“The Pivotal nature of award methods in green public procurement”

AUTHORS	Sofia Lundberg Per-Olov Marklund
ARTICLE INFO	Sofia Lundberg and Per-Olov Marklund (2011). The Pivotal nature of award methods in green public procurement. *Environmental Economics*, *2*(3)
RELEASED ON	Tuesday, 08 November 2011
JOURNAL	"Environmental Economics"
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
0	0	0

Environmental Economics, Volume 2, Issue 3, 2011

Sofia Lundberg (Sweden), Per-Olov Marklund (Sweden)

The pivotal nature of award methods in green public procurement

Abstract

Internationally there is a strong trend of implementing Green Public Procurement (GPP), and it is seen as an environmental policy tool. By its purchasing power public authorities are via environmental concerns in procurement believed to have the power of stimulating firms to reduce emissions, be resource efficient, and developing products proper in line with a sustainable society. In allocating public contracts authorities use award methods and scoring rules. This paper discusses the procedure of allocating contracts when GPP is applied. Departing from previous research on this topic the paper explicitly discusses the pivotal role of using suitable award methods and scoring rules for GPP to functioning as an efficient environmental policy tool. It is most important that GPP matches the preferences of the society (e.g., a municipality or a state authority). Assuming that GPP can be used as an environmental policy tool, the authors present general guidelines for choosing an efficient award method and scoring rule in perspective of welfare and sustainability. The necessity of such guidelines cannot be emphasized enough, partly because previous scientific literature on the topic is very scarce and partly, which is most serious, empirical data (Swedish public procurement data) indicate that scoring rules that violates necessary conditions for efficient outcomes are commonly used. It cannot be ruled out that this unfortunate circumstance causes the society substantial costs. In this perspective it is also noteworthy that Sweden is regarded as a frontrunner in successfully implementing GPP (Kahlenborn et al., 2011).

Keywords: award methods, economically most advantageous tender, environmental policy, lowest price, public procurement auctions, scoring rules, sustainability.
JEL Classification: H57, Q01, Q58.

Introduction

"… the criteria on which the contracting authorities shall base the award of public contracts shall be either: when the award is made to the tender, most economically advantageous from the point of view of the contracting authority, various criteria linked to the subject-matter of the public contract are in question, for example, quality, price, technical merit, aesthetic and functional characteristics, environmental characteristics, running costs, cost-effectiveness, after-sales service and technical assistance, delivery date and delivery period or period of completion, or the lowest price only. ... the contracting authority shall specify … the relative weighting which it gives to each of the criteria chosen to determine the most economically advantageous tender"

(Article 55, 2004/17/EC, Article 53, 2004/18/EC).

The overall purpose of this paper is to discuss a crucial issue connected to the procedure of allocating public contracts when so-called Green Public Procurement (GPP) is applied, i.e., to illustrate the importance of using an efficient *award method* when environmental characteristics are considered in the evaluation of sellers' bids.

Public procurement corresponds approximately to 16 percent of EU's total GDP (European Commission, 2008), which clearly is a significant part of EU's economy. Public authorities are, therefore, considered as having a considerable purchasing power that can be used to stimulate firms to reduce emissions and produce more resource efficiently. Also, it is believed that firms can be stimulated to develop goods and services that better fit the ambitions of public authorities being part of a sustainable society. In that way it is also believed that the authorities' own consumption will leave less ecological footprints if GPP is implemented. Internationally there is a strong trend of promoting the implementation of GPP. For instance, The European Commission is very clear in its ambitions, European Commission (2008; 2010), and the Green Paper (2011). A first step was taken in 2001 at the Gothenburg European Council (2001) and Member States were called for developing National Action Plans (NAPs) on GPP, European Commission (2003). Most states have also acted on this call. From the political point of view GPP is thereby regarded as an environmental policy tool.

For GPP to work as an efficient environmental policy tool the authorities need award methods and scoring rules to account for both price and environmental characteristics. Note that the EU directives regulate the call for tender to be evaluated in terms of the lowest price or the economically most advantageous tender (EMAT). Although the directives implicitly stipulate the use of a scoring rule a lot of freedom is leaved to the procuring authorities in how to design the specific scoring rule to be used. As a matter of fact, from a welfare economics and sustainability point of view, scoring rules are often poorly designed in practice (Bergman and Lundberg, 2011). Some of the most frequently used scoring rules can lead to an arbitrary award of contracts and, as shown in the present paper, in terms of GPP, to an inefficient environmental policy. It is in this perspective worrying that contracts, e.g., in Sweden in about one third of all procurement auctions, are allocated based on scoring rules of such nature. In light of this, it is also most striking that Sweden is identified as a frontrunner in successfully implementing GPP (Kahlenborn et al., 2011). As illustrated in this paper when evaluating GPP as an environmental policy tool and identifying frontrunners one cannot only rest on the fact that GPP is applied, one must also understand how it is applied.

The objective of this paper is specifically to discuss award methods and scoring rules in relation to GPP and, as we will assert, the choice of award method and scoring rule is of crucial importance in the context of environmental policy. Specifically, throughout this paper public procurement is here regarded as defined by the EU procuring directives, 2004/17/EC and 2004/18/EC: "*Public procurement means the measures implemented by a contracting authority with the aim of awarding a contract or concluding a framework agreement regarding products, services, or works*" (Article 13)[1]. By departing from the political objective within the EU of achieving sustainable development, we address the topic on a general level by putting GPP and the choice of award method and scoring rule within the perspective of welfare economics.

Strikingly, economics research is lagging considerably behind the political activities to implement and establish GPP. The literature in the field of economics addressing GPP is very scarce and, hence, many of the commonly used political arguments in favor of GPP are hardly supported by scientific findings (Lundberg et al., 2009).

Furthermore, most of the previous studies take GPP for granted and focus mainly on how it should be implemented to achieve a better environment, not considering GPP in a wider context, e.g., including the important question how to regard GPP in comparison to other environmental policy tools (see, e.g., Michelsen and de Boer, 2009; Nissinen et al., 2009; Geng and Doberstein, 2008; Parikka-Alhola, 2008; Thomson and Jackson, 2007; D'Amoto, 2006; Kippo-Edlund et al., 2005; Grolleau et al., 2004; Erdmenger, 2003; Clinch et al., 2002; Sterner, 2002; Marron, 1997). This paucity was more closely emphasized in Lundberg et al. (2009). However, most importantly, neither of these studies recognizes the crucial role of choosing the appropriate awarding method and scoring rule.

Scoring rules within public procurement are discussed by, e.g., Asker and Cantillon (2008; 2010) and Dini et al. (2006). However, neither of these studies explicitly addresses scoring rules in context of GPP. Therefore, the major contribution of this paper is to bring structure to this issue on a general level, by drawing mainly from Bergman and Lundberg (2011). Their previous research is here translated to the specific situation of GPP, and as such we can conclude something about when certain types of award methods and scoring rules should be used.

This paper is organized as follows. The institutional settings and principles of GPP are presented in section 1. Section 2 includes a simple theoretical presentation of how to get the best environmental value for money. Under the assumption of a utility maximizing procuring authority and no uncertainty about production costs the socially optimal trade-off between price and environmental quality is illustrated. Award methods and scoring rules are discussed in section 3. Uncertainty about production costs is introduced in section 4. GPP in practice and environmental policy consequences of badly designed scoring rules is discussed in section 5, followed by a short presentation of empirical findings in section 6. The final section concludes the paper.

1. Institutional settings

Following the GPA[2] and the European procurement directives[3], public contracts within the EU are allocated by competitive bidding. In general, the procurement directives stipulate sealed bidding and contracts are either awarded to the lowest bidding supplier or the supplier who is considered to have submitted the economically most advantageous tender (EMAT). Irrespective of the award principle mandatory requirements can be used in a first qualifying phase followed by a second phase where the bids are evaluated. The evaluation is either based on a combination of price and quality characteristics, price only, or quality only. When contracts are allocated based on a multiple set of criteria[4] these should be linked to the subject-matter of the public contract in question. In addition to price, e.g., quality, technical merit, aesthetic and functional characteristics, environmental characteristics, running costs, cost-effectiveness, after-sales service and technical assistance, delivery date and delivery period, or period of completion can be considered in the award

[1] This definition excludes auctions of tradable permits and nature conservation contracts.

[2] The Government Procurement Agreement (GPA) is a plurilateral agreement negotiated alongside multilateral trade agreements within the WTO. The European Community (27 member states), Hong Kong (China), Iceland, Israel, Japan, Korea, Liechtenstein, the Netherlands with respect to) Aruba, Norway, Singapore, Switzerland and the United States of America has to date signed the GPA.

[3] Directives 2004/17/EC (water, energy, transport and postal service sectors) and 2004/18/EC (public work contracts, public supply contracts and public service contracts).

[4] See, e.g., Che (1993) for more on multidimensional auctions.

decision. For simplicity reasons here we disregard other qualitative criteria but environmental, consequently, focus is only on the trade-off between price and environmental criteria (GPP). Furthermore, it is assumed that bids are either two dimensional, that is, potential suppliers submit a price scalar and an environmental quality scalar, or one dimensional[1]. In the latter case, bidders either submit only a price bid or only an environmental quality bid. The environmental quality could in itself of course been seen as multidimensional but it would only complicate the analysis and the main findings are not altered.

When contracts are awarded according to EMAT and contracts are allocated based on a combination of price and environmental characteristics some form of scoring rule must be applied. The procurement directives stipulate relative weighing of price and criteria and if this is impossible, that criteria (price included) are listed in descending order of priority[2]. The directives are in its formulation of the exact scoring rule vague giving a lot of freedom in its design to the procuring authority. Equal treatment, transparency, non-discrimination, proportionality, and mutual recognition are the guiding principles for public procurement within the EU and consequently for the scoring rule[3].

2. Best environmental value for money

The procuring authority represents the society, e.g., a local government or a state authority. Having decided to procure a certain product and to use the procurement auction as an environmental policy tool, the procuring authority faces the question of how to explicitly design the procurement auction and how to consider the environment in relation to price, all things equal. For one thing, according to the EU procuring directives, the procuring authority must clearly declare in the call for tender how bids will be evaluated, i.e., specify an award method. If contracts will be awarded on a combination of price and environmental quality a scoring rule must also be specified. The following discussion in this section is based on Bergman and Lundberg (2011).

From a welfare economics point of view, the scoring rule should truthfully represent the society's preferences for the procured product and the environment. Theoretically, the preferences are expressed by the utility function, which commonly is assumed to be quasi-linear in bids in purpose of abstracting away from the income effect[4]. Assuming no income effect the objective function in the society's utility maximization problem may be expressed as[5]:

$$\max_{Q} U(Q,C) = V(Q) + B - C(Q) . \quad (1)$$

where $V(Q)$ is the procurer's valuation of the environment (e.g., Q representing a certain environmental quality expressed as a mandatory requirement or evaluation criteria in the call for tender), B is the procuring authority's budget, and $C(Q)$ is the bid, or the seller's cost of producing the product. Furthermore, it is assumed that $C'(Q) > 0$, $C''(Q) > 0$, $V'(Q) > 0$, and $V''(Q) < 0$.

Besides exhibiting diminishing marginal utility in environmental quality, the utility function exhibits some basic properties, such as completeness, transitivity, monotonicity, and independence of irrelevant alternative bids. If we departure from full information and think of low environmental quality as associated with low production costs, $C'(Q) > 0$, the preferences for production costs and environmental quality can for a representative procuring authority be illustrated as in Figure 1[6]. As in Bergman and Lundberg (2011), the situation is for simplicity reasons illustrated for only three firms which are assumed to be heterogeneous in costs, that is, given Q_3^* ; $C_3 < C_2 < C_1$.

It is shown in Figure 1 that a utility maximizing procuring authority, as expressed in equation (1), indeed gets the highest environmental utility for the money. The contract will, given the assumptions made and that an efficient award method is chosen, be allocated to the supplier with the least environmental (marginal) adjustment cost, favoring the supplier with the environmentally cleanest production technology ex ante the procurement auction. That is, firm 3 is assigned the contract, the price paid is P_3^*, and the environmental quality achieved is Q_3^* [7]. The interesting question is then, how to allocate contracts in order to reach the $P_3^* Q_3^*$ outcome.

[1] An alternative is of course to let bidders submit price and quality vectors in line with so-called menu auctions (e.g., Bernheim and Whinston, 1986; Klemperer, 1999; Milgrom, 2004). However, Asker and Cantillon (2008) demonstrate that allocation of contracts based on scoring rules dominate beauty contests, price-only auctions and menu auctions.

[2] See Article 55 in 2004/17/EC or Article 53 in 2004/18/EC.

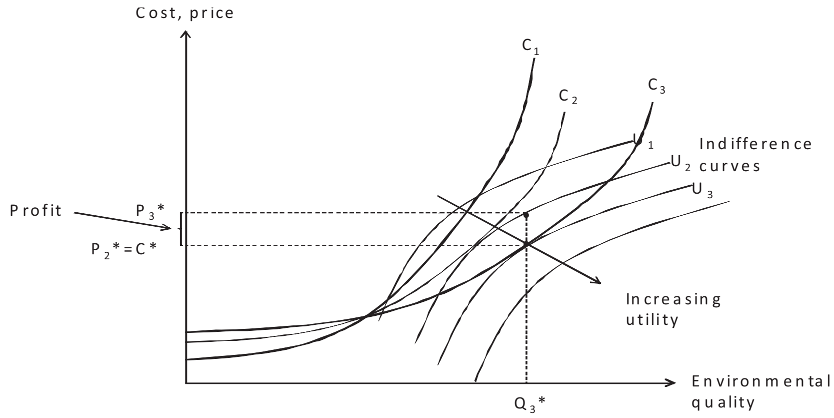
[3] Directives 2004/17/EC (water, energy, transport and postal service sectors) and 2004/18/EC (public work contracts, public supply contracts and public service contracts).

[4] The assumption of no income effects is reasonable in most procurement auctions. Commonly, the price of the procured product constitutes only a small fraction of the procuring authority's total budget.

[5] To keep the analysis simple we assume that among sellers the procured product is homogenous in other qualities than the environmental quality. Assuming the environmental dimension being the only quality dimension will not alter the outcome of the analysis.

[6] Note that the cost curves show the cost of producing a fixed quantity of the procured product, i.e., the intercepts, plus the cost of adjusting to environmental criteria. Consequently, the cost curves only vary with environmental criteria stringency.

[7] However, this is not to be interpreted as GPP necessarily being a cost-effective way of improving the environment (see Lundberg et al., 2009).



Source: Bergman and Lundberg (2011).

Fig. 1. Utility maximization given heterogeneous bidders

3. Award methods and scoring rules

Given that GPP is applied, there are four main categories of methods that public procurers can use to allocate contracts. Winning bids can be selected on the basis of: (1) Price only; (2) Environmental quality only; (3) Price-to-environmental quality scoring; and (4) Environmental quality-to-price scoring. All but price only belong to the rule of awarding the economically most advantageous tender (EMAT). Price only takes the form of traditional first price sealed bid auction which can be combined with mandatory quality criteria, e.g., environmental ones. In this case, bids are one-dimensional, and price only bids are submitted. Environmental quality only builds on the principle that the procuring authority specifies a price and then potential suppliers compete in the quality dimension only. Methods (3) and (4) require that a *scoring rule* is specified. As pointed out in Bergman and Lundberg (2011), scoring rules related to these two types of methods can be designed to be equivalent in outcome. However, yet there are convincing reasons to choose an environmental quality-to-price scoring rule.

The four main categories of methods and different scoring rules varies in their ability to give an efficient outcome, e.g., as manifested by the bundle (P_3^*, Q_3^*) in Figure 1. An efficient scoring rule is linear in the bid (Dini et al., 2006; Asker and Cantillon, 2008; 2010), transparent, consistent, and difficult to manipulate for strategic reasons (Bergman and Lundberg, 2011). A choice of an efficient scoring rule that considers both price and environmental aspects is a condition for a welfare efficient outcome, internalizing environmental externalities optimally. The four main categories of award methods will be presented in short below.

3.1. Price only. Let us start with a situation where bidders compete in price only. In this case the procuring authority specifies mandatory environmental requirements in the call for tender corresponding to Q_3^* in Figure 1. Then, bidders submit sealed bids that are pure price bids and the contract is allocated to the lowest bidding firm, which given the production cost curves, will be firm 3[1]. The price paid is P_3^*. This is equivalent to the classical sealed bid first price auction (e.g., Vickrey, 1961; Milgrom, 1989).

3.2. Environmental quality only. Another alternative is to let the bidders compete in environmental quality only. First the price is by the procuring authority set to P_3^* in Figure 1. Then, bidders submit sealed bids that are pure quality bids and given the cost curves, the contract is allocated to the bidder offering the highest environmental quality, which again will be firm 3. The environmental quality achieved is Q_3^*. In this case, the challenge is to design GPP so the price truthfully corresponds to the society's utility from consuming the procured product and the utility from environmental improvement (internalizing negative environmental externalities).

Given full information, price or environmental quality only competition will give the same socially optimal outcome ($P_3^* Q_3^*$).

3.3. Environmental quality-to-price scoring. Assuming no income effects, the utility function in equation (1) is specified as being quasi-linear in price. For a correct representation of the society's preferences, this means that scoring rules must also be linear in price. As shown in Bergman and Lundberg (2011), in the case of an environmental quality-to-prices scoring rule, and perfect competition, the utility function may be represented by the evaluation price, *EP*, generally as:

[1] Here, it is assumed that all firms placing a bid meet the mandatory requirements.

$$EP_i = C_i(Q_i) - V(Q_i), \text{ where } i = 1,\ldots,n \text{ bidders,} \quad (2)$$

where supplier i's production cost equals its submitted offer, $C_i(Q_i) = P_i$, and $V(Q_i)$, is the society's utility of supplier i delivering the environmental quality, Q_i, expressed in monetary units. Hence, the submitted offer, P_i is discounted with the environmental quality value (the environmental shadow price) and, accordingly, the supplier that submits the lowest evaluated offer, EP, will be allocated the contract (which however still will be paid according to $P_i = C_i(Q_i)$. Referring back to the only three supplier situation illustrated in Figure 1, the evaluation price will be $EP_3 = P_3^* - V(Q_3^*)$, $P_3^* > C_3(Q_3^*)$. The contract is allocated to firm 3, offering the highest utility according to equation (1).

Furthermore, accounting for environmental quality in equation (2) can be done by either a rewarding discount, i.e., suppliers offering quality in excess of a minimum quality required by the procurer are rewarded or, by a surcharge, i.e., offered quality in insufficiency of a maximum quality is punished. Also, the quality discount or surcharge can be defined in absolute or relative terms, which means they are either the same for all suppliers or that they are proportional to the supplier's own bid, respectively. However, as pointed out in Bergman and Lundberg (2011), an absolute approach is preferred. By using a relative approach it is assumed that the value of one unit quality varies with the submitted offer, which makes it difficult to get a perfect match between the society's ex ante value of environmental improvement and its ex post value.

An alternative to use money as the unit of measurement, as in the case of environmental quality-to-price scoring rules, one can use quality points as the unit of measurement by applying price-to-environmental quality scoring rules. The latter rules are discussed in the next section.

3.4. Price-to-environmental quality scoring. Again, starting from the assumption of no income effects, the utility is quasi-linear in price. From a purely linear algebraic point of view, a functional form that satisfies this assumption could be expressed generally as[1]

$$S_i = a + b[C_i(Q_i) - V(Q_i)], \text{ where } i = 1,\ldots,n \text{ bidders,} \quad (3)$$

where S_i denotes the evaluated score of a bid from supplier i. Consequently, price-to-environmental quality scoring rules, where the bid, $P_i = C_i(Q_i)$, is transformed from a monetary value to a scoring point, must satisfy the expression in (3).

The expression in (3) is basically the same as in equation (2) as a, the intercept, and b, the slope coefficient are constants. Therefore, assuming full information, in theory it does not matter whether an environmental quality-to-price or a price-to-environmental quality scoring rule is used. The evaluation and ranking of submitted bids will generate exactly the same outcome, i.e., referring to Figure 1, firm 3 is allocated the contract. However, it is in practice easier to transform quality scores into money than transforming money to scores, because we are probably more familiar with making decisions based on the price and quality trade-off in monetary terms. Note also that utility, $V(Q)$, in general is not linear and therefore the evaluation of suppliers' environmental performance by using scores will not be scale independent. For instance, following Bergman and Lundberg (2011), a given environmental quality increase will be evaluated differently depending on whether using a scale that measures this increase from 100 to 101 or using a scale measuring the increase from, e.g., 0 to 1. The environment will be valued higher in the latter case. Environmental quality-to-price scoring is therefore recommended.

In practice though, the challenge is to choose an award method that optimizes utility of the society, i.e., a local government, when there is uncertainty about production costs.

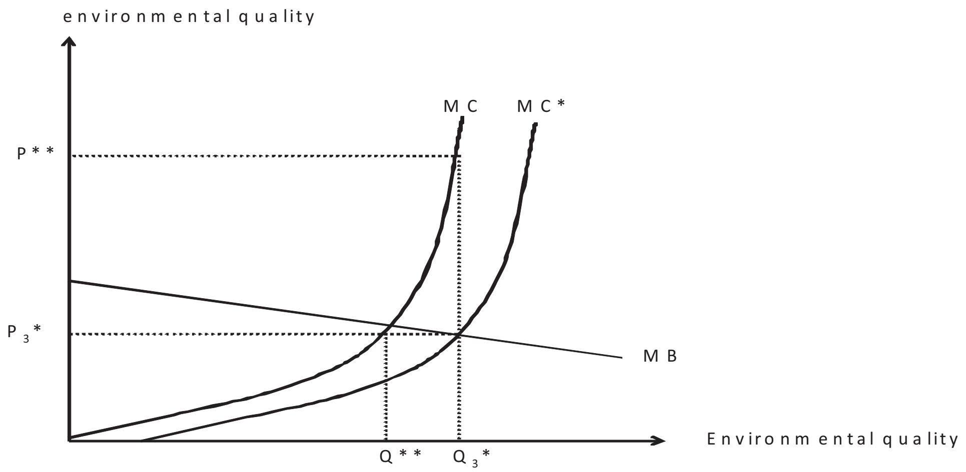
4. Introducing uncertainty

So far we have discussed award methods and scoring rules under the assumption of full information, i.e., the procurer knows the bidders' production costs. In this case, the four alternatives of award methods discussed above will be equivalent, generating the same outcome. However, as shown in Bergman and Lundberg (2011), if the procurer is uncertain about the cost of adjusting environmentally the outcome can divert considerably between the alternatives. For instance, in the case of underestimating the cost, by letting suppliers compete in price only the procurer may pay too high a price of environmental quality if misjudging the suppliers' true production costs in the first place (setting the environmental criteria to stringent in relation to the true production cost). Or, similarly, in the case of competition in environmental quality only the environmental impact of the procurement auction may be insignificant (setting the price too low in relation to the true production cost). Misjudging the true cost, a result lying somewhere in between will be achieved when letting suppliers com-

[1] Obviously, the environmental quality-to-price scoring rule in (2) is a special case of the more general expression in (3), setting $a = 0$, and $b = 1$.

pete in both price and environmental quality, i.e., when using price-to-environmental quality or environmental quality-to-price scoring rules to evaluate bids.

To address uncertainty, specifically in the case of underestimating environmental adjustment costs, as an example, the choice between price-only competition and quality-only competition is illustrated in Figure 2. The *MC** curve reflects the marginal cost curve that the procuring authority assumes is true. If this was the case the socially optimal outcome ($P_3^* Q_3^*$) would be achieved with both methods. The procuring authority is here assumed to know the society's value of the environmental improvement, the marginal utility curve, *MB*. If the true marginal cost is *MC* quality-only competition, at a price P_3^*, will result in a well fare loss since the environmental quality will be lower (Q^{**}) than the socially optimal one (Q_3^*). Price-only competition at a given environmental quality, Q_3^*, will result in a price (P^{**}) much higher than the socially optimal one. Referring back to Weitzman (1974), under uncertainty about production costs, environmental quality-only competition to a fixed price, P_3^*, is generally preferable to price-only competition to a given environmental quality, Q_3^*, if the slope of the bidder's marginal cost curve is steeper compared to the curve reflecting the procurer's marginal utility of the environmental quality[1]. If the relationship is reversed, the *MB* curve being steeper than the *MC* curve, price-only competition is preferable to environmental quality-only competition[2]. That is, if production costs rises steeply relative to the decrease in environmental utility the preferable award method is quality-only competition. However, if the environmental utility falls sharply compared to the rise of production cost, price-only competition is preferable as an award method. Additionally, if both changes in marginal cost and marginal utility are large a proper scoring auction, preferable using an environmental quality-to-price scoring rule, should be applied (Bergman and Lundberg, 2011).



Source: Bergman and Lundberg (2011).

Fig. 2. The choice between price-only and quality-only competition

5. GPP in practice

By representing the utility function in (1) with a scoring rule means that it, besides being linear in bids, also needs to satisfy the properties of completeness, transitivity, monotonicity, and independence of irrelevant alternatives (bids). The three former ones are hardly a problem in practice (Bergman and Lundberg, 2009). Hence, as a basis of judging scoring rules are mainly the desirable properties of being linear in bids and independent of irrelevant alternatives.

As discussed previously the property of being linear in bids follows from the reasonable assumption of no income effects and is a necessary condition for the scoring rule to represent the utility function (Asker and Cantillon, 2008). A scoring rule being independent of irrelevant alternatives means that if bid A is preferred to bid B when the irrelevant bid C (e.g., an unreasonable low/high bid) is not needed to be considered, bid A is still preferred to bid B if the irrelevant bid C also has to be considered. The particular problem of violating this property was first recognized in Lunander and Andersson (2004), and further highlighted in Lunander (2009), and Bergman and Lundberg (2009, 2011). Generally speaking, violating the property of independence of irrelevant alternatives is due to the procurer's ranking of bids being dependent on an endogenously given reference bid, or bids.

[1] Referring to Lundberg et al. (2009) this could be associated to GPP working as an economic environmental policy tool (compare fixing price to establishing an environmental tax rate). However, they do not explicitly consider environmental quality-only competition.

[2] Again, referring to Lundberg et al. (2009), in this case GPP could be categorized as a quantitative environmental policy tool (fixing quality).

Bergman and Lundberg (2009, 2011) found, based on Swedish public procurement data[1], that scoring rules often used in practice most unfortunately violates the main properties discussed above. For example, in about a third of the procurements studied the following price-to-quality scoring rule was applied:

$$PP_i = 5 \times \frac{P_{lowest}}{P_i} \text{ where } i = 1,\ldots,n \text{ bidders} \quad (4)$$

which aims at transforming the $C_i(Q_i) = P_i$ value in equation (3) into scoring points. The price scoring, PP_i, is the price score that relates the price bid, P_i, to the lowest submitted bid, P_{lowest}. Obviously, in this case the multiplicative property of the scoring rule entails no-linearity in price P_i and, therefore, the rule cannot represent the quasi-linear utility function in (1). Hence, the no-income effect property is violated. Furthermore, the reference price, P_{lowest} is endogenously given and, consequently, the evaluated bids will not be independent of irrelevant bids. Variants of the rule in equation (4) are scorings rules that, e.g., departure from mean of submitted bids, median of submitted bids, or the difference between highest and lowest submitted bid. The use of endogenous reference prices can open up for strategic manipulation.

Inherent in the problem of endogenous reference prices, as is the case of model (4), is that if a scoring rule that makes the evaluation of a bid contingent on another submitted bid, the model from a potential supplier's perspective is not transparent. For example, given a specific bid, P_i, if a potential supplier choses between producing the subject-matter with a low or a high environmental quality level, she does not know whether the low or high quality alternative will give her the largest probability to win the contract. That is, the supplier is left with an arbitrary choice of environmental quality, as the procurer's choice of winning bid will have an arbitrary outcome. Consequently, the scoring rule is inappropriate from an environmental policy perspective. There is a high probability that the outcome will be based on the environment being incorrectly valued, as the scoring rule hardly will reflect the preferences of the society correctly. Furthermore, non-transparency may deter entry from potential suppliers since non-transparency increases their risk and eventually having negative effect on their expected pay-offs.

Finally, since bids are two dimensional (see equation (1) or (3)), including the environmental quality also, some quality score (QP_i) is added to the price score (PP_i) and the bidder receiving the highest total score (S_i) is awarded the contract.

Obviously, a scoring rule that is non-linear in bids and/or dependent of irrelevant bids, such as in (4), suffers from considerable inefficiency from the economics point of view, accounting both the procurer-seller and the society point of view. Scoring rules involving GPP that risk an arbitrary choice of winner are associated with the risk of an inefficient environmental policy. The outcome could be that the true environmental externality is not fully internalized as well as more than internalized.

5.1. Why weights should not be used. The choice of a scoring rule that satisfies the properties of the utility function in (1) should, e.g., correspond to the expression in (2), $EP_i = P_i - V(Q_i)$, where both the subject-matter and environmental quality are valued in monetary terms (and not in scoring points). The bidder with the lowest evaluation price, EP_i, then wins the contract. However, in practice the environment and the price are also often given different weights. This is fully understandable given the formulation in the EC directives (see introductory quotation) which, however, is an unfortunate formulation. The actual relative weight given to the subject-matter and environmental quality is namely scale dependent, having unfortunate consequences on the evaluation process and, therefore, the possibilities of GPP working as an efficient environmental policy tool.

Let us illustrate this with an example. Consider a procuring authority that designs its procurement to target a specific environmental problem. The authority decides to base the evaluation process on a quality-to-price scoring rule according to equation (2), which has the prerequisites of satisfying the properties of the utility function in (1). However, as is commonly done, in order to internalize the negative environmental externality from producing and consuming the subject-matter, the authority also decides to use the price weight w and the environmental weight (1-w). Assume further, as in Bergman and Lundberg (2011), that every additional unit of environmental quality, Q_i, are valued to v and that quality is ranged from 0 to some value x. The evaluation price is then:

$$EP_i = wP_i + (1-w)(x - Q_i)v \,. \quad (6)$$

As illustrated in Bergman and Lundberg (2011) the choice of range of measurement of the environmental quality will affect the final actual weight of the same. Doubling the range will reduce the environmental quality score by half. From an environmental policy view this means that the ex post relative weight between the procured product and environmental quality

[1] The data originates from the period of 2002 to 2008 (varying in number of years depending on the type of contracts studied) and includes 85 procurements of elderly care, 33 procurements of waste disposal, 32 procurements of food wholesaling, and finally 39 procurements of 39 internal cleaning services.

will in most cases not correspond to the ex ante stipulated weight. That is, the price on the environment ex post the procurement auction will not correctly reflect the social environmental shadow price.

6. Empirical experiences

Based on Swedish data from four product areas (cleaning, waste disposal, food, and elderly care), Bergman and Lundberg (2011) report frequent use of badly designed scoring rules. This is confirmed by Swedish data on procurements of internal cleaning service contracts from 2009, most recently collected by us. Due to the lack of public availability, our data are collected by means of a survey. Information on the award method is in general available in the call for tender while the scoring rule and the award criteria are found in the additional documents governing a public procurement auction. Our data include 153 procurements and 317 contracts. The procurements have been classified as either having some sort of environmental concern taken in the allocation of contracts (GPP) or not (no GPP). The GPP was either expressed as mandatory criteria for qualification of bidders, award criteria for evaluation of bids, or as mandatory criteria in combination with award criteria. GPP in some form was present in 68 percent of the procurements.

As an illustration to the theoretical analysis provided in this paper, descriptive statistics on the award methods and scoring rules used in the procurements studied are presented. As illustrated by Table 1, the lowest price (price-only competition) was applied in 39.5 percent of all procurements. The highest quality (quality-only competition) was applied in four of the 153 procurements. Referring back to the discussion linked to Figure 2 above, this indicates rational behavior, as internal regular cleaning contracts are characterized by low uncertainty about production costs (Hyytinen et al., 2007)[1].

Given that EMAT was the award method the most commonly used scoring rule was price-to-quality scoring (41.4 percent). The price-to-quality scoring rule given in equation (4) was the most commonly used rule and applied in 29.6 percent of the procurements. Interestingly, this model was more frequently used in the GPP procurements (33.6 percent). Consequently, within the scope of EMAT, the single most commonly used model had all the undesirable properties listed in the previous section including the risk of an arbitrary environmental policy and it was even more represented when GPP was enforced.

Table 1. Applied award methods and scoring rules in regular cleaning service procurements, 2009. Expressed in percentage of number of procurements

Model	All	GPP	No GPP	z (Probability value)
Price only	39.5	36.5	44.9	2.3 (0.0212)
Price to quality	41.4	41.3	40.8	3.5 (0.0004)
Equation 4	29.6	33.6	20.4	4.0 (0.0001)
Quality to price	16.5	20.2	8.2	3.4 (0.0007)
Quality only	2.6	1.9	4.1	0.6 (0.5618)
N	153	104	49	

A proportion test[2] suggests that the difference in proportions of the different award methods and scoring rules for GPP procurements with environmental concern (GPP) and GPP procurements cannot be rejected in all cases but quality only procurements. The test statistic (z) and probability values (in parenthesis) are reported in the rightmost column.

It is observed in our data that choice of award methods and scoring rules to some extent make sense in relation to the theoretical findings. The representation of badly designed scoring rules is from an environmental policy perspective however worrying high. This illustrates how important it is to really understand the award methods and scoring rules applied in the process of identifying GPP frontrunners. When evaluating GPP as a policy tool it is not enough to just document if GPP is applied, one must also understand how.

[1] Note, however, that it is a question for future research to find out if the mandatory requirements were of such level of stringency that they actually were binding and required some sort of environmental adjustment of the production technology and design.

[2] The test statistic for the difference of the two proportions is, given a normally distributed test statistics $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$ where $\hat{p}_p = \frac{x_1 - x_2}{n_1 + n_2}$ and n is the sample size and x is the number of successes in each treatment (GPP or not GPP). See Wang (2000).

The context presented here should be viewed as a starting point for future empirical research. It is not our intention to give a full analysis and the choice of award method and scoring rule is likely based on several factors of which GPP can be, but need not to be, one. The very simple empirical relationships as illustrated here cannot give a complete picture of the GPP as environmental policy tool. This requires more in-depth analysis and a review of what lies behind the choice of evaluation method, a choice based probably on more factors than the environmental.

Discussion and conclusions

It can be questioned whether or not GPP should be used as an environmental policy tool (Lundberg et al., 2009), but given that it is used as an environmental policy tool its design in terms of contract allocation methods is of crucial importance. All things equal, pure price competition in combination with mandatory environmental (GPP) requirements is recommended when there is low or no uncertainty about costs associated with the adjustment to the GPP requirements, or when it is important to reach a minimum level of environmental quality. Quality competition, on the other hand, is recommended in situations where there is considerable uncertainty about environmental adjustment costs, and thereby from a welfare and sustainability perspective a risk of too high a cost associated with environmental quality.

When awarding contracts based on combinations of price and environmental quality, and to avoid complexity and non-transparency, quality-to-price scoring is preferred over price-to-quality scoring. According to the quality-to-price scoring approach the subject-matter of the auction and its environmental impact should be valued in monetary terms. Furthermore, the explicitly chosen scoring rule should be based on absolute discounts or surcharges, when the supplier offers environmental quality in excess of a minimum quality required and when offering a quality in insufficiency of a stated maximum quality, respectively. In addition, as discussed in Bergman and Lundberg (2011) weighing price and quality is not recommended since it is meaningless and only adds complexity. Also, any scoring rule that builds on endogenous reference prices should be avoided since these models, depending on their exact design, to a larger or lesser extent involves a high risk for inefficient outcomes in terms of environmental improvement. Since a scoring rule that evaluates one bid in relation to another submitted bid or combinations of other submitted bids is non-transparent the optimal combination of price and environmental quality not might be submitted. It is in this perspective worrying that contracts, in Sweden in about one third of all procurement auctions, are allocated based on scoring rules of such nature. The representation of badly designed scoring rules is from an environmental policy perspective worrying and this paper illustrates how important it is to really understand the award methods and scoring rules applied in the process of identify of GPP front-runners. When evaluating GPP as a policy tool it is not enough to just document if GPP is applied, one must also understand how. The best environmental ambitions can namely be undermined with a bad choice of award method and ill conceived scoring rule.

Acknowledgement

Financial support from the Swedish Research Council is grateguly acknowledged.

References

1. Asker, J. and E. Cantillon (2008). Properties of Scoring Auctions, *Rand Journal of Economics*, 39 (1), pp. 69-85.
2. Asker, J. and E. Cantillon (2010). Procurement when Price and Quality Matter, *Rand Journal of Economics*, 41 (1), pp. 1-34.
3. Bergman, M. and S. Lundberg (2009). Att utvärdera anbud – Utvärderingsmodeller i teori och praktik, The Swedish Competition Authority, Commission Report 2009:10 (Summary in English).
4. Bergman, M. and S. Lundberg (2011). Tender Evaluation and Award Methodologies in Public Procurement, Umeå Economic Studies, No 821, Umeå University.
5. Bernheim, B.D. and M. Whinston (1986). Menu Auctions, Resource Allocation, and Economic Influence, *Quarterly Journal of Economics*, 101 (1), pp. 1-31.
6. Che, Y.-K. (1993). Design Competition through Multidimensional Auctions, *Rand Journal of Economics*, 24 (4), pp. 668-680.
7. Clinch, J.P., K. Schlegelmilch, R.-U. Sprenger and U. Triebswetter (2002). *Greening the Budget: Budgetary Policies for Environmental Improvement*, Edward Elgar, Series: International Studies in Environmental Policy Making, Cheltenham.
8. D'Amoto, A. (2006). Environmental Issues in Public Procurement: How much Decentralization? *Revista di Politica Economica*, 96 (1-2), pp. 209-234.
9. Dini, F., R. Pacini and T. Valletti (2006). Scoring Rules, in Dimitri, N., G. Piga and G. Spagnolo (Eds.), *Handbook of Procurement*, Cambridge University Press.
10. Erdmenger, C. (2003). Buying into the Environment, Experiences, Opportunities and Potential for Eco-

Procurement, International Council on Local Environmental Initiatives (ICLEI), Greenleaf Publishing Limited, Sheffield.
11. European Commission (2003). Integrated Product Policy – Building Environmental Life-Cycle Thinking, Communication from the Commission to the Council and the European Parliament, COM (2003) 302 final.
12. European Commission (2008). Public Procurement for a Better Environment, Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions, COM(2008) 400 final.
13. European Commission (2010). Europe 2020 – A Strategy for Smart, Sustainable and Inclusive Growth, Communication from the Commission to the Council and the European Parliament, COM (2010) 2020.
14. Geng, Y. and B. Doberstein (2008). Greening Government Procurement in Developing Countries: Building Capacity in China, *Journal of Environmental Management*, 88 (4), pp. 932-938.
15. Gothenburg European Council (2001). Presidency Conclusions, SN 200/1/01 REV 1, Gothenburg, 15 and 16 June, 2001.
16. Green Paper (2011). On the Modernization of EU Public Procurement Policy – Towards a more Efficient European Procurement Market, COM (2011) 15 final.
17. Grolleau, G., N. Mzoughi and C. Nouira (2004). Public Purchasing and Eco-Labelling Schemes: Making the Connection and Reinforcing Policy Coherence, *Journal of Interdisciplinary Economics*, 15 (2), pp. 131-151.
18. Hyytinen, A., S. Lundberg, and O. Toivanen (2007). Politics and Procurement: Evidence from Cleaning Service Contracts. Discussion Paper No 196, Helsinki Center of Economic Research, Discussion Papers.
19. Kahlenborn W., C. Moser, J. Frijdal, and M. Essing (2011). Strategic Use of Public Procurement in Europe – Final Report to the European Commission MARKT/2010/02/C, Berlin: Adelphi.
20. Kippo-Edlund, P., H. Hauta-Heikkilä, H. Miettinen and A. Nissinen (2005). Measuring the Environmnetal Soundness of Public Procurement in Nordic Countries, TemaNord 2005, p. 505.
21. Klemperer, P. (1999). Auction Theory: A Guide to the Literature, *Journal of Economic Surveys*, Wiley Blackwell, Vol. 13 (3), pp. 227-286.
22. Lunader, A. (2009). En logisk fälla – Relativ poängsättning av pris vid anbudsutvärdering i offentlig upphandling, The Swedish Competition Authority, Commission Report 2009, p. 12 (Summary in English).
23. Lunander, A. and A. Andersson (2004). Metoder vid utvärdering av pris och kvalitet i offentlig upphandling – En inventering och analys av utvärderingsmodeller inom offentlig upphandling, The Swedish Competition Authority, Commission Report 2004, p. 1 (Summary in English).
24. Lundberg, S., P.-O. Marklund and R. Brännlund (2009). Green Public Procurement as a Policy Tool: Cost-Effectiveness and Competition Considerations, Umeå Economic Studies, No 775, Umeå University.
25. Marron, D.B. (1997). Buying Green: Government Procurement as an Instrument of Environmental Policy, *Public Finance Review*, 25 (3), pp. 285-305.
26. Michelsen, O. and L de Boer (2009). Green Public Procurement in Norway: A Survey of Practices at the Municipal and County Level, *Journal of Environmental Management*, 91 (1), pp. 160-167.
27. Milgrom, P. (2004). *Putting Auction Theory to Work*, Cambridge Books, Cambridge University Press.
28. Milgrom, P. (1989). Auctions and Bidding: A Primer, *Journal of Economic perspectives*, 3 (3), pp. 3-22.
29. Nissinen, A., K. Parikka-Alhola and H. Rita (2009). Environmental Criteria in the Public Purchases above the EU Treshhold Values by Three Nordic Countries: 2003 and 2005, *Ecological Economics*, 68 (6), pp. 1838-1849.
30. Parikka-Alhola, K. (2008). Promoting Environmentally Sound Furniture By Green Public Procurement, *Ecological Economics*, 68 (1-2), pp. 472-485.
31. Sterner, E. (2002). 'Green Procurement' of Buildings: A Study of Swedish Clients' Considerations, *Construction Management and Economics*, 20 (1), pp. 21-30.
32. Thomson, J. and T. Jackson (2007). 'Sustainable Procurement in Practice: Lessons from Local Government, *Journal of Environmental Planning and Management*, 50 (3), pp. 421-444.
33. Vickrey, W. (1961). Counterspeculation, Auctions and Sealed Competitive Tenders, *Journal of Finance*, 16 (1), pp. 8-37.
34. Wang, D. (2000). Confidence Intervals for the Ratio of Two Binomial Proportions by Koopman's Method, *Stata Technical Bulletin*, 58, pp. 16-19, sg154.
35. Weitzman, M.L. (1974). Prices vs. Quantities, *Review of Economic Studies*, 41(4), pp. 477-491.