choice perspective. However, it is entirely possible to view the choices of the authorities from an institutional perspective. How procurements are organized and what scoring rules are chosen may be driven by local habits or traditions or organizational routines. This leads up to the following hypothesis:

> *Hypothesis 3 (H3):* An authority will tend to use similar scoring rules when procuring different types of products.

Based on our brief discussion of the possible implications of transaction cost, we also propose the following hypothesis:

> *Hypothesis 4 (H4)*: Large authorities and authorities that are more directly accountable to citizens, such as municipalities, will tend to rely more often on complex supplier-selection methods (EMAT). Small authorities and authorities that are mainly accountable to the central government will tend to select winning bids on the basis of lowest price.

Even though our normative discussion does not support the idea that the scoring rule should explicitly consider quality when quality is seen as more important, and when containing costs is seen as less important, there is a popular perception that this is indeed the case. Hence, we propose the following hypothesis:

> *Hypothesis 5 (H5)*: Complex supplier-selection methods (EMAT) are used more often when quality is important relative to the desire to contain costs.

How to test Hypothesis 3 warrants a brief discussion. Given the nature of our dataset – and, we believe, also more generally – it is difficult to establish independently what award mechanism is optimal for a given product. As will be explained below, we ask procuring municipalities to assess, for each product, the severity of the moral hazard problem and the uncertainty concerning the cost of producing at different quality levels. There is a risk, however, that the replies reflect a tendency to rationalize the chosen award mechanism. Then the assessment of a product's characteristics would tend to be driven by the chosen award mechanism, rather than the award mechanism being driven by the characteristics of the product.

Figure 1 illustrates the problem. The hypotheses we want to test correspond to links from product characteristics and habits to the procurement design. However, neither habits nor product characteristics can be observed; we can only observe the authority's assessment of the product's characteristics and the identity of the authority. This would not be a concern, if it were not for the risk that the authority's assessment of product characteristics can be influenced both by its own habits and by the tendency to rationalize its choice of actions by making assessments of product characteristics that match with the chosen method, rather than with the characteristics of the product. These reverse causations are illustrated with the lightly shaded arrows pointing upward and to the left.

Figure 1. Determinants of procurement design
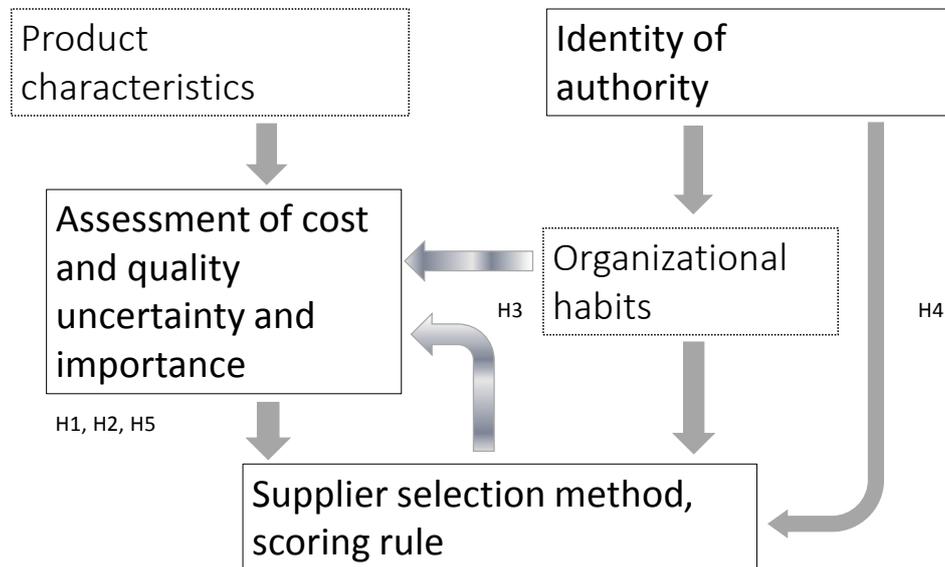
In order to disentangle the effect of habits (or institutions) and product characteristics we need assessments of product characteristics that are independent of the authority's actual choice of supplier selection methods. For this reason, as will be detailed below, we have conducted two surveys: one directed to authorities from which we have also collected detailed data from tender documents and one directed to other authorities.

3. THE SURVEYS AND THE DATA

We use data collected in two surveys in combination with data extracted from documents associated with the procurements covered in one of the surveys. The data which is the starting point for the first survey was collected in a two-step procedure. First, 40 Swedish authorities were randomly sampled with a probability proportional to their size. Second, procurements from the 2007 to 2008 period were randomly sampled (with equal probabilities) from the selected authorities' records. Of the sampled authorities, 39 made data available, by giving us access to procurement documents, i.e., the call for tenders, the decision records and the tender compilations.

In the end, the net sample consisted of 651 procurements undertaken by 14 municipalities, six counties and 19 central-government authorities, with a maximum of 20 procurements per authority and fewer if the authority made less than 20 procurements during the period studied.

The data includes detailed information for all of these procurements concerning the nature of the product, the design of the procurement, the bids etc. (See Bergman and Stake, 2014, for details.)

Although we have data on a large number of procurements in our original data, it covers only a limited number of contracting authorities and only in a few instances is the same product procured by more than one authority. Furthermore, we have no independent assessment of what is the optimal procurement mechanism for a particular product. Absent additional data, it will hence be difficult to disentangle the effects of habits from the effects of the characteristics of the products and our research design would be susceptible to rationalization effects.

To address these data limitations, we surveyed the contracting authorities as well as other authorities about their assessments of the products, in terms of cost and quality uncertainty. This gives us two possible ways to check the robustness of our results. First, to address a possible rationalization bias, we can use *other* authorities' assessment of product characteristics in a regression analysis. Second, if we have multiple assessments of each products' characteristics, we can estimate the separate effects of the identity of the authority and of the product on the *assessment* of product characteristics, with the latter effect corresponding to the right-to-left arrow in Figure 1.

Survey 1, which was sent to the 39 authorities, included three questions for each of the procurements for which we had data. We asked the Director of procurement to answer the questions or to delegate that task to a senior procurement staff member. (See Appendix 2 for a presentation of the survey questions.) The contact was initiated with a call to the Director, followed by an e-mail with a link to a web-based survey and then, if required, by up to three reminders.

The first question concerned to what extent it was important to get exactly the sought-for quality relative to how important it was to keep within the budget (Quality vs. budget). The second question was how knowledgeable the authority was about the cost, for different quality levels, of the product (Price uncertainty). Finally, we asked to what extent the quality of the product was *verifiable* ex post (Quality uncertainty).

The overall response rate to our questions in Survey 1 was 77 percent, corresponding to 30 responding authorities which, between them, had undertaken 467 out of the 651 procurements

on which we have data. We received about 400 answers to our questions. That is, the response rate for individual questions, conditional on responding to the survey, was just over 80 percent, for an effective average response rate of just over 60 percent for individual questions (procurements).

Some of the procurements included more than one type of products. For each of the individual procurements we identified a main product and for this product we asked the respondent to assess the three questions on a 1-9 scale. The scale for each question had the following meaning:

(i)   Quality vs. Budget (Price): 1 = Price is irrelevant, quality is crucial. 9 = Price is crucial, quality is irrelevant.

(ii)   Price uncertainty: 1 = No pre-tendering uncertainty about the cost of quality. 9 = Extreme pre-tendering uncertainty about the cost of quality.

(iii)   Quality uncertainty: 1 = No uncertainty about delivered quality. 9 = Extremely high uncertainty about delivered quality.

Survey 2 was directed to municipalities not included in the first survey. We asked them to respond to the same set of questions for up to 20 products that had been procured by one of the 14 municipalities in the gross sample from the first survey. We also asked whether the authority had itself purchased the product in question or not. We refer to the municipalities in the second survey as the *control group*, while the authorities in the first survey are referred to as the *contracting authorities*.[9]

In order to improve the fit between the contracting authorities and the control group, we first divided the municipalities into two size categories: the 100 largest municipalities, with populations in excess of approximately 25,000 and the 175 smallest. The sets of products procured by the ten large municipalities among the contracting authorities were used as templates for the questions sent to the 90 largest municipalities in the control group. Each of the ten municipalities among the contracting authorities was hence used for one tenth of the municipalities in the control group. Similarly, four small municipalities in the first survey were used as templates for other small municipalities in the second survey.

---

[9] We use the expression "control group" for simplicity, even though it may not be a control group in the strict sense. It is, however, a group of municipalities that is similar to the municipalities addressed in Survey I, but which have no reason to let their responses be influenced by design choices made by municipalities in the group of contracting authorities. I.e., this is a group of respondents that, we believe, should not be subject to rationalization bias when assessing product characteristics.

We received answers from 107 out of the 275 municipalities addressed in the control group, a response rate of one third. From the responding authorities we could in principle have gotten a maximum of 1,419 replies on a specific procurement; in fact we received 1,336 responses for individual procurements. However, matching individual procurements and questions gave us response rate that was lower; for most questions we received about 1,050 replies.

Out of the initial 651 procurements, 221 were undertaken by municipalities. However, since the response rate in Survey 1 was lower among municipalities we received replies for just 86 procurements organized by municipalities, for an overall response rate of 39 percent. Five products had to be deleted from the second survey because they were specific to one of the municipalities; one example was a large commuter-train tunnel. Since several municipalities in the control group were surveyed about the same procurements from the initial population, we got a response from at least one municipality for each of the remaining 216 municipal procurements in the control group. Figure 2 illustrates the relation between the two surveys, as seen from the perspective of individual procurements.

Figure 2. Populations, sampling frames and responses for the two surveys.[10]

---
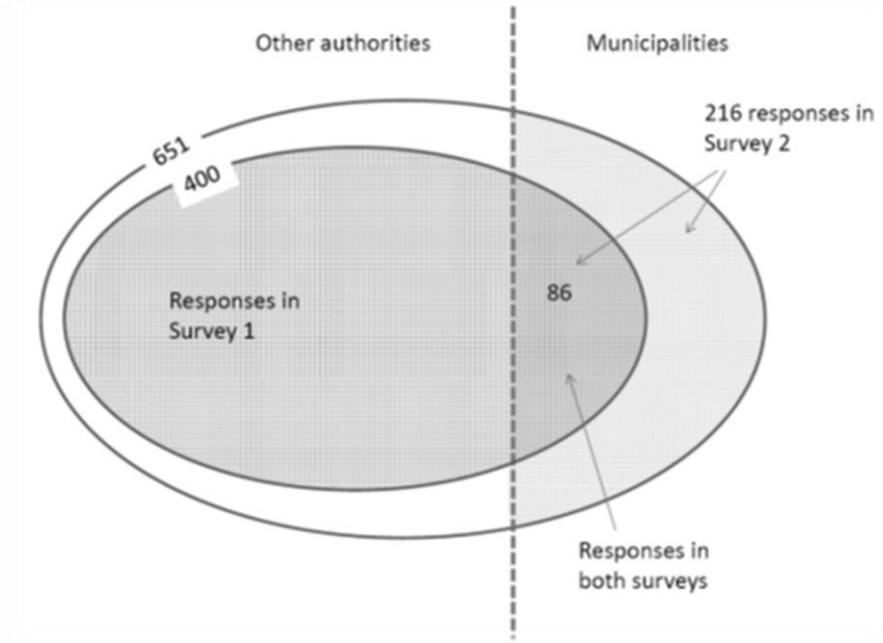
[10] The five procurements excluded from Survey 2, as discussed in the text, are ignored in the figure. Note also that in Survey 2 the same question was sent to multiple municipalities. Hence we had on average close to five responding municipalities for each of the 216 responses.

Descriptive statistics, the contracting authorities (Survey 1)

Due to the diverse nature of the population from which the procurements in the first survey were drawn, there is of course large heterogeneity among the products. Table A1 in the Appendix reports the number of procurements per two-digit level according to the CPV nomenclature.[11] The most common product categories are business services (12 percent), construction works (11 percent), medical equipment and pharmaceuticals (9 percent), health and social works services (9 percent) and architectural, constructions, engineering and inspection services (6 percent). Only three of the product categories in the nomenclature list are not observed in our sample.[12] Table 1 reports descriptive statistics for variables extracted from the procurement documents.

Table 1. Descriptive statistics, the procurement data.

Variable	Mean	Std.dev.	Min	Max	N
Supplier selection method					
Lowest price	0.38	0.49	0	1	651
EMAT	0.62	0.49	0	1	651
Price to score	0.51	0.50	0	1	651
Quality to price	0.10	0.29	0	1	651
Quality only	0.02	0.13	0	1	651
Other variables					
Joint procurement	0.07	0.26	0	1	651
Framework contract	0.26	0.44	0	1	651
Service contract	0.39	0.49	0	1	651
Goods contract	0.19	0.39	0	1	651
TED (published within EU)	0.40	0.49	0	1	631
Contract duration (years)	1.59	1.30	0	7	648
Extension period 1 (years)	0.79	0.74	0	4	648
Extension period 2 (years)	0.31	0.55	0	3	647

As can be seen, supplier selection was based on lowest price in 38 percent of the procurements and, consequently, EMAT was used in the remaining 62 percent. Within the

---

[11] "The CPV consists of a main vocabulary for defining the subject of a contract, and a supplementary vocabulary for adding further qualitative information. The main vocabulary is based on a tree structure comprising codes of up to 9 digits (an 8 digit code plus a check digit) associated with a wording that describes the type of supplies, works or services forming the subject of the contract." See http://simap.europa.eu/codes-and-nomenclatures/codes-cpv/codes-cpv_en.htm.

[12] These are: Agricultural, farming, fishing, forestry and related products (CPV 3000000-1); Installation services (except software) (CPV 51000000-9); Public utilities (CPV 65000000-3).

latter category, scoring rules that transform bid prices into price scores, "price to score", dominated with more than half of all tenders. Scoring rules that conversely assigned values to quality characteristics, "quality to price", came second, being used in about one tenth of all tenders, while quality-only scoring (or beauty contests) was used in no more than 2 percent of the procurements.[13]

About 7 percent of the procurements were jointly organized by two or more contracting authorities and about a quarter were procurements for framework agreements. Two out of five procurements were for service contracts, one out of five was for goods contracts and the rest was a combination of goods and services (including works).

Figure 3 illustrates the distribution of price weights in the sample. Besides showing that lowest price is much more common than highest quality, the figure also shows that when the P2S method is used, price is typically given a percentage weight in even tens in the 40 to 70 range. The price weight cannot be calculated when the quality-to-price method is used and the price weight was also missing for some other cases.

Figure 3. Frequencies of price weights in the sample. (*N*=556.)
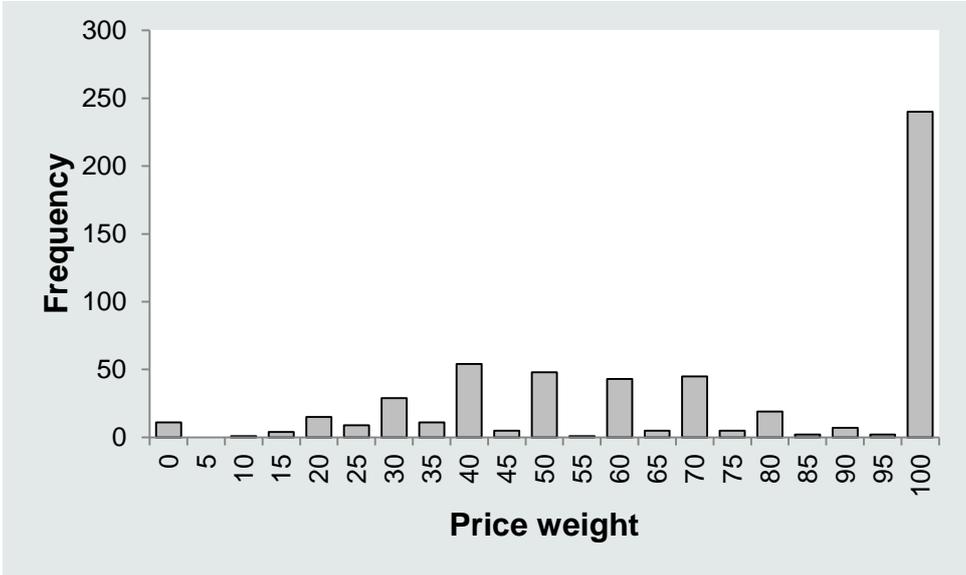
The choices of supplier selection method and scoring rule differ significantly between authorities, as can be seen in Figure 4. The most common scoring rule, when EMAT is used, is some sort of price to score rule.

---

[13] Note that since the original sample of 39 authorities was established with probability sampling, with a probability proportional to the size of the authority, we cannot claim that the numbers reported in Table 1 are best estimates of the distribution in the population of all Swedish procurements. However, our sample can be used to draw inferences on marginal effects of explanatory variables; see Cameron and Trivedi (2005, p 820).

Figure 4. Supplier selection method and scoring rule by authority
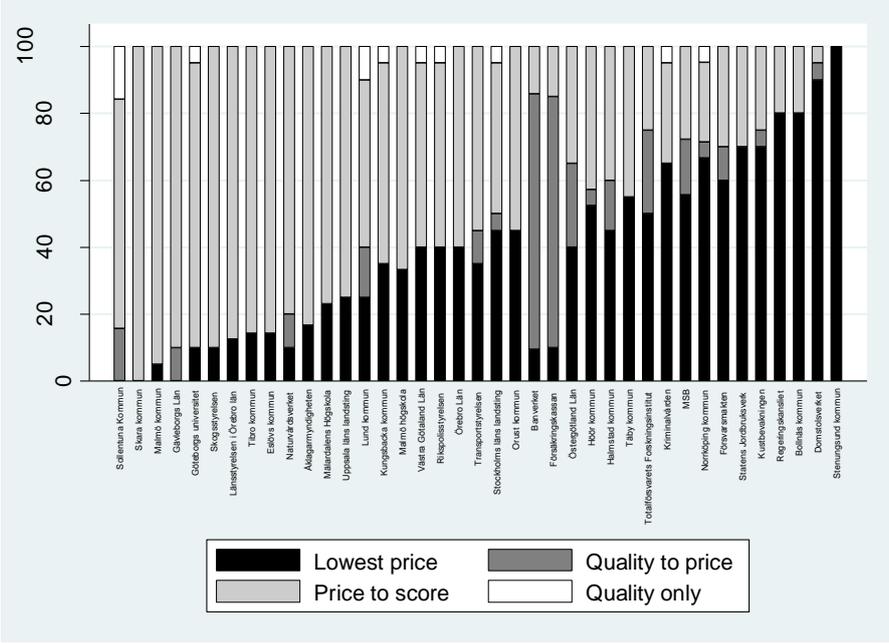
Regional authorities tend to prefer price-to-score methods; the average county used this method for about 70 percent of the procurements. Local authorities also use this method predominately, on average in about 55 percent of the procurements. In contrast, central government authorities have a higher tendency to use lowest price; the average central government authority uses lowest price in almost 50 percent of the procurements.

All public procurements within the EU with an expected contract value exceeding a certain threshold value must be published in Tenders Electronic Daily (TED), but different threshold values are used for different products. For example, the threshold value for "Works contracts, works concessions contracts, subsidized works contracts" was 500,000 EUR in 2012, while the threshold was only 200,000 EUR for "All service contracts, all design contests, subsidized service contracts, all supplies contracts". See EU (2012). *TED* is here defined as a dummy variable taking the value 1 if the procurement notice was published in TED (or at least should have been). This is also a proxy variable for contract value; information that in general is missing in our data. A majority of the procurements (60 percent) fall below the threshold. The average duration of the procurement contract was 1.59 years to which, on average, an

additional 0.79 years could be added in a first extension period and a further 0.31 years in a second extension period.[14]

Descriptive statistics for survey responses

The distributions of the responses from both surveys, for the three questions, are displayed in Figure 3 and descriptive statistics are reported in Table 2.

Very few of the authorities reported price to be irrelevant and quality to be crucial, corresponding to response 1 in panel *i* of Figure 5. It was more common, but still rare, to report price to be crucial and quality to be irrelevant. Extremely high uncertainty about price or quality was similarly rare, while the opposite – no uncertainty in these dimensions – was a little bit more common, as shown in Panels *ii* and *iii*.

Figure 5. Distribution for survey questions (scale 1 − 9).

(i)   Quality versus Budget (Price)

(ii)   Price uncertainty

(iii)   Quality uncertainty

(i)   Budget (cost) is important to the right, quality to the left.

(ii)   Price uncertainty increases to the right.

(iii)   Quality uncertainty (moral hazard) increases to the right.

---
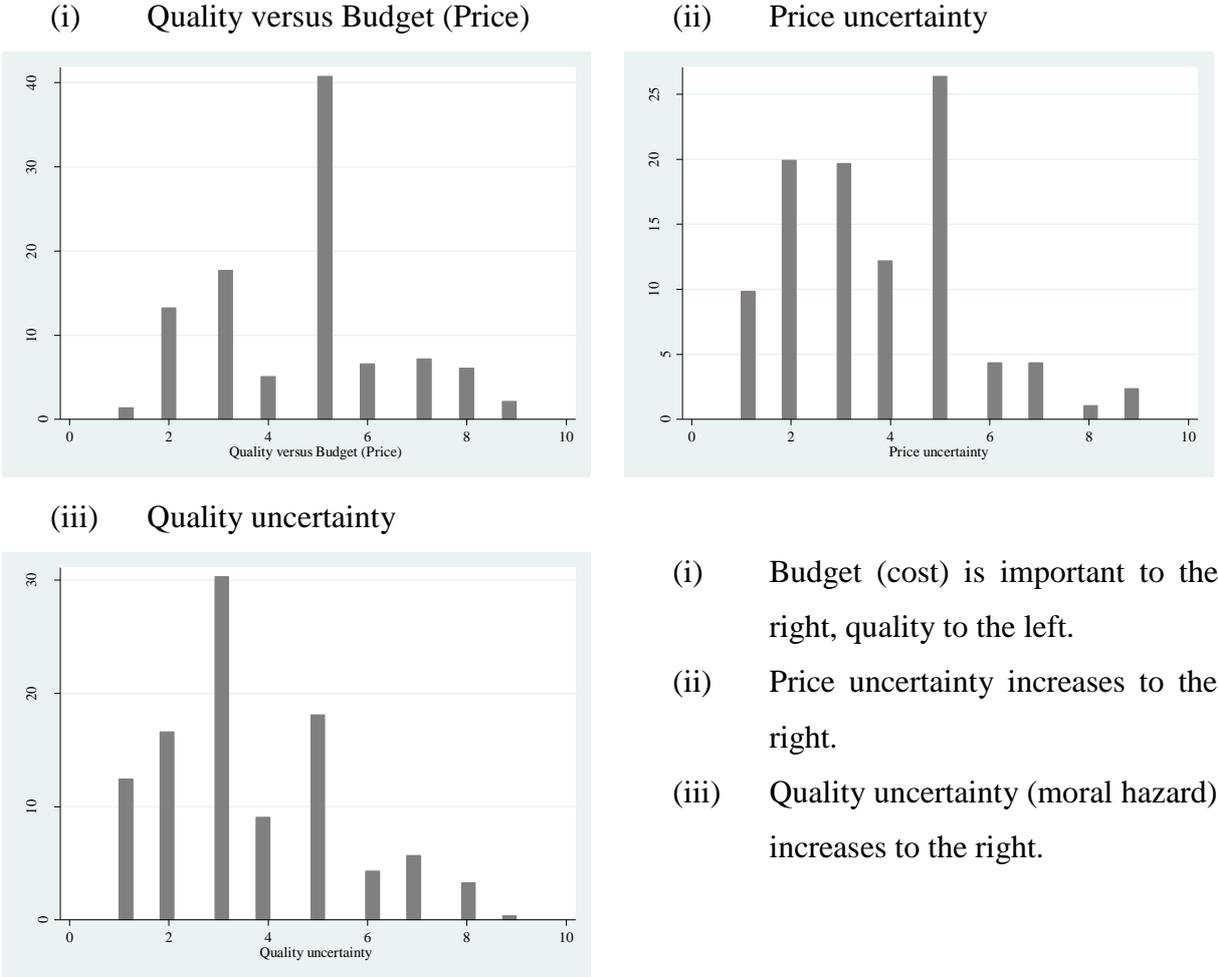
[14] The average is based on all procurements.

Table 2 reports average responses for the three survey questions, separately by type of product (services and goods), by choice of supplier selection method (lowest price and EMAT) and by the main categories of scoring rules (P2S, Q2P and Quality-only scoring).

On average, quality is perceived as more important than is keeping within the budget and the authorities seem to be relatively certain about price as well as about quality. The averages for "Price uncertainty" and "Quality uncertainty" range approximately between 3 and 4 with somewhat higher values (higher uncertainty) for goods contracts. The lowest figure is found for lowest-price procurements (2.92) and the highest is found in procurements of goods (4.20). It seems reasonable that the contracting authorities report low price uncertainty when it chooses to use the lowest-price method.

The difference between the control group and the contracting authorities is highly significant for Price uncertainty (*t*-value is |3.86|) and for Quality uncertainty (*t*-value is |5.14|), but only weakly so for the question concerning Quality versus Budget preferences (t-value is |1.93|).

Table 2. Descriptive statistics for the survey questions

	Survey 1	Survey 2		
	All	All	With experience	Without experience
Quality vs. Budget				
Mean	4.39	4.59	4.50	4.82
Std.dev.	1.90	1.62	1.66	1.47
N	396	923	659	258
Price uncertainty				
Mean	3.48	3.87	3.59	4.57
Std.dev.	1.55	1.78	1.70	1.77
N	398	921	659	256
Quality uncertainty				
Mean	3.41	3.95	3.68	4.65
Std.dev.	1.68	1.80	1.73	1.78
N	400	908	653	249

We also compared the reported values for "Price uncertainty" and "Quality uncertainty" in procurements using the lowest-price supplier selection method relative to those that used EMAT. There was a significant difference in means, with lowest-price procurements located closer to "No uncertainty about price" and "No uncertainty about delivered quality" than EMAT procurements.[15]

---

[15] The *t*-values are |3.52| and |4.52|, respectively.

We found, somewhat surprisingly, that quality uncertainty is higher for goods than for service contracts and that the focus on price is stronger for service contracts than for goods contract.[16] One would typically expect the opposite, since non-verifiable quality characteristics are likely to be more important for services.[17] The difference in mean in "Budget vs. Quality" between service and goods contracts is significant at the 10 percent level.

As noted, the number of respondents in the control group is 107 and between them they reported at least one value for each of the 216 procurements. In addition to grading the three questions, the municipalities in the control group were asked whether they had experience of procuring similar contracts. This information is available to us in 1,321 cases. The difference in mean between municipalities in the control group, with and without own experience, respectively, is significant for all three survey questions (Quality vs. Budget, Price uncertainty and Quality uncertainty.[18] Municipalities without own experience tend to emphasize cost control more, while being more uncertain about price and quality.

4. EMPIRICAL MODELS

We test our hypothesis on the data extracted from the procurement documents and the two surveys. We model the choice of supplier selection method with two different models:[19]

(i)   As a choice between EMAT and lowest price. Estimation method: Logit.
(ii)   As a choice of price weight, with the supplier-selection method following implicitly from that choice. Price weights equal to 100 percent correspond to lowest price and weights in between implies that the price-to-score method was used. The quality-to-price procurements are not included in the regression.[20] Estimation method: Tobit.

The outcomes from regression models *(i)* and *(ii)* can in principle allow us to reject, or not, Hypotheses 1, 2, 4 and 5.

---

[16] The difference is significant with a *t*-value of |4.23|.

[17] To some extent the result may be due to the classification of wholesale supply of goods as "goods", rather than "services".

[18] The *t*-values are |2.71| (Quality vs. Budget), |7.70| (Price uncertainty) and |7.45| (Quality uncertainty).

[19] We have also tried using four non-ordered alternatives, similar to the second step in model *i*, but with lowest price as the fourth alternative. Just like the second-step model, this model performed poorly.

[20] For any quality-to-price scoring rule an equivalent price-to-score mechanism can in principle be found. However, the weights for price and quality, respectively, cannot meaningfully be defined without simultaneously defining the scales along which these are measured and since the choices of scales are arbitrary, the price weight in the equivalent price-to-score model is also arbitrary. It follows that the price weight is less meaningful than it at first appears to be; a problem that we to a large extent ignore.

The dependent variable in the logit regression (*Model i*) is discrete and takes the value 1 if lowest-price is the supplier selection method and, consequently, the value 0 if EMAT is the observed method. The dependent variable in the Tobit regression (*Model ii*) is the price weight. This approach is motivated since weights can neither exceed 100 percent nor be lower than 0. Weights equal to 0 correspond to Quality-only scoring and weight equal to 100 correspond to supplier selection based on Lowest price.

We use the same set of control variables for both models, with the key explanatory variables drawn from the two surveys complemented with authority, procurement and contract characteristics. We include one variable for each of the three survey questions, a dummy for TED publication (yes = 1), the length of the contract period, the total length of the extension period, a dummy variable for type of contract (goods, services or mixed, with the latter as the reference category), a dummy variable for joint procurement (yes = 1), a time dummy variable (2008 = 1), a dummy variable for type of contracting authority (local/municipality, regional/county or central-government authority, with local authorities as the reference category) and, finally, a measure of the size of the contracting authority.[21]

Two different specifications are used. In the first one, Specification 1, the survey-question variables take the values reported by the respondents in Survey 1. This reflects the views on price and quality uncertainty of the contracting authorities, i.e., the authorities that actually made the procurements. In a second specification, Specification 2, we calculate per-product averages across all municipalities that provided responses for a particular product, irrespective of whether they actually procured the product or just responded to a hypothetical question.

In addition to these models we specify three equations with the responses to the three survey questions as dependent variables. The aim is two-fold. Firstly, to analyze the effect of experience on the perceived uncertainty about price and quality and on quality-versus-budget preferences. Secondly, to explore the links between the responses to the three survey questions. For this purpose we use Seemingly Unrelated Regression (SUR).

The SUR is estimated with a set of explanatory variables that includes the type of product, TED, contract duration and option period, whether the respondent in the control group had experience of similar procurements or not, type of contracting authority and a dummy variable that identifies the control group.

---

[21] We use population for municipalities and counties and number of employees for central-government authorities, normalized so that the sum of weights for central authorities is half of the total sum of weights.

To shed further light on what explains the responses to the survey questions we estimate models with authority and product fixed effects and calculate the partial R^2. This measure allows us to assess whether authority-specific or product-specific factors accounted for most of the otherwise unexplained variation. If respondent-specific factors explain more of the variation, this indicates that the authorities' choice of method tends to be driven by local habits and, consequently, is less well adapted to the product. If product-specific factors can explain more of the variation, a possible interpretation is that genuine differences between products are important in shaping the contracting authorities' behavior.

5. RESULTS

The results from estimating *Model i* are presented in Table 3. A high level of uncertainty concerning the cost of different levels of quality makes the use of the lowest-price supplier selection method less likely (*H1* cannot be rejected). Also, when quality is highly non-verifiable the contracting authorities are more likely to use EMAT (*H2* cannot be rejected).

Table 3. Results, logit regression. Lowest price vs. EMAT as supplier selection method. Lowest price=1 (yes) or zero (EMAT). $^*p< 0.10$, $^{**}p< 0.05$, $^{***}p< 0.01$

Lowest price	Specification 1, reported values on (1) – (3), Survey 1		Specification 2, average values on (1) – (3), both Surveys	
	Coef.	*SE*	Coef.	*SE*
Quality vs. Price	0.06	0.06	0.08	0.06
Price uncertainty	-0.19**	0.09	-0.21**	0.10
Quality uncertainty	-0.20**	0.09	-0.17*	0.10
TED = 1	-0.22	0.26	-0.43*	0.23
Contract period	0.39***	0.13	0.36***	0.11
Extension period	-0.47***	0.16	-0.56***	0.14
Joint procurement	-0.12	0.57	0.06	0.39
Goods contract	-0.44	0.35	-0.22	0.29
Gov. Authority	0.58	0.36	0.58*	0.30
County	-0.05	0.43	0.03	0.37
Authority size	0.02*	0.01	0.02**	0.01
Year	0.21	0.23	0.37*	0.20
Constant	-0.21	0.62	-0.21	0.56
LR chi2(12)	50.04			62.75
Prob> chi2	0.00			0.00
Pseudo R2	0.10			0.10
Log likelihood	-233.27			-301.58
N	392			503

The significantly positive effect, in one specification, of the government-authority dummy lends some support to *H4*, while the positive effect of the authority sign refutes our simple application of transaction-costs theory. That we find no significant effect of the importance of quality (*H5*) is in line with normative theory, but not with the popular view that the scoring rule should put a large weight on quality when it is important.

The results also suggest that lowest price is more common for contracts with long duration – but less common for contracts that include an optional extension period.

To evaluate the economic significance of the perceived uncertainty about price and quality and the municipalities' relative emphasis on budget or quality, we use the estimated model to predict the probabilities that lowest price is used for high and low response values, respectively. Table 4 reports, e.g., the probability that lowest price is used when the quality uncertainty is set at *i*) one standard deviation below the mean and *ii*) one standard deviation above the mean, with all other variables set at the sample mean. In this example, when the municipality's perception of quality uncertainty rises from low to high, the probability that lowest price is used falls from 45 to just below 30 percent. Increasing price uncertainty has a similar effect, while the effect of the municipality becoming highly focused on the budget has a smaller effect.

Table 4. Estimated probability that lowest price is the choice of supplier selection method at low and high values on the survey questions.

Question	Mean-1 std.dev.	Mean+1 std.dev.
Quality vs. Price	0.3433	0.3958
Price uncertainty	0.4384	0.3049
Quality uncertainty	0.4502	0.2949

Estimating Specification 2 of the model (the columns to the right in Table 3, based on average responses to the survey questions) makes little difference for the parameter estimates. However, more coefficients are now significant. According to this specification, lowest price is less likely to be observed when the contract value exceeds the threshold (TED=1), is more likely when central-government authorities procure and is more common in 2008 than it was in the preceding year.

The conclusion, so far, is that we can reject neither *H1* nor *H2*. The choice between lowest price and EMAT can, to some extent, be explained by the contracting authority's perception

of price and quality uncertainty and the sign of the effect is in accordance with the theoretical predictions of how a rational authority should behave.

Model ii, which uses Tobit regression, largely confirms these findings. See Table 5. A higher weight for quality is used when the contracting authority is more uncertain about costs or quality – in line with *H1* and *H2*. Compared to *Model i,* qualitatively similar effects are found for the contract period and contract extension variables. Also, at a 10 percent significance level, authorities that emphasize budget consequences rather than quality aspects tend to increase the price weight, in line with *H5*. In one specification and for one variable (size of authority), we find similar weak support for *H4*.

Table 5. Results, Tobit regression, $^*p< 0.10$, $^{**}p< 0.05$, $^{***}p< 0.01$. Dependent variable is the price weight.

Price weight	Specification 1		Specification 2	
	Coef.	*S.E*	Coef.	*S.E*
Quality vs. Budget	2.21	1.43	2.62*	1.44
Price uncertainty	-3.92**	1.93	-4.99**	2.25
Quality uncertainty	-5.20***	1.82	-4.99**	2.22
TED = 1	-6.80	5.95	-9.42	5.45
Contract period	8.27***	3.07	7.24***	2.56
Extension period	-9.02**	3.62	-12.26***	2.87
Joint procurement	-1.69	13.47	6.23	9.46
Goods contract	-6.84	7.71	-5.30	6.60
Gov. Authority	9.65	7.79	12.28	6.90
County	4.63	9.38	9.43	8.58
Authority size	0.28	0.21	0.40*	0.16
Year	3.92	5.18	5.52	4.68
Constant	91.68	14.04	93.16	13.42
Sigma	44.18	2.49	45.06	2.26
LR chi2(12)		43.77		60.27
Prob> chi2		0.00		0.00
Log likelihood		-1149.30		-1460.89
Pseudo R^2		0.02		0.02
N		347		447
4 left-censored observations			4 left-censored observations	
198 uncensored observations			250 uncensored observations	
145 right-censored observations			189 right-censored observations	

Using data from both surveys and simple fixed-effect models we seek to understand what determines – in a particular procurement – the procurer's attitudes concerning quality and

price uncertainty, as well as concerning the trade-off between quality and costs. Are the attitudes mainly linked to the authority itself or are they more strongly linked to the item procured? Since the survey responses, in turn, influence the choice between the two supplier-selection methods, the answer to this question provides a deeper understanding of the choices made by the procurers.

Table 6 reports R^2 and partial R^2 values for simple models that try to predict responses for the survey questions. This gives us information about the extent to which municipal and product effects, respectively, contribute to explain otherwise unexplained variation in quality over price preferences and in price and quality uncertainty. While we cannot test H3 directly, we can test the related hypothesis that a given authority tends to perceive different products to have similar characteristics. If there are "habits" in perceptions – what problems are perceived to be important – this will likely result in consistent patterns in actual behaviour. I.e., there will effectively be habits also in behaviour.

Table 6. R^2-values and partial R^2 from regressions with authority (municipal), product and fixed effects. See Table A2 in the Appendix for full presentation of results, N = 1005 – 1017.

	Controls	Excluded control, R^2-values			Excluded control, Partial R^2		
	Authority, product, experience	Authority	Prod.	Exp.	Authority	Prod.	Exp.
Product is CPV							
Quality vs. Budget	0.51	0.13	0.40	0.51	0.44	0.18	0.01
Price uncertainty	0.46	0.10	0.42	0.43	0.39	0.06	0.05
Quality uncertainty	0.44	0.08	0.41	0.43	0.40	0.05	0.03
Product is ID (1)							
Quality vs. Budget	0.63	0.29	0.40	0.63	0.47	0.38	0.00
Price uncertainty	0.56	0.27	0.42	0.54	0.39	0.23	0.04
Quality uncertainty	0.55	0.23	0.41	0.54	0.42	0.23	0.03
Product is name (2)							
Quality vs. Budget	0.59	0.22	0.40	0.58	0.47	0.31	0.01
Price uncertainty	0.53	0.22	0.42	0.51	0.40	0.19	0.04
Quality uncertainty	0.52	0.18	0.41	0.51	0.41	0.18	0.03

Each entry in the first four columns corresponds to a separate regression. In the first column three independent (sets of) variables are used: a dummy variable for each municipality, a product dummy variable and a dummy variable signaling that the municipally has actually procured the product in question; i.e., that it is not just responding to a hypothetical question. Columns 2-4 report similar regressions, with one (set of) variables excluded in each column.

The first three rows are based on the CPV classification of products, yielding a total of 41 dummy variables. The middle three rows are based on the most fine-grained classification, in which each actual template procurement is represented by a dummy variable, for a total of 209 dummy variables. In the bottom three rows apparently similar or identical products are indicated with a common dummy variable. For example, "media for library" and "library media" are fused into one category, as are "management of nursing homes" and "management of nursing homes; dementia care". This gave us 139 dummies – and this is our preferred specification.

The right-most three columns provide the corresponding partial R^2 values.[22] Table 6 shows that authority effects contribute more to the explanation of the variation in responses, even when the most fine-grained product dummies are used. The tendency that authority effects are important is particularly strong for price and quality uncertainty. Here, the partial R^2 for authority identity is more than twice as large as the corresponding value for the type of product in our preferred specification of product effects. The contribution from adding experience is negligible. We conclude that product characteristics as well as procurer identity influence the choice of supplier-selection method at least indirectly, via the preferences reflected in our survey questions. This conclusion is stable for the alternative definitions of product fixed effects. Indirectly, this supports our third hypothesis.

Finally, the SUR regressions that relate the replies to the three survey questions to each other, as well as to other variables, are reported in the three columns to the left in Table 7. According to the estimates, uncertainty about quality goes hand in hand with uncertainty about the price. High uncertainty about price is also associated with an emphasis on budget control and vice versa. Similarly, high uncertainty about quality is associated with an emphasis on quality and emphasis on quality is associated with uncertainty about quality.

Since there were authorities of various categories in the survey, while the control group solely consisted of local authorities (municipalities), the model was also estimated for a subset that includes only the municipalities. The results are reported in the three columns to the right. The outcome with respect the survey variables are in line with those for the full sample. A few other results are worth commenting. Based on the findings for the full sample, central government authorities seem to care relatively more about quality, to be less uncertain about

---

[22] Calculated as the fraction of the otherwise unexplained variation that can be explained by adding the variable or set of variables indicated at the top of the column to the model.

Table 7. Results, seemingly unrelated regression (SUR). $^*p< 0.10$, $^{**}p< 0.05$, $^{***}p< 0.01$

	All ($N= 1,298$)		Only local governments ($N = 999$)	
	Coef.	S.E	Coef.	S.E
Budget vs. Quality				
Price uncertainty	0.25***	0.03	0.22***	0.03
Quality uncertainty	-0.17***	0.03	-0.13***	0.03
TED = 1	0.04	0.11	0.26**	0.13
Contract period	0.09*	0.05	0.01	0.05
Extension period	-0.05	0.05	-0.07	0.06
Goods contract	-0.08	0.13	0.06	0.14
Gov. authority	-0.56***	0.21		
County	0.17	0.26		
Control group = 1	-0.19	0.19	-0.17	0.18
Experience	-0.30**	0.13	-0.31**	0.12
Constant	4.57***	0.27	4.62***	0.27
Price uncertainty				
Budget vs. Quality	0.17***	0.02	0.18***	0.03
Quality uncertainty	0.80***	0.02	0.78***	0.02
TED = 1	-0.27***	0.09	-0.41***	0.12
Contract period	0.03	0.04	0.07	0.05
Extension period	-0.09**	0.04	-0.13**	0.05
Goods contract	-0.55***	0.11	-0.63***	0.13
Gov. authority	0.55***	0.17		
County	0.26	0.21		
Control group = 1	0.21	0.15	0.20	0.16
Experience	-0.12	0.10	-0.12	0.11
Constant	0.02	0.24	0.10	0.27
Quality uncertainty				
Budget vs. Quality	-0.12***	0.02	-0.12***	0.03
Price uncertainty	0.87***	0.02	0.86***	0.03
TED = 1	0.24**	0.10	0.32***	0.12
Contract period	-0.03	0.04	-0.05	0.05
Extension period	0.09*	0.05	0.12**	0.05
Goods contract	0.52***	0.11	0.55***	0.13
Gov. authority	-0.89***	0.18		
County	-0.50**	0.22		
Control group = 1	-0.40**	0.16	-0.40**	0.17
Experience	-0.18	0.11	-0.20*	0.12
Constant	1.48***	0.24	1.50***	0.28

Equation	Pseudo R^2	chi2	P	PseudoR^2	chi2	P
Budget vs. Quality	0.02	93.64	0.00	0.01	54.71	0.00
Price uncertainty	0.26	1834.89	0.00	0.25	1322.98	0.00
Quality uncertainty	0.27	1802.98	0.00	0.22	1240.31	0.00

quality and more uncertain about prices – compared to local authorities. Low quality uncertainty speaks in favor of using lowest price, while being more uncertain about costs should have the opposite effect. As reported above, we find that central authorities do, in fact, use lowest price more often.

We see that authorities that have prior experience emphasize quality more and costs less, while reporting a lower degree of price and quality uncertainty. That confidence grows with experience is expected, but that experience also makes the buyer prioritize quality more strongly is interesting and perhaps encouraging.

Budget control is reported as being more important if the procurement contract in question stipulated longer contract periods. The length of the contract period is, however, not important when the respondents indicated their uncertainty about price and quality, respectively. The extension period matters, to some extent, for price and quality uncertainty. Uncertainty about price is higher for contracts with short extension periods, while the opposite effect is found for uncertainty about quality. Finally, goods contracts are associated with low price uncertainty but with high quality uncertainty.

6. CONCLUDING DISCUSSION

In this paper we focus on two well-known problems that a contracting authority faces and on how they are related to scoring-rule design. [Har knappt ändrat något här. Vi bör saga att vi har en tredje fråga, nämligen habits redan här.] The first of the two problems is that since important aspects of quality are often non-verifiable, there is an incentive for the supplier to shirk on (ex-post) quality after the contract has been secured. I.e., the contracting authority has to deal with moral hazard – as well as a risk for adverse selection. The second problem is that the contracting authority may face a substantial uncertainty concerning the providers' – and, eventually, its own – cost of meeting alternative (ex-ante) quality levels. If quality standards are set too high, quality may on the margin cost more than it is worth.

We argue that, from a normative point of view, lowest price is an adequate method when the degree of uncertainty is low, for example because the procured products are standardized and since quality can be verified. When there is significant cost uncertainty it is better to use the so-called Economically Most Advantageous Tender Method (EMAT). (And then preferably a method which assigns monetary values to different quality levels.) If there is significant uncertainty concerning delivered quality, the contracting authority should retain a degree of

discretion, so as to be able to award good quality performance in observable but non-measurable quality dimensions; options to extend the contract and subjective assessments of quality are two possibilities. Based on survey data we test the hypothesis that a high level of uncertainty concerning the cost of different levels of quality makes the use of the lowest-price supplier selection method less likely. Using the same data we also test the prior that highly non-verifiable quality makes the contracting authority more inclined to use high quality weights.

Based on procurement and contract characteristics and on our survey data, using a simple empirical framework with logistic, Tobit and SUR regressions, our main findings are that EMAT and more complex scoring rules are used when the contracting authorities report that they experience substantial uncertainty concerning delivered quality and actual costs and that these factors tend to decrease the weight given to price – in line with our predictions.

The findings that product effects are important in explaining survey responses, that extension periods are positively associated with quality uncertainty and that EMAT is used more often for high-value (above-threshold) procurements are also in line with normative theories for public procurement.

In contrast, the finding that authority effects matter more for survey responses than do product effects suggests that institutional habits also play a role. Similarly, the findings that central-government authorities and large authorities tend to favour lowest-price procurements are less easily explained by normative theory and more in line with habitual behavior, as it would seem that these presumably more professional authorities should more easily be able to manage the complexities of EMAT.

To summarize, the choice between lowest price or EMAT as the supplier selection method can, at least to some extent, be explained by procurement and contract characteristics and by the contracting authority's stated uncertainty about price and quality. However, our results also support the conclusion that institutional habits are important determinants of tendering design, in line with the organizational-process model.

REFERENCES

Albano, G. L., C. Bernardino, and I. Alberto, 2011, *Relational Procurement Contracts: A Simple Model of Reputation Mechanism*, CEIS Tor Vergata Research Paper Series, 9(10), No. 209.

Allison, Graham, 1971, *Essence of Decision: Explaining the Cuban Missile Crisis*, Boston: Little, Brown

Allison, Graham and Philip Zelikow, 1999, *Essence of Decision: Explaining the Cuban Missile Crisis*, New York: Addison Wesley Longman.

Arrowsmith, S., 2006, The Past and Future Evolution of EC Procurement Law: from Framework to Common Code?, *Public Contracts Law Journal*, 35, 337-384.

Asker, J., Cantillon, E., 2008. Properties of scoring auctions. *Rand Journal of Economics* 39, 69–85.

Asker, J. and E. Cantillon, 2010, Procurement when Price and Quality Matter, *RAND Journal of Economics*, 41, 1-34.

Banfield E. C., 1975, Corruption as a Feature of Governmental Organization, *Economics of Law and Crime*, 18, 587-605.

Bergman, M. and S. Lundberg, 2013, Tender Evaluation and Award Methodologies in Public Procurement. Forthcoming in *Journal of Purchasing and Supply Management. http://dx.doi.org/10.1016/j.pursup.2013.02.003i.*

Bergman, M. and S. Lundberg, 2012, Sourcing for Government Goods and Services: Theory and Evidence from Swedish Cities and Authorities, in *Charting a Course in Public Procurement Innovation and Knowledge Sharing*, edited by Guan Luigi Albano, Keith F. Snider and Khi V. Thai., 2012, PrAcademic Press, Boca Raton, FL, USA.

Bergman, M. and J. Y. Stake, 2014, *Bid Distribution and Transaction Costs*, manuscript, Södertörns University.

Calzolari, G. and G. Spagnolo, 2009. *Relational Contracts and Competitive Screening,* CEPRDp No. 7434.

Cameron, A. C. and P. K. Trivedi, 2005, *Microeconometrics. Methods and Applications*, Cambridge University Press, Cambridge.

Che, Y-K, 1993, Design Competition through Multidimensional Auctions, *Rand Journal of Economics,* 24, 668-680.

Chen, T. H., 2008, An Economic Approach to Public Procurement, *Journal of Public Procurement,* 8, 407-430.

Coggburn, J. D., 2003, Exploring Differences in the American States' Procurement Practices, *Journal of Public Procurement*, 3, 3–28.

Cripps, M. and N. Ireland, 1994, The Design of Auctions and Tenders with Quality Thresholds: The Symmetric Case, *The Economic Journal*, 104, 316-326.

de Boer, Luitzen, Leo van der Wegen and Jan Telgen, 1998, Outranking Methods in Support of Supplier Selection, *European Journal of Purchasing & Supply Management*, 4 (2‑3), 109‑118.

Dimitri, Nicola, Gustavo Piga and Giancarlo Spagnolo, 2011, *Handbook of Procurement*, Cambridge: Cambridge University Press.

Directive 2004/17/EC of the European Parliament and of the Council of 31 March 2004 coordinating the procurement procedures of entities operating in the water, energy, transport and postal services sectors (30.04.2004)

Directive 2004/18/EC of the European Parliament and of the Council of 31 March 2004 on the coordination of procedures for the award of public works contracts, public supply contracts and public service contracts (30.04.2004).

Green, Mark and Fred Thompson, 2001, Organizational Process Models of Budgeting, *Research in Public Administration*, 7, 55-81.

Hart, O., A. Shleifer, R. W.Vishny, 1997, The Proper Scope of Government: Theory and an Application to Prisons, *Quarterly Journal of Economics*, 112, 1127-1161.

Hyytinen, A., S. Lundberg, S. and O. Toivanen, 2007, *Politics and procurement: Evidence from cleaning contracts*. Discussion Paper No 196, Helsinki Center of Economic Research.

Kelman, S., 1990, *Procurement and Public Management: The Fear of Discretion and the Quality of Public Performance.* Washington: The AEI Press.

Kelman, S., 2002, Remaking Federal Procurement, *Public Contracts Law Journal*, 31, 581-62.

Kelman, S., 2007, Public Administration and Organization Studies, in Arthur Brief and James P. Walsh (Eds) *Academy of Management Annals*. New York: Erlbaum.

Klein, B., and K. B. Leffler, 1981, The Role of Market Forces in Assuring Contractual Performance, *Journal of Political Economy*, 89, 615-641.

Klemperer, Paul, 2004, *Auctions: Theory and Practice*, Princeton: Princeton University Press.

Krishna, Vijay, 2002, *Auction Theory*, Academic Press, Elsevier.

Laffont, J.-J., Tirole, J., 1993. *A Theory of Incentives in Procurement and Regulation*. The MIT Press, Cambridge.

Macaulay, S., 1963, Non-contractual relations in business: A preliminary study, *American Sociological Review*, 28, 55-67.

MacLeod, W. B., 2007, Reputations, Relationships, and Contract Enforcement, *Journal of Economic Literature*, 45, 595-628.

March, James G. and Johan P. Olsen, 1989, *Rediscovering Institutions – the Organizational Basis of Politics*, The Free Press.

Mateus, R., J. A. Ferreira, J. Carreira, 2010, Full Disclosure of Tender Evaluation Models: Background and Application in Portuguese Public Procurement, *Journal of Purchasing & Supply Management*,16, 206–215.

Milgrom, Paul, 2004, *Putting Auction Theory to Work*, Cambridge: Cambridge University Press

Orszag, P. R., 2009, *Improving Government Acquisition*. Office of Management and Budget (M-09-25, July 29, 2009). Available at www.whitehouse.gov.

Potoski , M., 2008, State and Local Government Procurement and the Winter Commission, *Public Administration Review*, 68, 58-69.

Powell, Walter W. and Paul J. DiMaggio (Eds.), 1991, *The New Institutionalism in Organizational Analysis*, Chicago: The University of Chicago Press.

Sánchez Graells, Albert, 2013, *The Difficult Balance between Transparency and Competition in Public Procurement: Some Recent Trends in the Case Law of the European Courts and a Look at the New Directives*, University of Leicester School of Law Research Paper No. 13-11.

Schooner S. L., 2011, A Random Walk: The Federal Circuit 2010 Government Contracts Decisions, *American University Law Review*, 60, 1067-2011.

Simon, Herbert A. 1947. *Administrative Behavior*. New York: Macmillan.

Spagnolo, G., 2012, Reputation, Competition, and Entry in Procurement, *International Journal of Industrial Organization*, 30, 291‑296.

Strand, Ivar, Paula Ramada and Erik Canton, 2011, *Public Procurement in Europe. Cost and Effectiveness*. Report to the EU Commission.

Tadelis, S., 2012, Public Procurement Design: Lessons from the Private Sector, *International Journal of Industrial Organization*, 30, 297-302.

Telgen, J. and F. Schotanus, 2010, *The Effects of Full Transparency in Supplier Selection on Subjectivity and Bid Quality*, Proceedings of the 4th International Public Procurement Conference, Seoul.

Thompson, F. J., 2008, State and Local Governance Fifteen Years Later: Enduring and New Challenges, *Public Administration Review*, 68, 8 - 19.

Verdeaux, J.-J., 2003, Public Procurement in the European Union and the United States: A Comparative Study, *Public Contract Law Journal*, 32, 713-738.

Weitzman, M. L., 1974, Prices vs. Quantities, *Review of Economic Studies*, 41, 477-91.

Williamson, Oliver E., 1981, The Economics of Organization: The Transaction Cost Approach, *The American Journal of Sociology*, 87(3), 548-577

APPENDIX1

Table A1. Product categories, based on the CPV-codes.

Category	Freq.	Percent		Freq.	Percent
CPV classification missing	7	1.08	Construction structures and materials; auxiliary products to construction (except electric apparatus)	10	1.54
Agricultural, farming, fishing, forestry …	1	0.15	Construction work	72	11.06
Petroleum products, fuel, electricity …	6	0.92	Software package and information systems	6	0.92
Mining, basic metals and related products	2	0.31	Repair and maintenance services	13	2.00
Food, beverages, tobacco and related products	12	1.84	Hotel, restaurant and retail trade services	12	1.84
Agricultural machinery	3	0.46	Transport services (excl. Waste transport)	21	3.23
Clothing, footwear, luggage articles and accessories	10	1.54	Supporting and auxiliary transport services …	5	0.77
Leather and textile fabrics, plastic, rubber materials	3	0.46	Postal and telecommunications services	2	0.31
Printed matter and related products	15	2.30	Financial and insurance services	10	1.54
Chemical products	3	0.46	Real estate services	4	0.61
Office and computing machinery, equipment …	15	2.30	Architectural, construction, engineering and inspection …	40	6.14
Electrical machinery, apparatus, equipment …	5	0.77	IT services: consulting, software development…	20	3.07
Radio, television, communication, telecommunication …	12	1.84	Research and development services …	4	0.61
Medical equipments, pharmaceuticals products	58	8.91	Administration, defense and social security services	5	0.77
Transport equipment and auxiliary products to transportation	20	3.07	Agricultural, forestry, horticultural, aquacultural and apicultural services	9	1.38
Security, fire-fighting, police and defense equipment	12	1.84	Business services: law, marketing, consulting, recruitment, printing and security	75	11.52
Musical instruments, sport goods, games, toys, handicraft, art materials and accessories	1	0.15	Education and training services	9	1.38
Laboratory, optical and precision equipments…	11	1.69	Health and social work services	56	8.60
Furniture (incl. office furniture), furnishings, domestic appliances (excl. lighting) and cleaning products	18	2.76	Sewage, refuse, cleaning and environmental services	31	4.76
Industrial machinery	9	1.38	Recreational, cultural and sporting services	13	2.00
Machinery for mining, quarrying, construction equipment	1	0.15	Other community, social and personal services	10	1.54

Table A2. R2-values from regression with authority (municipal) and product fixed effects.

Controls	Dependent		
	Quality vs. Budget $N = 1017$	Price uncertainty $N = 1017$	Quality uncertainty $N = 1005$
1. Municipality	0.3970	0.3847	0.3913
2. Product (CPV)	0.1212	0.0189	0.0402
3. Experience	0.0065	0.0542	0.0420
1 and 2	0.5057	0.4251	0.4257
1 and 3	0.4000	0.4203	0.4138
2 and 3	0.1283	0.1027	0.0776
1, 2 and 3	0.5085	0.4560	0.4448
Coef. Experience	negative and sign.	negative and sign.	negative and sign.
1. Municipality	0.3970	0.3847	0.3913
2. Product name (CFT 1)	0.2860	0.2529	0.2130
3. Experience	0.0065	0.0542	0.0420
1 and 2	0.6254	0.5380	0.5389
1 and 3	0.4000	0.4203	0.4138
2 and 3	0.2905	0.2747	0.2313
1, 2 and experience	0.6263	0.5557	0.5513
Coef. Experience	negative and sign.	negative and sign.	negative and sign.
1. Municipality	0.3970	0.3847	0.3913
2. Product name (CFT 2)	0.2152	0.1873	0.1648
3. Experience	0.0065	0.0542	0.0420
1 and 2	0.5825	0.5102	0.5064
1 and 3	0.4000	0.4203	0.4138
2 and 3	0.2236	0.2173	0.1848
1, 2 and experience	0.5859	0.5311	0.5203
Coef. Experience	negative and sign.	negative and sign.	negative and sign.

Note that Product fixed effects are first based on the CPV code (CPV), alternatively on the label in the call for tender (CFT 1) and finally on the label in the call for tender but if different descriptions are used for the same product these are merged into one category (CFT 2).

APPENDIX 2 – SURVEY TO THE CONTROL GROUP

Introduction

In this survey we will ask you questions concerning the characteristics of some products. There will be three questions for each product and the questions will concern quality. In this survey we interpret quality to include all characteristics of the product except the price: quality in a traditional sense, including, e.g., durability and energy efficiency, but also delivery times, product range of a supplier, attitudes towards the customer, ethical aspects, knowledgeable staff and so on.

The questions are addressed to the manager in charge of the municipality's procurement or to a staff member that works at the central procurement office. We seek the spontaneous assessment of an experienced and senior staff member, rather than the opinion of personnel working directly with the procured product or the municipality's official policy. We are primarily interested in the differences between the products, not in determining the levels in an absolute sense. Some questions will concern quite broad product areas, where characteristics may vary; if so, please respond for typical or, in terms of value, dominant products within these areas.

Please note that the names of the products are provided at the top of the pages.

[Name of product/area, e.g., construction of a new day-care building]

Has the municipality procured this product? Yes_____ No _____

Comment: _____

Question 1. For this product, how important is it i) that the delivered quality is exactly the desired quality, relative to ii) that the cost is close to the allocated budget? Please note that we are not asking for the relative weights of price and quality used to evaluate the bids.

How important is it to get the desired quality relative to getting a cost that is close to the allocated budget?

Quality is critical, price is irrelevant				Euqally important				Price is critical, quality is irrelevant
○	○	○	○	○	○	○	○	○

Comment: _____

Question 2. For this product, how knowledgeable is the procurement unit and its staff concerning the costs of different quality levels and concerning the product as such? Assume a situation where the authority has made normal preparations prior to a procurement, rather than the current level of knowledge concerning this product. Assume a normal business-cycle situation.

How knowledgeable is the procurement unit and its staff concerning the costs of different quality levels and concerning the product as such?

No uncertainty about cost (price)				Intermediate				Extremely large uncertainty about cost (price)
○	○	○	○	○	○	○	○	○

Comment: _____

Question 3. For this product, to which extent can the quality for the product be described in the call for tender and in the contract and to what extent can quality be verified upon or after delivery, so that the quality level will be exactly as intended? Our question is not about actual occurrences of quality deficiencies and not about whether the authority has developed routines for post-delivery control; our question is whether or not it would, in principle, be possible to objectively specify (in a measurable way) and verify quality so that it would be feasible to sue the supplier for breach of contract if delivered quality is deficient.

To what extent can the quality for the product be described in the call for tender and in the contract and to what extent can quality be verified upon or after delivery, so that the quality level will be exactly as intended?

No uncertainty about actually delivered quality				Intermediate				Extremely large uncertainty about actually delivered quality
○	○	○	○	○	○	○	○	○

Comment: _____