EasyChair Preprint

№ 1120

The Importance of Non-Price Criteria in Government Contracting with For-Profit Enterprises

Amandine Lerusse and Steven Van de Walle

June 8, 2019

The Importance of Non-Price Criteria in Government Contracting with For-Profit Enterprises

Results from a discrete choice experiment among local governments

Amandine Lerusse (KU Leuven)

Steven Van de Walle (KU Leuven)

Draft

Abstract

Public institutions mostly favor neo-classical contracting in public procurement by choosing the external service provider offering the cheapest option. However, non-price criteria such as environmental, innovative and social criteria are increasingly taken into consideration when contracts are awarded. The objective of this study is to examine local civil servants' valuation of non-price criteria when awarding contracts to external service providers. For this purpose, we conduct a discrete choice experiment in the field of waste collection at the municipal level. By applying random utility theory, we claim that local civil servants make a trade off between the four criteria. Consequently, local civil servants intend to maximize the utility of the municipality by awarding the contract to the external service provider producing the highest utility. We analyze local civil servants' valuation of non-price criteria through a conditional logistic regression and a marginal willingness to pay analysis. Although the price criterion still remains at the center of public procurement, the results also show that local civil servants take non-price criteria into consideration when awarding contracts to for-profit enterprises. More specifically, local civil servants are willing to pay more to increase the level of the environmental, innovative and social criteria.

Keywords: Random utility theory, contracting out, discrete choice experiment, local governments.

Introduction

Public institutions usually follow the doctrine of neo-classical contracting out and prioritize the price criterion over more quality-oriented criteria in public procurement. Consequently, public institutions typically award contracts to the cheapest providers. This choice is not solely driven by economic considerations. As a matter of fact, decision makers also believe that basing their decision on subjective non-price criteria could jeopardize the process of public procurement by exposing it to abuse and preventing its fair access (Keulemans and Van De Walle 2017).

It is often argued that choosing the less expensive provider may negatively impact the quality of the public service because cheap bids may undermine other public values that could be worth pursuing through public procurement (Mazzucato 2011). In this context, we observe increasing interest for including non-price criteria in public procurement as a way to promote environmentally friendly policies, social justice, good governance and innovation (Walker and Brammer 2009; Uyarra et al. 2014). Yet, many obstacles to the integration of non-price criteria remain within public procurement (Walker and Brammer 2009; Uyarra et al. 2014).

Although including non-price criteria in public procurement can be considered an essential policy instrument that governments can use to pursue a range of policy objectives, the extent to which civil servants make use of this tool remains unclear. Furthermore, limited studies have examined civil servants' behavior in public procurement (Grandia 2016). The goal of this study is therefore to examine how local civil servants value non-price criteria when awarding contracts to external service providers. The stated behavior of local civil servants for non-price criteria are examined through random utility theory. We argue that local civil servants make a trade off between the environmental, innovative, price and social criteria and aim to maximize the utility of the municipality by awarding the contract to the external service provider producing the highest utility. As random utility theory finds its roots in the field of psychology,

we intend to address the lack of connection between public administration and psychology highlighted by Simon (1955) (Olsen 2015; Grimmelikhuijsen et al. 2016).

We analyze local civil servants' valuation of non-price criteria through a discrete choice experiment. This stated preferences method, rarely employed in the field of public administration, allows us to measure individuals' stated preferences that cannot be directly observed (Mangham, Hanson, and Mcpake 2009; Louviere, Flynn, and Carson 2010). We examine the awarding of contracts in the field of bulky waste items collection at the municipal level. We surveyed the heads of the environmental department or the environmental advisors of the municipality in four European countries. The focus is on for-profit enterprises because they constitute the most common external service provider chosen by local authorities to deliver public services (Schoute, Budding, and Gradus 2018).

This article is divided into five sections. The first section starts with a review of the literature on non-price criteria. The second section elaborates on random utility theory as the theoretical background explaining how local civil servants take the four criteria into consideration. The third section describes the design of the discrete choice experiment. The fourth section presents the empirical findings. The last section discusses the main findings and draws some conclusions.

Beyond the price criterion in public procurement

The price criterion has for many years been at the center of public procurement decisions. Public institutions argue that, compared to more subjective criteria, awarding contracts based on the price criterion was more objective and easier to justify (Thomson and Jackson 2007; Keulemans and Van De Walle 2017). However, the past years have seen the development of sustainable and innovative public procurement as a policy instrument to attain environmental,

social and innovative objectives that would be difficult to reach differently (Rolfstam 2009; Testa et al. 2012; Keulemans and Van De Walle 2017).

Sustainable public procurement includes environmental and social criteria and aim to promote environmental respect, social justice and good governance (McCrudden 2004; Walker and Brammer 2009; Furneaux and Barraket 2014; Keulemans and Van De Walle 2017). The inclusion of an environmental criterion into public procurement intends to reduce the environmental pressures and to stimulate the development of more environmentally friendly public goods and services (Testa et al. 2012). Promoting low-carbon vehicles, developing renewable energies and low energy building are examples of environmental criteria that can be included in tender documents (European Commission 2016). The environmental criterion appears to be more and more included in public procurement. In their analysis of the uptake of green procurement among European countries, Renda et al. (2012) observe, that 60 percent of the EU member states, 67 percent of EU regional governments and 67 percent of local governments have already included an environmental component in their public procurement policy.

With regards to the social criterion, Loosemore (2016, 133) argues that "social procurement differs from traditional procurement in the use of procurement to leverage extra social benefits and create 'social value' in local communities". Examples of social criteria in public procurement consist of objectives such as promoting the employment of disadvantaged groups, reducing unemployment or improving employment practices (McCrudden 2004). Just as the environmental criterion, public institutions are increasingly including the social criterion in tender documents. For instance, in their report on the integration of social values in public procurement in the United Kingdom, Temple, Wigglesworth, and Smith (2014) argue that 62 percent of local authorities take social values into consideration when procuring public services.

Public authorities can also take a lead in creating innovation by stimulating it through public procurement (Mazzucato 2011). By including an innovative criterion into the contract, public institutions intend to correct for market failures and to promote research and development (Edler and Georghiou 2007; Edquist and Zabala-iturriagagoitia 2012). More specifically, through innovative public procurement, public institutions aim to deliver higher quality public services at lower prices, meet new market needs, modernize public services, stimulate the development of starts-ups and small and medium-sized enterprises (SMEs) (European Commission 2018). The European Union has recently demonstrated a renewed interest in developing innovative public procurement further by publishing a Commission notice. This documents provides guidance to public institutions on how to implement innovative public procurement (European Commission 2018).

Although public institutions increasingly value environmental, social and innovative criteria, there are economic, cognitive and political barriers to their implementation in public procurement policies. First, one of the main obstacles to the inclusion of non-price criteria in public procurement remains their cost (Walker and Brammer 2009; Thomson and Jackson 2007). Awarding contracts based on other criteria than price might substantially increase the cost of the public service and might put additional fiscal pressures on governments.

A second barrier is the lack of awareness from public institutions on how to include and evaluate non-price criteria into public procurement (Walker and Brammer 2009; Erridge and Hennigan 2012). Some civil servants claim that they do not always have the capacity and the expertise to evaluate non-price standards specified in the bid (Walker and Brammer 2009; Testa et al. 2012). Similarly, other civil servants state that they lack the necessary trainings and information to implement non-price criteria in public procurement (Thomson and Jackson 2007; Testa et al. 2012). As a matter of fact, while the impact of the price criterion is relatively easy to identify and quantify, analyzing the impact of non-price criteria is more complex and

requires more knowledge on their consequences on public goods and services. In addition, regulations on how to include and evaluate non-price criteria in public procurement remain unclear to many civil servants (Erridge and Hennigan 2012).

Third, Thomson and Jackson (2007) underline that non-price criteria are often disregarded by high level public officials who do not always prioritize non-criteria in their public procurement policies. Finally, Uyarra et al. (2014, 640) underline additional barriers such as "a lack of interaction with procuring organizations, the use of rigid as opposed to outcome-based specifications, low competences of procurers and a poor management of risk".

In light of the abovementioned barriers, it is crucial to examine the extent to which civil servants are willing to take non-price criteria into consideration when awarding contracts. Moreover, limited studies have analyzed the extent to which public institutions simultaneously consider environmental, innovative, price and social criteria in public procurement. Although Grandia (2016, 189) points out that "the application of sustainable procurement is indeed directly influenced by the behavior (and thus decisions) of the procurers", studies examining the stated preferences of civil servants for non-price criteria remain scarce. The present research therefore aims at filling these gaps by studying civil servants' stated preferences for the environmental, innovative and social criteria when awarding contracts.

Theoretical framework

Local civil servants' decision making in public procurement

The awarding of contracts is a typical field of decision making in public administration. Local civil servants who have to choose, on behalf of their institution, between competing external service providers make use of a decision rule to process the information they have at their disposal. The literature reports four main decision rules: dominance, satisfaction, lexicographic rules and utility (Slovic, Fischhoff, and Lichtenstein 1977; Svenson 1979; Ben-Akiva and

Lerman 1985). In contrast with dominance, satisfaction and lexicographic rules, that are known as non-compensatory decision making rules, utility is considered a compensatory decision rule where decision makers are expected to make a trade off between attributes. This last category of decision rule assumes that each alternative can be reduced to a single index of attractiveness that can be referred to as utility. By their choice, the main objective of decision makers is to maximize this utility (Ben-Akiva and Lerman 1985).

In the present case, we claim that local civil servants make a trade off between the environmental, innovative, price and social criteria and that local civil servants intend to maximize the utility of the municipality by choosing the external service provider that offers the best combination of the criteria. In addition, due to the growing importance given to non-price criteria by public institutions (Keulemans and Van De Walle 2017), we argue that local civil servants' decision rule is compensatory rather than non-compensatory. Therefore, we assume that the utility decision rule is the most adequate to describe their choice and that the competing external service providers among which they can choose from are reduced to utility values. Local civil servants choose the external service provider that produces the highest utility or, in other words, the best combination of price, environmental, innovative and social criteria. This means that local civil servants are willing to award the contract to a more expensive external service provider if one or more non-price criteria is/are superior for one of the external service providers.

Random utility theory

We argue that local civil servants make use of a compensatory decision making rule and make a trade off between the environmental, innovative, price and social criteria. Moreover, we assume that observed factors do not solely influence local civil servants' stated preferences. As a matter of fact, we claim that a random component, composed of unobserved factors, also

affects local civil servants' stated preferences. This random component can include variables that are linked to the environment of the municipality (e.g. financial and/or political situation of the municipality) and/or to the local civil servants themselves (e.g. their personal experiences with the evaluation of tenders). Consequently, we decide to employ random utility theory to examine local civil servants' valuation of non-price criteria because it assumes that latent preferences are composed of an observed and a random component.

Random utility theory has its foundations in psychology and probabilistic choice theory (Thurstone 1927; Manski 1977; Kjær 2005). Probabilistic choice theories were first developed to clarify why some decision makers had inconsistent and non-transitive preferences that would go against the concept of rational behavior (Ben-Akiva and Lerman 1985). Inconsistent preferences mean that under the exact same conditions, a local civil servant would not choose the same for-profit enterprise. Non-transitive preferences imply that if enterprise A is preferred over enterprise B and enterprise B is preferred over enterprise C this does not necessarily indicate that enterprise A is preferred over enterprise C (Ben-Akiva and Lerman 1985). Probabilistic choice theories therefore assert that there is a degree of uncertainty surrounding individuals' choice. As a consequence, probabilistic choice theories do not predict which alternative decision makers will opt for (Kjær 2005; Cascetta 2009). Instead, the model predicts the probability that the alternative with the greatest utility for decision makers will be chosen over the other possible alternatives (Kjær 2005; Cascetta 2009). In order to explain these behavioral inconsistencies, a dichotomy of probabilistic choice theories started to emerge (Ben-Akiva and Lerman 1985; Kjær 2005). The first view assumes that individuals' behavior is intrinsically probabilistic and fluctuates according to internal and external characteristics. The second view claims that researchers cannot observe individuals' behavior with certainty.

Random utility theory is built on the second view where "utilities are treated as random variables not to reflect a lack of rationality in the decision maker but to reflect a lack of

information regarding the characteristics of alternatives and/or decision makers on the part of the observer" (Manski 1977, 229). Random utility theory is therefore consistent with neoclassical economic theory that considers individuals to be rational and utility-maximizers. As a matter of fact, researchers predict the probability that the individuals will opt for the alternative which produces the highest level of utility for decision makers (Manski 1977; Kjær 2005). Random utility theory assumes that the indirect utility function can be decomposed into two utility functions. The first one comprises all the characteristics that are known by researchers while the second one represents the characteristics influencing the choices made by decision makers but that are not observed by researchers (Kjær 2005). More specifically, random utility theory states that latent preferences, also called latent utilities, are associated with all choices under scrutiny and exist for all individuals (Louviere, Flynn, and Carson 2010). Those latent preferences are composed of an observed and a random component. The observed component comprises the factors clarifying variation in individuals' choices (covariates) as well as in the choice of alternatives (attributes) that one can observe while all the undetermined variables, influencing individuals' choice, compose the random component (Louviere, Flynn, and Carson 2010).

Local civil servants' stated behavior in public procurement

In light of the relatively recent importance given to non-price criteria in public procurement (Keulemans and Van De Walle 2017), we assume that local civil servants award contracts to external service providers based on the utility decision rule. We argue that local civil servants make a trade off between the four criteria and, therefore, aim to maximize the utility of the municipality by choosing the external service provider that produces the highest utility. Nevertheless, a random component, composed of unobserved factors (e.g. characteristics linked to the environment of the municipality and/or to the local civil servants themselves), also determines local civil servants' stated preferences. We are therefore unable to predict local

civil servants' criteria preferences with certainty. Instead, we predict the probability that the external service provider with the highest utility for a local civil servant is chosen over the other external service provider. In order to investigate how local civil servants value non-price criteria when awarding contracts to external service providers, we use a discrete choice experiment. This method is based on random utility theory and is further introduced in the next section (Louviere, Flynn, and Carson 2010).

Data and Method

Discrete choice experiments aim at eliciting individuals' stated preferences for products or services in existing and non-existing markets (Lancsar and Louviere 2008; Louviere, Flynn, and Carson 2010). In discrete choice experiments, individuals' stated preferences are derived from a survey where respondents have to choose between two or more options across several choice sets given a hypothetical scenario. A combination of attributes which can range from one to several levels describe the options displayed to the respondents (Lancsar and Louviere 2008; Hoyos 2010; Hauber et al. 2016; Lancsar, Fiebig, and Hole 2017). For instance, when awarding a contract to a for-profit enterprise, a significant attribute for civil servants to choose the contractor might be the price of the service. This price can take several levels such as 50 euros, 100 euros or 150 euros.

While choosing between two or more options, individuals have to make a trade-off between different attributes. One can examine individuals' stated preferences by measuring the impact of every attribute level on choice. In addition, including a monetary measure in the discrete choice experiment allows the researcher to identify respondents' marginal willingness to pay for a level change in one attribute (Ryan 2004; Hoyos 2010). The marginal willingness to pay is defined as the maximum amount of money an individual is inclined to pay for a good or a service (Gafni 1998). One can employ the marginal willingness to pay to analyze how much

individuals value a good or a service. The concept marginal indicates that the marginal willingness to pay is always measured relative to a baseline good or service. Although many fields such as economics, health economics, marketing and transportation conducted discrete choice experiments, this particular method has rarely been used in the field of public administration.

Research setting

The focus of the study is on how municipal civil servants value non-price criteria when awarding contracts to for-profit enterprises in the field of waste collection and, more particularly, bulky waste items collection. The research focuses on for-profit-enterprises. Among the different external service providers, local authorities most often award the contract to for-profit enterprises (Schoute, Budding, and Gradus 2018). We selected the field of bulky waste items collection as this is a rather straightforward non-sensitive public service where contracting out with for-profit enterprises is relatively common (Brown and Potoski 2005; Bel and Costas 2006; Schoute, Budding, and Gradus 2018). Bulky waste items are a particular type of waste which is too big to be placed in standard waste containers such as old televisions, washing machines, coaches, etc. They can therefore be placed on the pavement to be picked up by waste collection companies or brought to container parks. In addition, this service is rather homogeneous across countries making data collection on different countries possible.

In the hypothetical scenario presented to local civil servants, we specify that the municipality for which they are working has decided to change to a different bulky waste items collector. The scenario requires local civil servants to select which one, out of the two presented for-profit enterprises, should become the new bulky waste items collector of their municipality. The for-profit enterprises are described by price, environmental, innovative and social criteria.

Design of the discrete choice experiment

Attributes and levels selection

Non-price criteria can be used to pursue environmental, innovative and social objectives in public procurement. In the field of waste collection, several initiatives have been promoted with regards to non-price criteria. For instance, in Canada, hybrid and silent waste collection trucks have been developed to reduce greenhouse gas and noise pollution. Regarding innovation, the city of Grenoble in France has installed sensors on the glass bottle banks. The sensors are linked to a smartphone app where individuals can accumulate bonus points if they correctly recycle their glass bottles. They can then spend these bonus points in neighboring shops. In Belgium, several bulky waste items enterprises take the social criterion into consideration by exclusively hiring vulnerable groups.

The for-profit enterprises are described by price, environmental, innovative and social criteria. Each criterion takes different levels. We operationalized the criteria by conducting exploratory semi-structured interviews with waste collection experts from waste collection agencies, a bulky waste items collection enterprise, inter-municipal associations and a municipality. We asked the interviewees to specify which parameters were the most appropriate and realistic to operationalize the four criteria. Following these interviews, we derived four attributes and their respective levels that are displayed in table 1. The first attribute, the price criterion, describes the price per ton of bulky waste items collected and has two levels. These levels are 250 euros and 270 euros per ton of bulky waste items collected. Second, the environmental criterion, showing the average age of the fleet of vehicles, is composed of three levels: 6 years, 3 years and 0 years (a new fleet). Third, the innovative criterion has two levels and indicates whether the for-profit enterprise offers or not an app for smartphone to the users of the service. The last attribute, the social criterion, has two levels and

displays whether the for-profit enterprise is currently involved or not in a training scheme for long-term unemployed.

The attributes and their respective levels correspond to the reality of the market and, except from the price per ton of bulky waste items collected that has continuous levels, the other attributes have categorical levels. We included the price per ton of bulky waste items to also be able to calculate local civil servants' marginal willingness to pay for a change in one of the non-price attribute level. In addition, except for the price attribute that was adapted to the context of the country, the attributes and their respective levels are exactly the same for every country.

Table 1. Attributes and their respective levels

	Level 1	Level 2	Level 3
Price per ton of bulky waste items collected (price criterion)[1]	The price per ton is 250 EUR	The price per ton is 270 EUR	
The average age of the fleet of vehicles (environmental criterion)	The average age of the fleet of vehicles is 6 years.	The average age of the fleet of vehicles is 3 years.	The average age of the fleet of vehicles is 0 years (an entire new fleet will be put into circulation).
Digital services (innovative criterion)	The for profit enterprise does not offer an app for smart-phones for service users.	The for-profit enterprise offers an app for smart-phones for service users (this app has a calendar, informs on pick-up days and offers a contact tool,…).	
The for-profit enterprise's involvement in the professional integration of vulnerable groups (social criterion)	The for-profit enterprise is not currently involved in a training scheme for long-term unemployed.	The for-profit enterprise is currently involved in a training scheme for long-term unemployed.	

Fractional factorial design

The number of hypothetical scenarios in the discrete choice experiment is 24 (three attributes with two levels and one attribute with three levels) and there are 276 possible choice sets to be

[1] As the standard of living is lower in Estonia in comparison with the other analyzed countries, the price of the service was adapted to 80 euros and 90 euros. In Norway, the price of the service was adjusted to 2400 Norwegian krone (~ 250 euros) and 2600 Norwegian krone (~ 270 euros) as Norway is not part of the Eurozone.

shown to the respondents. We performed a fractional factorial design to reduce the number of possible choice sets. Lancsar and Louviere (2008, 667) define a fractional factorial design as "a sample from the full factorial selected such that all effects of interest can be estimated (at a minimum, the main effects, but also as many higher-order interaction effects as possible)". We solely identified the main effects because we do not expect any interaction between the criteria. The decision to award the contract is often forced implying that local civil servants cannot opt out from the decision making process. We therefore did not include an opt-out option in the study design. In addition, we did not label the choice alternatives to keep the for-profit enterprises as generic as possible. The aim was not to influence local civil servants towards a particular brand.

We generated the fractional factorial design by computing an orthogonal main effect array with the rotation method in R (Aizaki 2012). This type of design insures that the attributes are statistically independent from each other (orthogonality) and that every level has the same probability to appear throughout the choice sets (level balance) (Ryan et al. 2012). The fractional factorial design reduced the number of possible choice sets to 12 and divided them into two blocks of 6 choice sets (see the appendix for an example of a choice set). We randomly assigned the local civil servants to one of the two blocks where each of them had to evaluate six choice sets.

Sample and data collection

According to the exploratory semi-structured interviews we conducted, the head of the environmental department and the environmental advisor of a municipality are the best respondents. As a matter of fact, they are the ones designing tender documents and evaluating bids in the area of waste collection on behalf of their local institution. Where there were no

heads of the environmental department or no environmental advisors, we selected the environmental alderman of the municipality or someone with an equivalent position.

We conducted the discrete choice experiment in Belgium, Estonia, Germany and Norway. These four countries demonstrate similar trends regarding public procurement (OECD 2017). First, the figures indicate that the four countries present comparable general government procurement as a percentage of GDP. Second, the four countries have also all developed, either at the central level or by procuring entities, contracting out strategies or policies to support green public procurement, small and medium-sized enterprises (SMEs) and innovative goods and services. Finally, the figures point out that the four countries spend similar percentages of their general government procurement spending on environmental protection (OECD 2017). It is worth noting that the objective of this article is not to explain cross-country differences regarding the relative significance of the different criteria to local civil servants. Alternatively, we aim to study the stated preferences of local civil servants, coming from countries with similar public procurement trends, when awarding contracts to for-profit enterprises.

The discrete choice experiments were translated in Dutch, French, Estonian, German and Norwegian by natives speakers. We sent personalized emails to the local civil servants who had to fill in the survey experiment electronically via Qualtrics. In Belgium, Estonia and Norway, we sent the discrete choice experiment to every municipality which consists of a population of 546 Belgian[2] local civil servants, 77 Estonian[3] local civil servants and 418 Norwegian[4] local civil servants. Germany is composed of more than 12.000 municipalities. We have therefore decided to exclusively examine municipalities with more than 20.000

[2] Belgium is composed of 589 municipalities but contact details were not available for every municipality and we deleted a respondent as this person pretested the discrete choice experiment. Therefore, the total sample of Belgian municipalities is 546.
[3] Estonia is composed of 79 municipalities but we deleted 2 respondents from the sample as they pretested the discrete choice experiment. Therefore, the total sample of Estonian municipalities is 77.
[4] Norway is composed of 422 municipalities but contact details were not available for every municipality and we deleted the respondents who pretested the discrete choice experiment. Therefore, the total sample of Norwegian municipalities is 418.

inhabitants and the counties which are also responsible for waste collection in Germany. The German sample consists of 901 municipalities and counties[5].

We verify the internal validity of the answers given by the local civil servants once the data collection is completed. Thomas and Clifford (2017) argue that ex-post exclusion allows the researchers to compare the respondents that are excluded to the ones that are not excluded and yields more transparent results compared to ex-ante exclusion. We decided to perform ex post exclusion of the respondents by examining the time local civil servants have taken to fill in the survey experiment. The minimum time to fill in the discrete choice experiment is approximately four minutes. As a consequence, we will verify all the choices made by local civil servants who have taken less than four minutes to complete the survey experiment. Greszki, Meyer, and Schoen (2015, 471) find that "removing "too fast" responses does not alter marginal distributions, irrespective of which speeder-correction technique is employed".

Statistical analysis

In the discrete choice experiment, the choice made by a local civil servant in a given choice set is either 0, for the for-profit enterprise that was not chosen, or 1, for the for-profit enterprise that was chosen. Data from forced-choice discrete choice experiments composed of two alternatives are best investigated by employing a limited dependent-variable model such as a conditional (fixed effects) logistic regression (Hauber et al. 2016). In addition, McFadden (1974) demonstrates that this model is consistent with random utility theory. The conditional logistic regression assumes the independence from irrelevant alternatives property stating that "the introduction or removal of a choice has no effect on the proportion of probability assigned to each of the other choices" (Ryan et al. 2012, 37). Stated differently, this property does not

[5] In Germany, the sample of respondents was 988 but contact details could not be found for every municipality and county. Moreover, we deleted two respondents from the sample as they pretested the discrete choice experiment. Therefore, the total sample of German municipalities is 901.

allow the error of an alternative to be related to the error of another alternative (Train 2002). We also analyze the relative importance of every attribute level to local civil servants by estimating the marginal willingness to pay and the changes in probabilities of choosing the baseline for-profit enterprise when the level of one of the for-profit enterprise has been modified.

Empirical Analysis

As mentioned in the previous section, we conduct an ex post exclusion of the respondents by examining the answers of the local civil servants who took less than four minutes to complete the survey experiment. One respondent was problematic because she/he took less than four minutes to complete the discrete choice experiment and selected the second alternative in all the choice sets. We therefore decided to delete this respondent from the data. After the completion of the internal validity check, the response rates is 30.8 percent in Belgium, 61 percent in Estonia, 13.9 percent in Germany and 28.2 percent in Norway. A total of 456 local civil servants' responses are therefore examined; 166 in Belgium, 47 in Estonia, 125 in Germany and 118 in Norway. The six choices taken by every local civil servant, also called the number of event, is 996 in Belgium, 282 in Estonia, 750 in Germany and 708 in Norway. The number of observations (the number of events times two as respondents could choose between two for-profit enterprises per choice set) is 1992 in Belgium, 564 in Estonia, 1500 in Germany and 1416 in Norway.

Conditional logistic regression and changes in probabilities of winning the contract for a for-profit enterprise

Table 2 presents the results of the conditional logistic regression and displays the odds ratios (see the appendix for the coefficients and the standard errors), the number of observations, the

number of local civil servants, the likelihood ratio and the McFadden pseudo R^2 [6] for the four countries. An odds ratio below 1 means that the attribute has a negative effect on local civil servants' decision to award the contract to a for-profit enterprise while an odds ratio above 1 means that the attribute has a positive effect on local civil servants' decision to award the contract to a for-profit enterprise. Table 3 shows the change in the probability of awarding the contract to the baseline for-profit enterprise (highest price, average age of the fleet of vehicles of 6 years, no app and no involvement of the for-profit enterprise in a training scheme) when the level of one of the for-profit enterprise attributes has been changed. While table 2 provides information on the likelihood of an attribute level to be chosen by local civil servants, table 3 displays more detailed information on the probability of a for-profit enterprise to win the contract when the level of one the attributes has been changed.

[6] Contrary to most regression models, the conditional logistic regression does not allow the interpretation of an adjusted R squared. Instead, the log-likelihood, which "is an indicator of the relative explanatory power of a model", is used to calculate a goodness of fit measure called the McFadden's pseudo R^2 (Hauber et al. 2016, 307). The McFadden's pseudo R^2 can take a value between 0 and can never attain 1.

Table 2. Conditional logistic regression (Odds ratios)

	Belgium	Estonia	Germany	Norway
Price criterion	0.962***	0.890***	0.936***	0.994***
Environmental criterion (3years) Ref. Cat[7].: 6 years	2.283***	1.847**	2.516***	3.070***
Environmental criterion (0yrs) Ref. Cat.: 6 years	3.139***	1.503	3.280***	4.868***
Innovative criterion Ref. Cat.: no app	2.688***	2.868***	3.325***	2.908***
Social criterion Ref. Cat.: no training scheme	3.568***	0.974	2.146***	2.231***
ASC	1.141	0.992	0.934	1.123
Number of observations	1992	564	1500	1416
Number of local civil servants	166	47	125	118
LR Chi² (6)	538.71	109.11	437.95	408.72
Prob > Chi²	0.0000	0.0000	0.0000	0.0000
McFadden pseudo R²	0.390	0.279	0.421	0.416

*** p < 0.001, ** p < 0.01, * p < 0.05

[7] Ref. Cat. refers to the reference category.

Table 3. Changes in probabilities

	Changes in probabilities[8]			
Change from baseline	Belgium	Estonia	Germany	Norway
Price (250 euros, 70 euros and 2400 Norwegian krone)	0.373***	0.523***	0.580***	0.509***
	(0.039)	(0.060)	(0.038)	(0.042)
Environment (3yrs) Ref. Cat.: 6 years	0.391***	0.297**	0.431***	0.509***
	(0.051)	(0.098)	(0.060)	(0.056)
Environment (0yrs) Ref. Cat.: 6 years	0.517***	0.201	0.533***	0.659***
	(0.049)	(0.104)	(0.060)	(0.049)
App Ref. Cat.: no app	0.458***	0.483***	0.538***	0.488***
	(0.037)	(0.063)	(0.041)	(0.044)
Involved in a training scheme Ref. Cat.: no training scheme	0.562***	-0.013	0.364***	0.381***
	(0.033)	(0.077)	(0.046)	(0.046)
Number of observations	1992	564	1500	1416
Number of local civil servants	166	47	125	118
LR Chi² (6)	538.71	109.11	437.95	408.72
Prob > Chi²	0.0000	0.0000	0.0000	0.0000
McFadden pseudo R²	0.390	0.279	0.421	0.416

*** p < 0.001, ** p < 0.01, * p < 0.05

Price criterion

The odds ratios, displayed in table 2, point out that the price criterion remains very much at the center of the awarding of contracts. Local civil servants are more likely to award the contract

[8] Calculations of the *P*-values and the standard errors were performed by the nlcom-command in Stata. This command bases its calculations on the delta method.

to a for-profit enterprise that has a cheaper public service. Table 3 shows that reducing the price of the public service from 270 euros to 250 euros increases the probability of a for-profit enterprise to win the contract by 37.3 percentage points in Belgium and by 58 percentage points in Germany. In Estonia, decreasing the price of the service by 10 euros raises the probability of a for-profit enterprise to win the contract by 52.3 percentage points while decreasing the price by 200 Norwegian krone in Norway increases the probability of a for-profit enterprise to win the contract by 50.9 percentage points. These results are statistically significant. However, although the price criterion remains very much at the center of the awarding of contracts across the four countries, the results also demonstrate that local civil servants consider non-price criteria when awarding contracts to for-profit enterprises. The next sections present local civil servants' stated preferences for the environmental, innovative and social criteria.

Environmental criterion

The results point out that Belgian, Estonian, German and Norwegian local civil servants all take the environmental criterion into consideration when awarding contracts to for-profit enterprises. The analysis in table 2 indicates that local civil servants across the four countries are more likely to award the contract to a for-profit enterprise that puts into circulation a fleet of vehicles of an average age of 3 years compared to a for-profit enterprise that puts a fleet of vehicles of an average age of 6 years into circulation. In addition, table 3 shows that reducing the average age of the fleet of vehicles from 6 years to 3 years increases the probability of a for-profit enterprise to win the contract by 39.1 percentage points in Belgium, 29.7 percentage points in Estonia, 43.1 percentage points in Germany and 50.9 percentage points in Norway. These findings are statistically significant.

The odds ratios, displayed in table 3, point out that, except from Estonia, local civil servants are more likely to award the contract to a for-profit enterprise that puts into circulation a new fleet of vehicles (0 years) over an enterprise that puts into circulation a fleet of vehicles of an

average age of 3 years. The analysis, presented in table 3, also indicates that putting a new fleet of vehicles into circulation compared to a fleet of vehicles of an average age of 3 years does not substantially increase an enterprise's probability of winning the contract. Some for-profit enterprises might therefore be more reluctant to put a new fleet of vehicles into circulation compared to a fleet of an average age of 3 years. As a matter of fact, reducing the average age of the fleet of vehicles from 3 years to a new fleet of vehicles (0 years) raises the probability of a for-profit enterprise to win the contract by 12.6 percentage points in Belgium, 10.2 percentage points in Germany and 15 percentage points in Norway. These results are statistically significant though.

One can argue that examining the environmental criterion in the discrete choice experiment may slightly bias the results upward due to the fact that the respondents work in the environmental field. Yet, we believe that not including the environmental criterion in a discrete choice experiment on waste collection, where environmental issues are important, would have resulted in a larger bias.

Innovative criterion

The odds ratios, in table 2, point out that local civil servants across the four countries consider the innovative criterion when awarding contracts to for-profit enterprises. The results show that local civil servants are more likely to award the contract to a for-profit enterprise that offers an app for smart-phones for service users (this app has a calendar, informs on pick-up days and offers a contact tool,…). Belgian, Estonian, German and Norwegian local civil servants are respectively 2.7, 2.9, 3.3 and 2.9 times more inclined to award the contract to a for-profit enterprise that offers an app for smart-phones. Furthermore, as displayed in table 3, a for-profit enterprise that offers an app for smart-phones to the users of the service compared to a for-profit enterprise that does not offer it increases its probability of winning the contract by 45.8

percentage points in Belgium, 48.3 percentage points in Estonia, 53.8 percentage points in Germany and 48.8 percentage points in Norway. These results are statistically significant.

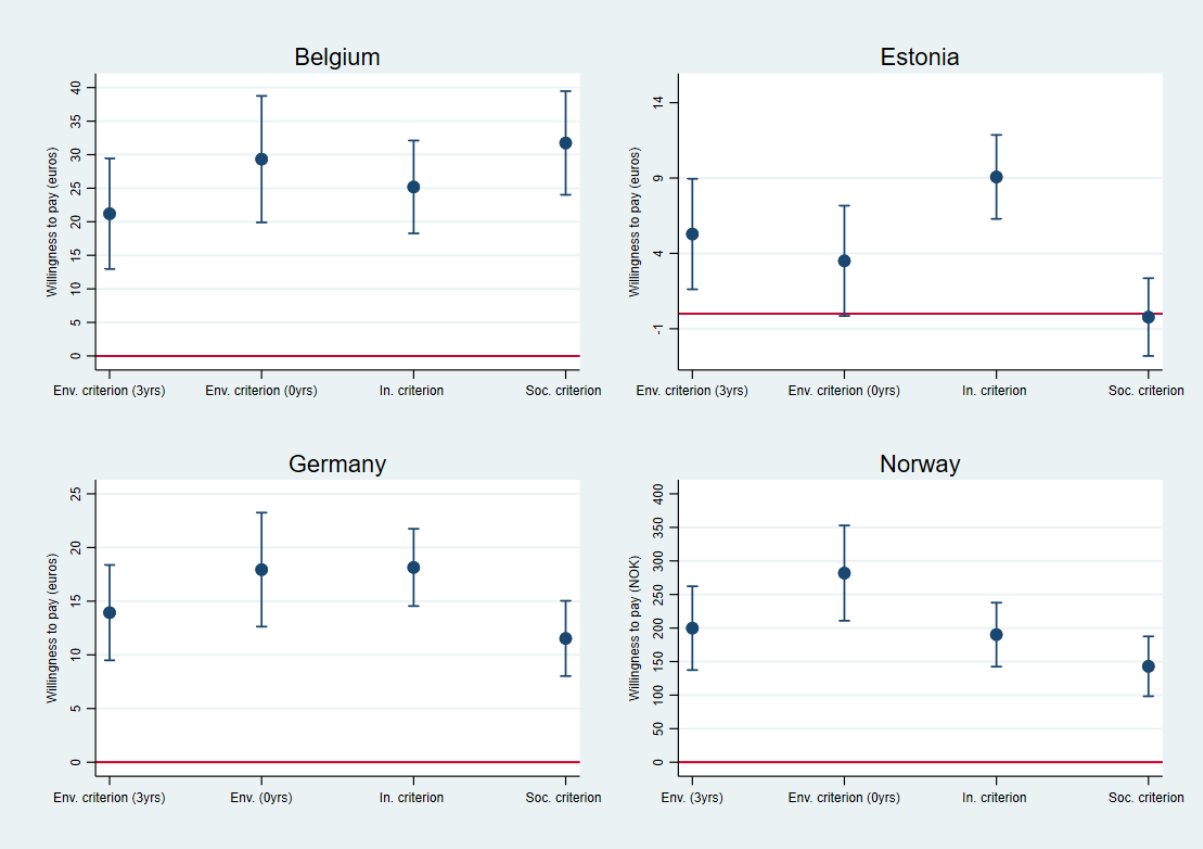
Social criterion

The results, displayed in table 2, show that Belgian, German and Norwegian local civil servants are more inclined to award the contract to a for-profit enterprise that is currently involved in a training scheme for long-term unemployed. The odds ratios indicate that Belgian, German and Norwegian local civil servants are respectively 3.6, 2.1, and 2.2 times more likely to choose a for-profit enterprise that is currently involved in a training scheme for long-term unemployed. These results are statistically significant. Moreover, the probability of winning the contract for a for-profit enterprise raises by 56.2 percentage points in Belgium, 36.4 percentage points in Germany and 38.1 percentage points in Norway if the for-profit enterprise is currently involved in a training scheme for long-term unemployed. These results are statistically significant. It is worth mentioning that Estonian local civil servants do not take the social criterion into consideration when awarding contracts to for-profit enterprises.

Marginal willingness to pay

Figure 1 displays the marginal willingness to pay estimates and the 95 percent confidence intervals for the four countries. With the marginal willingness to pay, we are able to compare the relative importance of every attribute level to local civil servants in every country.

Figure 1. Marginal willingness to pay estimates and 95 percent confidence intervals[9]



The results indicate that local civil servants in the four countries all look beyond the price criterion when awarding contracts to for-profit enterprises. However, the findings also show that local civil servants show different stated preferences regarding the criteria. In Belgium, the social criterion is the most important attribute. Belgian local civil servants are willing to pay 32.4 euros more for a for-profit enterprise that is currently involved in a training scheme for long-term unemployed compared to a service provider that does not include this service. The most important attribute for Estonian local civil servants is the innovative criterion. They are willing to pay 9.1 euros more for a for-profit enterprise that offers an app for smart-phones to the users of the service. In Germany, local civil servants consider the environmental criterion

[9] Calculations of the confidence intervals were performed by the nlcom-command in Stata. This command bases its calculations on the delta method.

as equivalently important as the innovative criterion. For the environmental criterion, they are willing to pay 17.9 euros more for a for-profit enterprise that puts into circulation a new fleet of vehicles (0 years) compared to a for-profit enterprise that puts into circulation a fleet of vehicles of an average age of 6 years. With regards to the innovative criterion, they are willing to pay 18.1 euros more for a for-profit enterprise that offers an app for smart-phones to the users of the service. Norwegian local civil servants also pay greater attention to the environmental criterion and are willing to pay 281.9 Norwegian krone more for a for-profit enterprise that puts into circulation a new fleet of vehicles (0 years) compared to a for-profit enterprise that puts into circulation a fleet of vehicles of an average age of 6 years. It should be reminded that these marginal willingness to pay estimates are not directly comparable to each other.

The abovementioned results demonstrate that local civil servants do not only take the price criterion in consideration when awarding contracts to for-profit enterprises. The analysis points out that Belgian, German and Norwegian local civil servants are willing to pay more to improve the environmental, innovative and social criteria while Estonian local civil servants are willing to pay more to improve the environmental and innovative criteria. By not solely taking the price criterion into consideration, one can argue that local civil servants make a trade off between the environmental, innovative, price and social criteria. Consequently, they intend to maximize the utility of the municipality by awarding the contract to the for-profit enterprise offering the best combination of the criteria

Robustness check

The conditional logistic regression assumes the independence from irrelevant alternatives (IIA). This property, often considered too restrictive, implies that “the unobserved portion of utility for one alternative is unrelated to the unobserved portion of utility for another alternative” (Train 2002, 39). Therefore, to check whether we can draw analogous conclusions

with regards to local civil servants' marginal willingness to pay by relaxing the IIA assumption, we conduct a mixed logit (see table 3 in the appendix) that allows the errors to be correlated to each other.

The results, derived from the mixed logit, indicate that the direction of the statistically significant estimated parameters do not differ from the conditional logistic regression. Furthermore, the marginal willingness to pay estimates (see figure 2 in the appendix), resulting from the mixed logit model, demonstrate that local civil servants' most preferred criterion remains identical to the ones derived from the estimates of the conditional logistic regression. The outcome of the mixed logit model also points out that the marginal willingness to pay estimates are very similar to the ones we derived from the conditional logistic regression. As a consequence, these results suggest that relaxing the IIA property in our analysis does not have a large influence on the marginal willingness to pay estimates.

Discussion and conclusion

According to the doctrine of neo-classical contracting out, public institutions typically award the contract to the external service provider offering, ceteris paribus, the cheapest option. Yet, public institutions recently show a growing interest in taking non-price criteria, such as environmental, innovative and social criteria into consideration when awarding contracts to external service providers. Non-price criteria have increasingly become essential policy instruments for public institutions to pursue policy objectives that are difficult to pursue otherwise. For instance, by taking non-price criteria into consideration, public institutions can increase environmental standards in private companies, develop the employability of vulnerable groups, boost innovation, support start-ups, or favor local or national industries. It has therefore become crucial to examine how public institutions make use of these policy options by studying how they value non-price criteria in comparison with the criterion they

usually look at when awarding contracts: price. Furthermore, researches studying civil servants' behavior in public procurement is relatively scarce (Grandia 2016).

In order to shed more light on this question, we examine, through a discrete choice experiment, local civil servants' valuation of non-price criteria when choosing between alternative public service providers. Following Simon's (1955) call for connecting psychological research with public administration, we analyze local civil servants' valuation of non-price criteria through the lens of random utility theory (Olsen 2015; Grimmelikhuijsen et al. 2016). We assume that local civil servants make a trade off between the environmental, innovative, price and social criteria and, as a result, aim to maximize the utility of the municipality by choosing the for-profit enterprise that offers the best combination of the criteria.

Based on a conditional logistic regression, we first find that the price criterion still plays a role in how local civil servants award contracts to for-profit enterprises. However, in line with the literature highlighting the growing importance of environmental, innovative and social criteria, we observe that local civil servants also consider non-price criteria when awarding contracts. As a matter of fact, including higher levels of environmental, innovative and social criteria into tender documents increases the stated chances of a for-profit enterprises to win the contract. Second, we find evidence that local civil servants make a trade off between the four criteria and intend to maximize the utility of the municipality by awarding the contract to the for-profit enterprise that produces the highest level of utility. Indeed, Belgian, German and Norwegian local civil servants are willing to pay more to improve the environmental, innovative and social criteria while Estonian local civil servants are willing to pay more to improve the environmental and innovative criteria.

The results imply that local civil servants value non-price criteria in public procurement. Yet, the legal framework around public procurement often lacks clarity and its dissemination

to local authorities is scarce, especially when dealing with the relatively new developments around non-price criteria. Providing local civil servants with the necessary resources, such as a clear legal framework and the necessary trainings, to be able to take non-price criteria into consideration when they contract out with external service providers could therefore help them in their decision making.

Our study shows that local civil servants are willing to move from contracts solely awarded based on the price criterion. This however may have several implications. First, the quality of the public service delivered to the population might substantially improve. As a matter of fact, the cheapest public service provider may often be of low quality while public service providers with higher levels of environmental, innovative and social criteria may be of higher quality. Second, non-price criteria, often considered more subjective than the price criterion, should be clearly laid down in the contract before tenders are evaluated to not open the door for abuse and favoritism in the field of public procurement. Third, local civil servants usually have to deliver public services within tight budgets but awarding contracts based on non-price criteria may put pressure on the budget. Public institution have therefore to ensure that, if non-price criteria are considered in tender evaluations, local civil servants have the sufficient budget to prevent the potential additional costs to be passed on the users of the public service. Finally, one might wonder whether demand for non-price criteria from the public administration will meet the private sector's supply of these criteria. Some for-profit public service providers might indeed be reluctant to integrate environmental, innovative and social standards to their enterprises as this can increase their overall production costs.

The research contributes to our understanding of the decision process behind the awarding of contracts at the municipal level, which is a topic that has hitherto received very little empirical attention. By assuming that local civil servants make a trade off between criteria and, therefore, aim at maximizing the utility of the municipality, random utility theory helps us to

have a more comprehensive understanding of local civil servants' stated preferences when awarding contracts. Furthermore, the discrete choice experiment, that is based on random utility theory, is a promising method to understand civil servants' stated behavior.

Although we highlight the importance of non-price criteria in public procurement, there is a need to develop further research on the topic. First, we examine the awarding of contracts to for-profit enterprises in the field of waste collection, and more particularly bulky waste items collection. Civil servants' stated preferences might vary under different public procurement settings. It is therefore essential for further research to analyze civil servants' stated preferences when contracts are awarded in a different service area and to another type of provider. Second, this article is limited to local civil servants' valuation of just three non-price criteria. Other criteria or criteria mixes should be explored further. Finally, we do not know which factors explain variation in the valuation, and whether these are related to civil servants' intrinsic characteristics (ideology, education), or to the institutional and policy setting within which they work.

References

Aizaki, Hideo. 2012. “Basic Functions for Supporting an Implementation of Choice Experiment in R.” *Journal of Statistical Software* 50 (2):1–24. https://doi.org/2115/57786.

Bel, Germà, and Antón Costas. 2006. “Do Public Sector Reforms Get Rusty? Local Privatization in Spain.” *The Journal of Policy Reform* 9 (1):1–24. https://doi.org/10.1080/13841280500513084.

Ben-Akiva, Moshe, and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge: MIT Press.

Brown, Trevor L., and Matthew Potoski. 2005. “Transaction Costs and Contracting: The Practitioner Perspective.” *Public Performance & Management Review* 28 (3):326–51.

Cascetta, Ennio. 2009. “Random Utility Theory.” In *Transportation System Analysis*, 89–167. Boston MA: Springer International Publishing.

Edler, Jakob, and Luke Georghiou. 2007. “Public Procurement and Innovation — Resurrecting the Demand Side.” *Research Policy* 36 (7):949–63. https://doi.org/10.1016/j.respol.2007.03.003.

Edquist, Charles, and Jon Mikel Zabala-iturriagagoitia. 2012. “Public Procurement for Innovation as Mission-Oriented Innovation Policy.” *Research Policy* 41 (10). Elsevier B.V.:1757–69. https://doi.org/10.1016/j.respol.2012.04.022.

Erridge, Andrew, and Sean Hennigan. 2012. “Sustainable Procurement in Health and Social Care in Northern Ireland.” *Public Money and Management* 32 (5):363–70. https://doi.org/10.1080/09540962.2012.703422.

European Commission. Buying Green ! A Handbook on Green Public Procurement. Brussels: The European Union, 2016. Accessed March 10, 2019. https://doi.org/10.2779/246106.

European Commission. Commission notice. Guidance on Innovation Procurement. Brussels: The European Union, 2018. Accessed March 15, 2019. https://ec.europa.eu/digital-single-market/en/news/commission-notice-guidance-innovation-procurement-available-now-all-eu-languages

Furneaux, Craig, and Jo Barraket. 2014. "Purchasing Social Good(s): A Definition and Typology of Social Procurement." *Public Money & Management* 34 (4):265–72. https://doi.org/10.1080/09540962.2014.920199.

Gafni, Amiram. 1998. "Willingness to Pay: What's in a Name?" *PharmacoEconomics* 14 (5):465–70. https://doi.org/10.2165/00019053-199814050-00001.

Grandia, Jolien. 2016. "Finding the Missing Link: Examining the Mediating Role of Sustainable Public Procurement Behaviour." *Journal of Cleaner Production* 124. Elsevier Ltd:183–90. https://doi.org/10.1016/j.jclepro.2016.02.102.

Greszki, Robert, Marco Meyer, and Harald Schoen. 2015. "Exploring the Effects of Removing 'Too Fast' Responses and Respondents From Web Surveys." *Public Opinion Quarterly* 79 (2):471–503. https://doi.org/10.1093/poq/nfu058.

Grimmelikhuijsen, Stephan, Sebastian Jilke, Asmus Leth Olsen, and Lars Tummers. 2016. "Behavioral Public Administration: Combining Insights from Public Administration and Psychology." *Public Administration Review* 77 (1):45–56. https://doi.org/10.1111/puar.12609.

Hauber, A Brett, Juan Marcos González, Catharina G M Groothuis-oudshoorn, Thomas Prior, Deborah A Marshall, Charles Cunningham, Maarten J Ijzerman, and John F P Bridges. 2016. "Statistical Methods for the Analysis of Discrete Choice Experiments : A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force." *Value in Health* 19 (4):300–315. https://doi.org/10.1016/j.jval.2016.04.004.

Hoyos, David. 2010. “The State of the Art of Environmental Valuation with Discrete Choice Experiments.” *Ecological Economics* 69 (8):1595–1603. https://doi.org/10.1016/j.ecolecon.2010.04.011.

Keulemans, Shelena, and Steven Van De Walle. 2017. “Cost-Effectiveness, Domestic Favouritism and Sustainability in Public Procurement: A Comparative Study of Public Preferences.” *International Journal of Public Sector Management* 30 (4):328–41. https://doi.org/10.1108/IJPSM-10-2016-0169.

Kjær, Trine. A Review of the Discrete Choice Experiment - with Emphasis on Its Application in Health Care. Health Economics Papers: 2005. Accessed March 23, 2018. https://www.sdu.dk/~/media/52E4A6B76FF340C3900EB41CAB67D9EA.ashx

Lancsar, Emily, Denzil G. Fiebig, and Arne Risa Hole. 2017. “Discrete Choice Experiments: A Guide to Model Specification, Estimation and Software.” *PharmacoEconomics* 35 (7):697–716. https://doi.org/10.1007/s40273-017-0506-4.

Lancsar, Emily, and Jordan Louviere. 2008. “Conducting Discrete Choice Experiments to Inform Healthcare Decision Making: A User’s Guide.” *PharmacoEconomics* 26 (8):661–77. https://doi.org/10.2165/00019053-200826080-00004.

Loosemore, Martin. 2016. “Social Procurement in UK Construction Projects.” *International Journal of Project Management* 34 (2). Elsevier Ltd and Association for Project Management and the International Project Management Association:133–44. https://doi.org/10.1016/j.ijproman.2015.10.005.

Louviere, Jordan J, Terry N Flynn, and Richard T Carson. 2010. “Discrete Choice Experiments Are Not Conjoint Analysis.” *Journal of Choice Modeling* 3 (3). Elsevier:57–72. https://doi.org/10.1016/S1755-5345(13)70014-9.

Mangham, Lindsay J, Kara Hanson, and Barbara Mcpake. 2009. “How to Do (or Not to Do) . . . Designing a Discrete Choice Experiment for Application in a Low-Income Country.” *Health Policy and Planning* 24 (2):151–58. https://doi.org/10.1093/heapol/czn047.

Manski, Charles F. 1977. “The Structure of Random Utility Models.” *Theory and Decision* 8 (3):229–54. https://doi.org/10.1007/BF00133443.

Mazzucato, Mariana. 2011. *The Entrepreneurial State*. London: Demos.

McCrudden, Christopher. 2004. “Using Public Procurement to Achieve Social Outcomes.” *Natural Resources Forum* 28 (4):257–67. https://doi.org/10.1111/j.1477-8947.2004.00099.x.

McFadden, Daniel. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontiers in Econometrics*. New York: Academic Press.

OECD. Government at a Glance 2017. Paris: OECD, 2017. Accessed January 30, 2019. https://www.oecd-ilibrary.org/governance/government-at-a-glance-2017_gov_glance-2017-en

Olsen, Asmus Leth. 2015. “‘Simon Said,’ We Didn’t Jump.” *Public Administration Review* 75 (2):325–26. https://doi.org/10.1111/puar.12330.

Renda, Andrea, Jacques Pelkmans, Christian Egenhofer, Lorna Schrefler, Giacomo Luchetta, and Can Selçuki. The Uptake of Green Public Procurement in the EU27. Brussels: The European Union, 2012. Accessed July 22, 2018. http://ec.europa.eu/environment/gpp/pdf/CEPS-CoE-GPP%20MAIN%20REPORT.pdf

Rolfstam, Max. 2009. “Public Procurement as an Innovation Policy Tool: The Role of Institutions.” *Science and Public Policy* 36 (5):349–60. https://doi.org/10.3152/030234209X442025.

Ryan, Mandy. 2004. “A Comparison of Stated Preference Methods for Estimating Monetary Values.” *Health Economics* 13 (3):291–96. https://doi.org/10.1002/hec.818.

Ryan, Mandy, Julie R. Kolstad, Peter C. Rockers, and Carmen Dolea. User Guide With Case Studies: *How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas*. Geneva: World Health Organization, 2012. Accessed February 2, 2018. https://www.who.int/hrh/resources/DCE_UserGuide_WEB.pdf?ua=1

Schoute, Martijn, Tjerk Budding, and Raymond Gradus. 2018. “Municipalities ’ Choices of Service Delivery Modes : The Influence of Service, Political, Governance, and Financial Characteristics.” *International Public Management Journal* 21 (4). Taylor & Francis:502–32. https://doi.org/10.1080/10967494.2017.1297337.

Simon, Herbert A. 1955. “A Behavioral Model of Rational Choice.” *The Quarterly Journal of Economics* 69 (1):99–118.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein. 1977. “Behavioural Decision Theory.” *Annual Review of Psychology* 28:1–39. https://doi.org/10.1146/annurev.ps.28.020177.000245.

Svenson, Ola. 1979. “Process Descriptions of Decision Making.” *Organizational Behavior and Human Performance* 23 (1):86–112. https://doi.org/10.1016/0030-5073(79)90048-5.

Temple, Nick, Charlie Wigglesworth, and Chris Smith. Communities Count: The Four Steps to Unlocking Social Value. London: Social Enterprise UK, 2014. Accessed July 21, 2018. https://www.socialenterprise.org.uk/Handlers/Download.ashx?IDMF=e1b5399d-3e59-4724-bc97-e989ce9a3d6a

Testa, Francesco, Fabio Iraldo, Marco Frey, and Tiberio Daddi. 2012. “What Factors Influence

the Uptake of GPP (Green Public Procurement) Practices? New Evidence from an Italian Survey." *Ecological Economics* 82. Elsevier B.V.:88–96. https://doi.org/10.1016/j.ecolecon.2012.07.011.

Thomas, Kyle A, and Scott Clifford. 2017. "Validity and Mechanical Turk: An Assessment of Exclusion Methods and Interactive Experiments." *Computers in Human Behavior* 77. Elsevier Ltd:184–97. https://doi.org/10.1016/j.chb.2017.08.038.

Thomson, Joyce, and Tim Jackson. 2007. "Sustainable Procurement in Practice: Lessons from Local Government." *Journal of Environmental Planning and Management* 50 (3):421–44. https://doi.org/10.1080/09640560701261695.

Thurstone, L.L. 1927. "A Law of Comparative Judgment." *Psychological Review* 34:273–86.

Train, Kenneth E. 2002. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Uyarra, Elvira, Jakob Edler, Javier Garcia-estevez, Luke Georghiou, and Jillian Yeow. 2014. "Barriers to Innovation through Public Procurement : A Supplier Perspective." *Technovation* 34 (10). Elsevier:631–45. https://doi.org/10.1016/j.technovation.2014.04.003.

Walker, Helen, and Stephen Brammer. 2009. "Sustainable Procurement in the United Kingdom Public Sector." *Supply Chain Management: An International Journal* 14 (2):128–37. https://doi.org/10.1108/13598540910941993.

Appendix

Figure 1. Example of a choice set displayed to respondents (English version)

Question 1 out of 6

Which company do you prefer?

	Company 1	Company 2
Company involvement in the professional integration of vulnerable groups	The company is currently involved in a training scheme for long-term unemployed.	The company is not currently involved in a training scheme for long-term unemployed.
The average age of the fleet of vehicles	The average age of the fleet of vehicles is 3 years.	The average age of the fleet of vehicles is 6 years.
Digital services	The company does not offer an app for smart-phones for service users.	The company offers an app for smart-phones for service users (this app has a calendar, informs on pick-up days and offers a contact tool, ...).
Cost per ton of bulky waste items collected	The cost per ton is of 270 EUR.	The cost per ton is of 250 EUR.

Please only tick one box.

- ○ I prefer company 1
- ○ I prefer company 2

Table 1. Descriptive statistics

	Belgium	Estonia	Germany	Norway
Gender				
n	164	46	113	117
Female	52.44%	32.61%	19.47%	23.93%
Male	47.56%	67.39%	80.53%	76.07%
Age				
n	160	47	114	115
Mean	44.94	46.49	54.02	52.37
Standard deviation	9.62	11.27	7.97	8.47
Minimum	24	27	28	29
Maximum	67	70	65	69
Tenure				
n	164	46	114	117
Mean	16.96	14.67	26.04	20.52
Standard deviation	9.76	9.10	10.71	11.33
Minimum	0	0	1	0
Maximum	39	48	48	44
Education				
n	164	47	97	118
Lower secondary	0.61%	/	/	/
Upper secondary	6.10%	4.26%	/	1.69%
Post-secondary non tertiary	0.61%	4.26%	/	5.93%
Short-cycle tertiary	/	/	/	11.02%
Bachelor or equivalent	39.02%	36.17%	40.21%	34.75%
Master or equivalent	51.22%	55.32%	59.79%	46.61%
Doctoral or equivalent	2.44%	/	/	/

Table 2. Conditional logistic regression (estimates and standard errors into parentheses)

	Belgium	**Estonia**	**Germany**	**Norway**
Price criterion	-0.039***	-0.116***	-0.066***	-0.006***
	(0.005)	(0.017)	(0.006)	(0.0006)
Environmental criterion (3years) Ref. Cat.: 6 years	0.825***	0.614**	0.923***	1.122***
	(0.121)	(0.216)	(0.147)	(0.152)
Environmental criterion (0yrs) Ref. Cat.: 6 years	1.144***	0.407	1.188***	1.583***
	(0.135)	(0.217)	(0.168)	(0.174)
Innovative criterion Ref. Cat.: no app	0.989***	1.054***	1.201***	1.068***
	(0.094)	(0.165)	(0.116)	(0.116)
Social criterion Ref. Cat.: no training scheme	1.272***	-0.026	0.764***	0.802***
	(0.095)	(0.153)	(0.105)	(0.108)
ASC	0.132	-0.008	-0.068	0.116
	(0.091)	(0.147)	(0.107)	(0.109)
Number of observations	1992	564	1500	1416
Number of local civil servants	166	47	125	118
LR Chi² (6)	538.71	109.11	437.95	408.72
Prob > Chi²	0.0000	0.0000	0.0000	0.0000
McFadden pseudo R^2	0.390	0.279	0.421	0.416

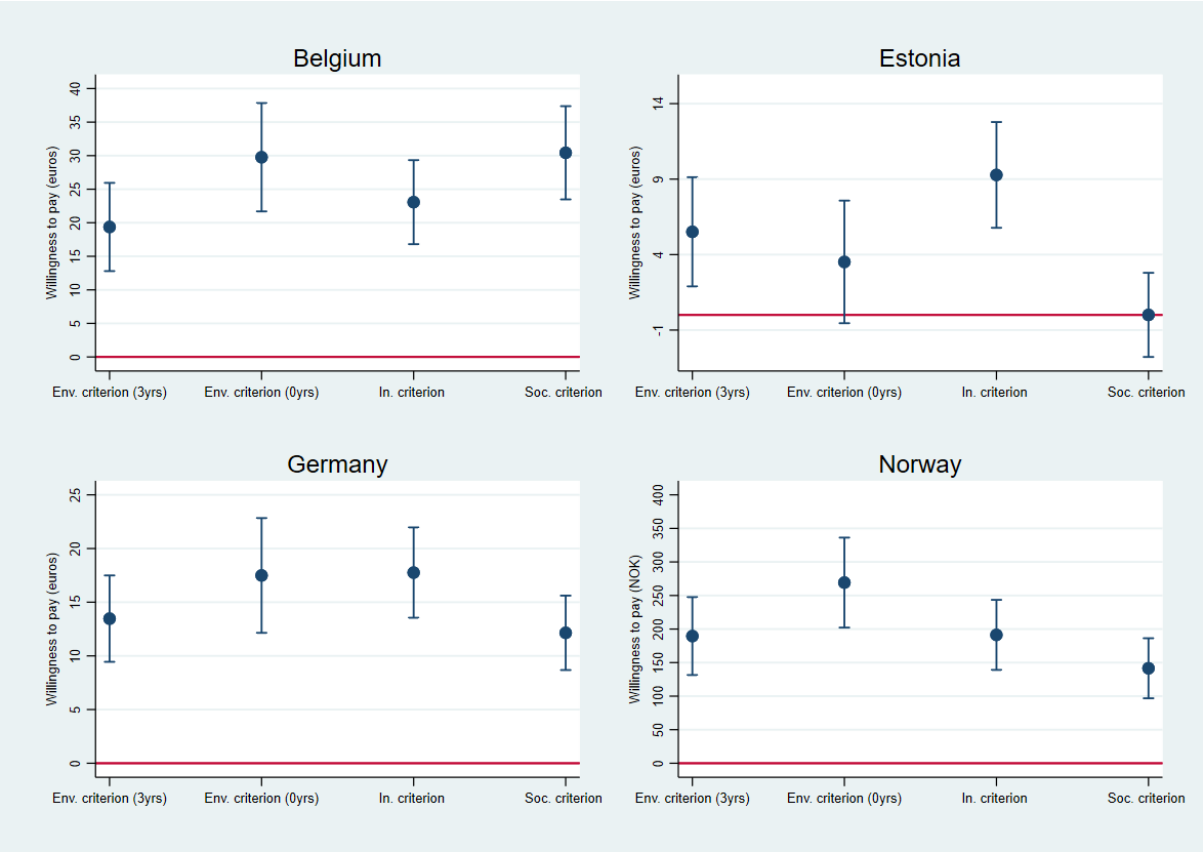
*** p < 0.001, ** p < 0.01, * p < 0.05

Table 3. Robustness check – Mixed logit model with 500 Halton draws (estimates and standard errors into parentheses)

	Belgium	Estonia	Germany	Norway
Price criterion	-0.126***	-0.168***	-0.130***	-0.010***
	(0.031)	(0.034)	(0.027)	(0.002)
Environmental criterion (3years) Ref. Cat.: 6 years	2.432***	0.928**	1.757***	1.986***
	(0.606)	(0.345)	(0.414)	(0.435)
Environmental criterion (0yrs) Ref. Cat.: 6 years	3.737***	0.592	2.281***	2.821***
	(0.962)	(0.366)	(0.557)	(0.601)
Innovative criterion Ref. Cat.: no app	2.896***	1.562***	2.316***	2.004***
	(0.686)	(0.357)	(0.486)	(0.456)
Social criterion Ref. Cat.: no training scheme	3.820***	-0.0009	1.584***	1.483***
	(0.935)	(0.239)	(0.393)	(0.356)
ASC	0.493	0.089	-0.189	0.375
	(0.296)	(0.224)	(0.195)	(0.224)
Number of observations	1992	564	1500	1416
Number of local civil servants	166	47	125	118
LR Chi² (4)	91.89	11.41	36.76	26.21
Prob > Chi²	0.0000	0.0223	0.0000	0.0000

*** p < 0.001, ** p < 0.01, * p < 0.05

Figure 2. Robustness check – Marginal willingness to pay estimates and 95 percent confidence intervals[10] derived from the mixed logit model



[10] Calculations of the confidence intervals were performed by the nlcom-command in Stata. This command bases its calculations on the delta method.