

Some Award Methods are More Prone to Problems than Others

An Empirical Analysis of the Relationship Between Award Methods in the Economically Most Advantageous Tender and Problems in Dutch Public Procurement.

Master thesis – Public administration (Specialization Public Management)

UNIVERSITEIT TWENTE.



First supervisor

Prof. Dr. Jan Telgen

Second supervisor

Dr. Harry van der Kaap

Robbert Hoogeveen

September, 2014

kennisplatform
CROW

First supervisor

Ad van Leest (A&C)

Second supervisors

Paul van Bruggen (A&C) and
Xylander Kroon (ABK)

Summary

Recent theoretical developments suggest that some award methods are more susceptible to problems than others. This proposition is important to public organizations in the application of award methods in Dutch public procurements. In order to contribute to the current process and practices within the award phase, we have selected ten frequently occurring practices (hereafter named 'elements') based on theoretical grounds and sought to identify whether the presence or exclusion of these elements is significantly related to problems. We hypothesize that these ten elements are more present in problematic procurements than non-problematic procurements. Therefore, this research is focused on the following main question:

To what extent do award methods (in the award phase of Dutch public tender between May 2010 and April 2014) include the 10 selected elements, and to what extent do these elements result in 'problems'?

The rationale for this research stems from the available leeway that contracting authorities have to implement the award phase. In current public procurement law, few regulations refer to the selection, use and application of award methods. A consequence of this leeway is that contracting authorities use various decision-making criteria and combine, weigh, score and measure criteria differently. As a result, many contracting authorities struggle with the pressure to explain their choices and therefore rely on some form of formal method, while the effects of these methods are not always known. For instance, certain methods or elements of these methods can result in undesirable, unpredictable or unforeseen outcomes.

In this study we sought to identify which of these elements are related to the outcome problems. Since there are many possible aspects, elements and choices in the award phase, this study aimed at explaining problems from the occurrence of the ten elements. The latter is examined by means of an empirical study of 445 Dutch public procurements. More specifically, a *stratified random sample* consisting of 402 non-problematic procurements and a sample of 43 problematic procurements. These two samples are compared, in the presence of the ten elements and their outcome problems. In this study the dependent variable is 'problems', more specifically 'problems resulting from the award method'. Problems are measured with two possibilities that suppliers have to oppose to the used award method, they are serious actions and indicate severe problems with regard to the award method; (1) procurements that resulted in a lawsuit and (2) in questions that were addressed by the Commission of Procurement Experts.

The results indicate that six elements are significantly more present in problematic procurements than non-problematic procurements. These elements concern the presence of the (1) Weighted Factor Score, (2) relative scoring rules, (3) 'flat' award criteria, (4) sub-criteria for award criterion price and the absence of (5) weights and (6) scoring rules. These six elements have an increased likelihood of problems. With the use or exclusion of these elements (i.e. negative impact) the chance of problems is twice to four times higher, varying in their susceptibility to problems from 21% to 60%. However, this does not imply that there is a causal relationship between these elements and problems in public procurement. These elements are often applied in public procurement and many times these do not result in problems. Therefore, these elements are more prone to problems than their counterparts. Moreover, two outliers were indicated, not indicating weights has a probability of

60% to result in problems and not indicating the score method 32%. In case these two elements are not indicated in request for quotations, the outcome problems is expected. In addition, not indicating these elements is a violation of public procurement regulations.

The other four elements are not significantly present in problematic procurements. Though, the presence of (7) lower boundaries is directed towards problems. This implies that lower boundaries are more present in problematic procurements, only not significantly. The other three elements; (8) type of performance levels, (9) form of the score graph and (10) scoring on the basis of ranking are not directed towards problems and consequently not significantly present in problematic procurements. Important to note; theoretically these four elements are more susceptible to problems.

Overall, award phases which lack weights or scoring rules are like to result in problems. Therefore, it is recommended to practitioners to apply award methods in such a way, that all parameters of the awarding mechanism are published in request for quotations. Furthermore, the Weighted Factor Score, relative scoring rules, 'flat' award criteria, sub-criteria for award criterion price and lower boundaries are susceptible to problems. Therefore, the intended relative impact of award criteria (price and quality) should not be affected by certain types of scoring rules or rating scales. Practitioners can withhold problematic elements, or apply these elements in non-problematic ways. This can be achieved by means of a transparent award method, which includes all parameters of the awarding mechanism.

Managementsamenvatting

Recente theoretische ontwikkelingen suggereren dat bepaalde gunningsmethodieken gevoeliger zijn voor problemen dan anderen. Deze stelling is belangrijk voor aanbestedende diensten in de toepassing bij het beoordelen van offertes van het gunningscriterium Economisch Meest Voordelige Inschrijving. Om bij te dragen aan de voorgaande stelling, hebben wij tien veelvoorkomende elementen geselecteerd op basis van hun theoretisch negatieve effect. Vervolgens hebben wij getracht te bepalen of de aanwezigheid of uitsluiting (een negatieve impact) van de elementen significant is gerelateerd aan problemen. Wij stellen dat de geselecteerde elementen meer aanwezig zijn in problematische aanbestedingen dan niet-problematische aanbestedingen. Het onderzoek is gericht op de beantwoording van de volgende hoofdvraag:

In welke mate bevatten gunningsmethodieken (in de gunningsfase van Nederlandse aanbestedingen tussen mei 2010 and april 2014) de 10 geselecteerde elementen, en in welke mate leiden deze elementen tot ‘problemen’?

De aanleiding voor dit onderzoek komt voort uit de beschikbare speelruimte die aanbestedende diensten hebben omtrent de uitvoering van gunningsfase. De huidige wetgeving bevat slechts enkele algemene bepalingen over het gebruik en toepassing van gunningscriteria in de gunningsfase. Een gevolg van deze speelruimte is dat aanbestedende diensten gebruik maken van verschillende gunningscriteria, deze anders combineren, wegen, scoren en meten. Dit resulteert in een vrijheid, waar veel aanbestedende diensten worstelen met de druk om hun keuzes te verantwoorden. Daarom geeft de aanbestedende dienst vaak invulling aan deze vrijheid door zich te beroepen op een bepaalde formele gunningsmethodiek, terwijl de effecten van deze methoden niet altijd bekend zijn. Zo kunnen bepaalde werkwijzen of elementen van deze werkwijzen resulteren in onvoorziene, onvoorspelbare of ongewenste resultaten.

In deze studie hebben we getracht te onderzoeken welke gunningsmethodieken of elementen uit de gunningsmethodieken zijn verbonden aan de uitkomst problemen. Aangezien er vele mogelijke aspecten, keuzes en elementen in de gunningsfase zijn, is deze studie gericht op het verklaren van problemen door middel van de aanwezigheid van tien geselecteerde elementen. Dit wordt onderzocht door middel van een empirisch onderzoek van gunningsfase van 445 Nederlandse overheidsopdrachten. Specifiek, een steekproef bestaande uit 402 niet-problematische aanbestedingen en een groep van 43 problematische aanbestedingen. Deze twee groepen worden met elkaar vergeleken in de aanwezigheid van de tien elementen en hun resultaat ‘problemen’. De afhankelijke variabele in dit onderzoek is ‘problemen’, specifiek ‘problemen met betrekking tot de gunningsmethodiek’. Problemen worden gemeten aan de hand van twee mogelijkheden die leveranciers hebben om zich te verzetten tegen de gehanteerde gunningprocedure. Beide acties zijn serieuze maatregelen en geven ernstige problemen weer met betrekking tot de gunningprocedure. Het zij (1) aanbestedingen die hebben geleid tot een rechtszaak of (2) tot vragen/klachten die in behandeling zijn genomen door de Commissie van Aanbestedingsexperts.

De resultaten geven aan dat zes elementen significant meer aanwezig zijn in problematische aanbestedingen. Deze elementen betreffen (1) de gewogen factor score (puntenmethode), (2) relatieve scores, (3) vlakke beoordeling, (4) subcriteria voor het gunningscriterium prijs (5) het niet vermelden van de gewichten en tot slot (6) het niet vermelden van scoremethodiek.

Deze zes elementen hebben een verhoogde kans op problemen. Bij het gebruik of uitsluiting van deze elementen (negatieve impact, bijvoorbeeld; het weglaten van gewichten en het toepassen van relatieve scores) is de kans op problemen twee tot viermaal hoger. Dit varieert per element, tussen 21% en 60% kans op problemen. Echter, dit betekent niet dat er een causaal verband tussen deze elementen en problemen is. Deze elementen worden vaak toegepast in aanbestedingsprocedures en vele malen leiden deze niet tot problemen (althans, meetbaar). Daarom kunnen wij alleen concluderen dat deze zes elementen gevoeliger zijn voor problemen dan hun tegenhangers. De resultaten geven wel twee uitschieters aan; het niet vermelden van het gewicht in de gunningsmethodiek heeft een kans van 60% op problemen. Dit is 32% voor het niet vermelden van de scoremethodiek. In geval deze elementen niet zijn vermeld in aanbestedingsdocumenten, wordt de uitkomst problemen verwacht. Daarnaast is het niet aangegeven van deze elementen in strijd met de wet (transparantiebeginsel).

De andere vier elementen zijn niet significant aanwezig in problematische aanbestedingen. Hoewel de aanwezigheid van (7) extra ondergrenzen wel meer aanwezig is in problematische aanbestedingen. De andere drie elementen; (8) prestatietype gunningcriterium, (9) bijzondere vorm van de scoregrafiek en (10) scores op basis rangorde zijn niet anders aanwezig in problematische aanbestedingen ten opzichte van niet-problematische aanbestedingen. Het is belangrijk om op te merken dat theoretisch gezien al deze elementen vatbaar zijn voor problemen.

Samengevat, gunningsmethodieken die het nalaten gewichten of scoremethodieken te vermelden leiden tot problemen. Het is daarom aanbevolen om alle parameters van een beoordelingsmechanisme te publiceren in de aanbestedingsdocumenten. Naast het niet vermelden van gewichten of scoremethodieken, zijn de gewogen factor score, relatieve scores, vlakke beoordelingen en subcriteria voor het gunningscriterium prijsgevoelig voor problemen. Het is aanbevolen om de beoogde impact van elementen te waarborgen in de gunningsmethodiek. Het is van belang dat dit niet wordt aangepast door bepaalde scoremechanismes of beoordelingsschalen. Tot slot, beoefenaars kunnen de problematische elementen onthouden of deze elementen in niet-problematische manieren toepassen. Dit kan worden bereikt door middel van een transparante en objectieve gunningsmethodiek, waarbij alle parameters van het beoordelingsmechanisme van tevoren bekend zijn gemaakt.

Preface

At the end of this one-year master thesis, which I started in 2013, I complete my master's degree at the University of Twente. I have enjoyed all courses within the master Public Administration, especially those around public management and an essential focus of this specialization; public performance. This is also the reason why I approached my first supervisor Professor Jan Telgen. After broadly and enthusiastically sharing his insights in topics and research, it soon became clear that the research subject would be public procurement (i.e. government supplier selection). A topic that is well aligned with my courses and my interests.

After discussing my preferences and research possibilities in public procurement, contact was sought with the organization CROW (the technology platform for transport, infrastructure and public space). This organization works together with government and businesses with regard to public procurement. CROW is active in research and in issuing regulations, focusing on all stakeholders in public procurement. This made CROW the perfect organization for an graduate internship. The organization has provided me with a license to their database "aanbestedingskalender.nl" which enabled me to execute a quantitative research with respect to elements in the award phases of Dutch public tenders.

First of all, I would like to thank my first supervisor Professor Jan Telgen for his guidance, support and constructive criticism during the last year. My thanks also go to my second supervisor Dr. Harry van der Kaap for his expertise and efforts in analyzing data.

Second, I would like to thank the organization CROW, both teams 'Aanbesteden & Contracteren' and 'Aanbestedingskalender.nl' in which I have spend many weeks. Especially, my gratitude goes to my three supervisors, Ad van Leest, Paul van Bruggen and Xylander Kroon for their time, feedback and insights in practice. The knowledge that I have gained at CROW during my graduation internship, extends far beyond the scope of this thesis.

Furthermore, I want to thank my friends and family which have supported me a lot during all the months I wrote my thesis. I appreciate all the times we discussed the progress of my thesis during 'dinner table sessions'.

In particular and most of all, my gratitude goes to my grandfather, who is no longer among us. Without him I would not be studying at the University at the first place. He stimulated, encouraged and promoted studying and taught me that perseverance will be rewarded.

Robbert Hoogeveen

September, 2014

Inhoud

Summary	2
Managementsamenvatting	4
Preface.....	6
1. Introduction.....	9
1.1 Research motivation.....	9
1.2 Scientific and practical relevance	11
1.3 Research questions.....	11
1.4 Structure.....	12
2. Research Approach.....	13
2.1 The setting: Dutch public procurement	13
2.2 Quantitative data (population and sample)	16
2.3.1 Data sources	19
2.3 Measures	20
2.3.1 Measuring ‘problems’ in procurements (dependent variable)	20
2.3.2 Measuring 10 elements in public procurements (independent variable).....	20
2.4 Data analysis.....	21
2.5 Research Method	21
2.6 Conclusion	22
3. Elements in the award phase	23
3.1 Characteristics in the award phase	23
3.1.1 Element 1 – Supplier selection methods (Award criteria: ‘lowest price’ or ‘EMAT’)	25
3.1.2 Element 2 – Award criteria (Sub-criteria).....	27
3.1.3 Element 3 – Publication of overall weights	27
3.1.4 Element 4 – Score method not published in tender documents	28
3.1.5 Element 5 – Form of the score graph (price)	28
3.1.6 Element 6 – Award criterion contains relative scores.....	29
3.1.7 Element 7 – Scores for an award criterion are awarded on the basis of ranking.	30
3.1.8 Element 8 – Flat award criteria	30
3.1.9 Element 9 – Extra lower boundaries	30
3.1.10 Element 10 – Award criterion price consist of various financial sub-criteria.....	31
3.2 Conclusion	31
4. Problems (Lawsuits and Questions addressed by the Committee)	32
4.1 Commission of Procurement Experts.....	32

4.2	Lawsuits	33
4.3	Conclusion	34
5.	Results	35
5.1	Presence of elements	35
5.2	The number of jointly applied elements	35
5.3	Hypothesis	37
5.4	Conclusion	43
6.	Conclusion and discussion.....	45
6.1	Conclusion	45
6.2	Recommendations and implications for practitioners.....	48
6.3	Limitations	48
6.4	Further (empirical) research.....	49
	References.....	51
	Appendix A: The setting - TED Europe descriptive information.....	53
	Appendix B: Population – All announcements ABK	55
	Appendix C: Sample Group 1 – non-problems	56
	Appendix D: Sample Group 2 – problems	57
	Appendix E: Sample Group 2 ‘problems’ – bias check removal of cases with insufficient information.....	60
	Appendix F: Research format	61
	Appendix G: Population and Sample (bias check).....	63
	Appendix H: Data analysis – 10 hypothesis.....	67

1. Introduction

In public procurement it is obligatory to buy via public tendering, a procurement method that aims at enabling market access in a transparent and objective manner. In the procurement procedure a request for quotation (invitation to bid, hereafter RfQ) is placed in the market. Contractors, suppliers or service providers who would like to obtain the contract are asked to place a detailed bid. The procurement procedure is typically divided into several phases, one of these phases is the award phase (subject of this research). In the award phase, contracting authorities or executing agencies (public bodies or public institutions/entities) determine the bid that is best aligned with the intended goals and requirements in the RfQ. Contracting authorities then select the best supplier, based on the formulated decision criteria and scoring methods. These parameters are usually published in the tender documents attached to the RfQ.

Important to note; the award phase is preceded by a selection process. In procurement procedures, selection criteria are used to identify suppliers which are deemed capable of performing tasks, and are therefore related to the quality of the supplier. The large number of available suppliers is hence reduced to a short list. In this research, the award phase of a public tender is the unit of analysis.

1.1 Research motivation

The rationale for this research stems from the available leeway of contracting authorities to give their own interpretation to the award phase. This can result in choices which may be different from government intentions. Within European directives and Dutch public procurement law, two obligatory rules are included to enhance five principles; non-discrimination, equal treatment, mutual recognition, proportionality and transparency. First, contracting authorities are obliged to make use of the Economically Most Advantageous Tender (EMAT) unless a good motivation is given to award with the lowest price criterion (Art. 2.114 Aanbestedingswet 2012). Second, if EMAT is used, it is obligatory to publish further award criteria and their relative weights (Art. 53 of the European Directive 2004/18/EG, Art. 2.115 Aanbestedingswet 2012).

According to the European Directives and the Dutch public procurement Act, the choice of contract award criteria is consequently twofold; the EMAT and the lowest-price (art. 53 of the European Directive 2004/18/EG; Art. 2.115. Aanbestedingswet 2012). Lowest-price tender evaluation is in principle straight-for-ward, since only a single criterion (price) is of importance in the award phase. However, awarding on the lowest-price criterion does not immediately lead to the best price-performance ratio. It has been argued that lowest-price scoring rule created an atmosphere in which suppliers reduce their performance to the minimum requirements in order to be able to offer a low price (Bergman and Lundberg, 2013). Therefore, the award criterion EMAT, which uses both price and quality in supplier selection, can enhance the efficiency of public procurement. However, using EMAT adds complexity to the procedure (Bergman and Lundberg, 2013).

In public procurement law, just a single regulation is listed on selection of further award criteria. This regulation states that award criteria need to be linked to the subject-matter of the contract (European Directive art. 53 2004/18/EG). Further, no specific regulations are listed on selection of award criteria within the EMAT criterion. This means that the contracting authority has the right to determine which further decision-making criteria are applicable to a public tender. This also applies to the weights, measurement and scoring of the separate further decision-making criteria.

In practice, choosing further award criteria and their weights are hardly ever overlooked since publication of these elements is mandatory, even though they might be given too little serious attention. However, relating criteria to each other and selecting scoring methods for criteria are not always acknowledged in practice and are consequently often neglected. (De Boer et al., 2006). In short, contracting authorities are largely free determining to what extent and in which way award criteria and award methods are applied and constructed in the award phase. A consequence of this leeway is that contracting authorities use various further decision-making criteria and combine, weigh, score and measure criteria differently.

As a result, many contracting authorities struggle with the pressure to explain their choices, especially when trying to identify the EMAT. Therefore contracting authorities rely on some type of formal method (De Boer et al., 2006). However, the application of formal decision methods is not without problems and is vulnerable to misuse. This can be explained by the fact that there are many aspects playing a role in supplier selection. Additionally, many decision criteria and methods can be considered while the effects of these aspects and methods are often unknown (Telgen and Schotanus, 2010). Moreover, many decision makers and stakeholders with different perspectives can or must be involved, which further complicate the decision-making process (De Boer et al., 2006). Overall, various formal decision methods, decision elements, and various quantitative and qualitative decision criteria for supplier selection can be taken into account (e.g. De Boer et al., 1998, 2001; Ho et al., 2010), resulting in lots of options and hence choices to be made.

Therefore, the aim of this research is to gain a deeper understanding of the process and practices within the available leeway of the award phase. We seek to identify which of the involved choices cause problems regarding the award method. For example, problems can arise when bidding suppliers are confronted with award methods lacking transparency and/or objectivity in determining the winner, such as, not indicating weights or (using) certain scoring rules. Since there are many aspects and choices in the award phase, it would be impossible to test the effects of all of them. Therefore, 10 frequently occurring elements (options and choices in constructing the award method) are selected based on a recent publication, 'Gunningsmethodieken voor Uitzendwerk' of the Public Procurement Research Centre (PPRC). The publication of the PPRC provides theoretical expectations on the selected elements, these elements are subject to leeway in the award phase and are expected to be problematic (extensively discussed in chapter 3). In this empirical study, it is determined whether the presence or exclusion (i.e. negative impact) of these 10 elements is related to problems (Figure 1 presents an overview).

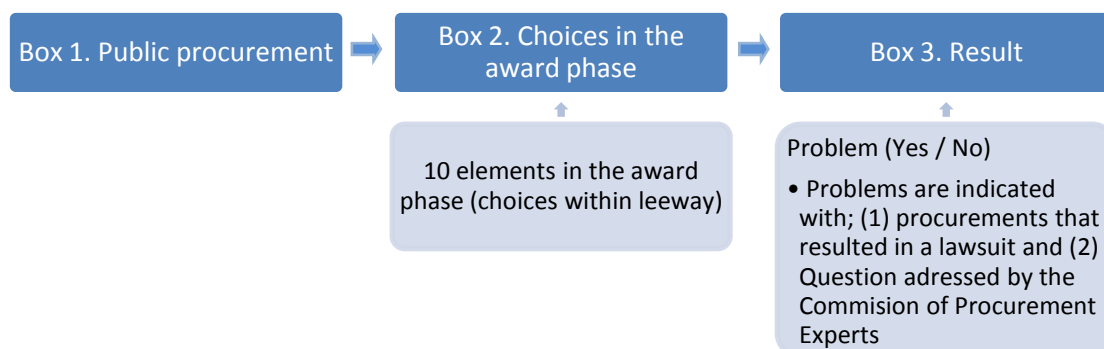


Figure 1: Overview research

As presented in Figure 1, the second box 'Choices in the award phase', addresses 10 elements which serve as independent variables (further discussed in chapter 3). Box number three 'Results' addresses the dependent variables (extensively discussed in chapter 4). Here, the dependent variables are 'problems in the award phase relating to the award method'. Problems become evident by the consequent actions of the bidding suppliers; opposing to the awarding mechanism of the contracting authorities. There are two possibilities for suppliers to oppose; (1) procurements that resulted in a lawsuit (district court, in Dutch: Kort geding bij de voorzieningenrechter) and (2) procurements involving questions that are addressed by the Commission of Procurement Experts. Both options are serious actions and indicate severe problems.

This research aims to explain the occurrence of problems, applying the negative impact of 10 elements in the award phase by means of an empirical study of the award method of Dutch public procurements. The main hypothesis is that we expect problems to occur when single or more of these 10 elements are present. *Is this true? Can evidence be provided that supports this expectation?* To examine the latter, this empirical study addresses a random selection of Dutch public procurements and a broad range of Dutch public procurements that resulted in problems. Furthermore, given the fact that inferences are drawn about the effect of different elements on problems, each procurement is analyzed at one specific point in time, making this study cross-sectional.

1.2 Scientific and practical relevance

This study makes a novel contribution to science in two ways. First, empirical information is provided about the relationship between 10 elements in the award phase and their effect on causing problems. The specific goal is to improve future procurement practices in the award phase. This explanatory information serves scientific purposes, as well as policy makers (current procurement policy and future policy development) and contracting authorities. Secondly, besides the examined relationships, public tenders are analyzed on award methods and criteria, this provides descriptive information. This information is particularly useful for CROW (the national platform for infrastructure, traffic, transport and public spaces) and Aanbestedingskalender.nl (ABK – a national platform for the publications of RfQ), as it regards parameters of the awarding mechanism.

1.3 Research questions

To improve future procurement practices in the award phase, the following main research question is formulated:

To what extent do award methods (in the award phase of Dutch public tender between May 2010 and April 2014) include the 10 selected elements, and to what extent do these elements result in 'problems'?

To answer the research question properly, this research is built on a number of sub-questions which provide structure to the main research question:

1. Which 10 elements are selected on theoretical grounds that have a negative impact in the award phase and are expected to result in problems?
2. Are there differences in the degree to which the 10 elements are present in non-problematic and problematic procurements?

3. Is there a causal relationship between the identified elements in the award phase and problems in public procurement?

1.4 Structure

The research is structured as followed; chapter two explains the research design, data collection and data analysis. In chapter three, the theoretical expectations are discussed and the 10 elements in the award phase are described; each element is discussed in terms of its negative impact on the result of public tenders. In chapter four, the outcome of public tenders are discussed based on the two options suppliers have to oppose to awarding by contracting authorities. In chapter five, the data regarding award mechanisms of 445 Dutch public tenders is discussed and each hypothesis is statistically analyzed. Finally, in chapter six conclusions are drawn, the practical implications for policy makers and practitioners are discussed. In addition, it provides suggestions for future (empirical) research.

2. Research Approach

This chapter outlines the research approach and addresses the central questions of causation related to the comparison of elements across both groups of procurements. First, the research setting, research population and available data are discussed. The focus lies on the way this research was carried out, and in particular why certain choices were made. Second, the measures of the 10 identified elements (independent variables) and measures of problems (dependent variables) are made measurable by operationalizing them. Subsequently, the data analysis is described. This section describes how the collected data is analyzed by comparing it to formulated hypotheses. Finally, the research method is described: the use of qualitative and quantitative research methods is explained regarding their use in this research.

2.1 The setting: Dutch public procurement

In this research award phases of Dutch public procurements between May 2010 and April 2014 – are studied. The population in this research consist of solely Dutch tenders which are published in the database Aanbestedingskalender.nl (ABK)¹ within these four years. There are several reasons for selecting this time period. First, to provide an accurate and complete overview of the current procurement practices that occurred most recently. Second, to control for fluctuations and variations in public procurements, more specific, in the elements used in the award phase. Third, to control for the new Dutch Public Procurement Act, which entered into force on 1 April 2013. And last, to have a sufficient sample of procurements that resulted in ‘problems’ (lawsuits and questions addresses by the Commission of Procurement Experts – regarding the award method to enable proper statistical analysis). The sample of procurements that resulted in problems is covers the same time period.

The sampled population is divided in three subgroups; organizations, procedures and contract subject. These groups are defined due to their expected influence on the applied award method. For example, various contracting authorities (and decision makers and stakeholders within contracting authorities) may have different perspectives on applied award methods in the award phase, which, in turn, leads to different use of award methods in practice. The type of procurement procedure is also expected to affect award methods. Dependent on the type of procurement procedure, guidelines and rules are included in European Directives, the Dutch Public Procurement Act and the Guide of Proportionality. These rules and guidelines may have impact on the suitability of the selected award method and hence decision. At last, the contract subject may have impact on the selected award method. For example, the organization CROW states that the award method “Awarding on Value” is common for works, however, this may be different for goods and services. Overall, these three subgroups are considered to account for the possible confounding effect of organization, procedure and contract subject (i.e. all three subgroups are taken into account to control for bias). The subgroup ‘contract subject’ has an additional purpose, namely; all formulated hypothesis are analyzed for both works, goods and services. In other words, the possibility exists that certain elements are solely causing ‘problems’ in any of these contract subjects.

¹ ABK – aanbestedingskalender.nl (In this database all RfQ’s of published procurements are present. Important to note; not all public procurements are published, for example; contracting authorities are not obligated to publish negotiated tenders and single calls for proposals)

1. Organizations

The organizational context of this analysis covers Dutch contracting authorities, public bodies or public institutions. These are multipurpose authorities providing education, social care, regulatory services (such as land use planning and waste management), housing, welfare benefits, recreation, and cultural services. In this study contracting authorities are divided into seven categories (contracting authorities – Art. 1.1 Aanbestedingswet 2012);

- The state (national government);
- Provinces;
- Municipalities;
- Water boards;
- Public entities;
- Collaborations of the previous governments or public entities;
- And other: such as education and port companies.

However, the exact number in the population with regard to organization types, cannot be provided, as the database ABK does not have the option to analyze these details from publications. Therefore, the exact number in the population remains unclear and cannot be directly linked to the sample. By drawing a random sample is from the population, it is attempted to obtain an equal distribution of contracting authorities used for this analysis. To correct for bias, the database Tenders Electronic Daily (TED)² is used to indicate the distribution of contracting authorities. Here it has to be noted that this information is only available for European public procurements; the precise distribution in national procedures remains unidentified. In Appendix A (Table I) an overview is presented with regard to the exact number of contracting authorities in European procurements within the research time frame. The Appendix includes percentages to enable comparison of the research population with the research sample (further discussed in paragraph 2.2).

2. Procurement procedures

In Dutch public procurement, contracting authorities deal with different procurement types. The European Directives state that tenders whose monetary value exceeds a certain amount, are of cross-border interest. Each procurement that exceeds this threshold needs to be submitted in the European Union (Art. 7 of the European Directive 2004/18/EG, and European Directive 2004/17/EG). For procurements under this threshold, the European Directives and Dutch Procurement Act do not impose selection rules regarding procurement procedures, although, guidelines concerning the selection and use of procedures are included in the Guide of Proportionality.³ The following types are included in the population:

National procedures:

1. Open (unrestricted) procedure
2. Restricted procedure
3. Negotiated procedure (publication not mandatory)
4. Single call for proposals (publication not mandatory)

European procedures:

5. Open (unrestricted) procedure
6. Restricted procedure

² Ted.europa.eu (The online version of the 'Supplement to the Official Journal of the European Union', dedicated to European public procurement.

³ Gids Proportionaliteit (the guide of proportionality provides guidelines for the selection of a procurement process/type. This guide is designated as mandatory guideline to follow.

To analyze the award phase of Dutch public procurements it is necessary that all relevant tender documents have been published, which is not always the case. In restricted procedures, negotiated procedures and single calls for proposals, the relevant documents are not always freely accessible on ABK. This means that only the selected suppliers receive the award documents. In addition, contracting authorities are not obligated to publish RfQ's in case of negotiated procedures and single call for proposals. Therefore, the sample contains predominantly national open and European open procedures. The accurate number of all procurement procedures types is unidentified, as the database (ABK) is unable to distill this information from RfQ's. However, it is possible to make a general distinction between national and European procedures. In the database TED all procedure types can be deduced from publications, however, as discussed previously, this only applies to European procurements). In Table 1, this limited overview is presented of both national and European procedures.

Procurement procedures (numbers in the research population)			
Procedure type	All announcements	RfQ's	Award decisions
European procedures ABK	41.748	X	16.682
National procedures ABK	25.484	X	1.845
Total ABK	67.232	X	18.525
European procedures TED	37.340	15.644	14.861

Table 1: Overview procurement procedures in the research population (NL: May 2010 – April 2014).

The procedure types in Table 1 distinguishes between national procedures and European procedures. The type of publication is divided in three types: all announcements, RfQ's and award decisions. In the option 'all announcements' overlap occurs, since RfQ's and award decisions are also both considered announcements. However, the publication types RfQ's and award decisions are two reliable indicators of the number of European procurements in the population (ABK: 16.682 / TED: 15.664 and 14.861). As discussed previously this is not the case for national procedures (1.845), because, not all national procedures are published. In addition, if national procedures are published, the award decision is not always uploaded to ABK. Therefore, the exact number with regard to national procurements remains unidentified. In Appendix A (Table II) more information is provided on different subgroups within European procedures. This information is useful to when correcting for bias (further discussed in paragraph 2.2).

3. Contract subjects

In public procurement, three contract subjects exist; works, goods and services. To correct for the possible confounding effects of function and task, the population and consequently sample are divided into three subgroups. These subgroups are spread over the publication type 'all announcements' within the database ABK. This consideration is made since 'all announcements' include both national and European procurements. In addition, in the database ABK, all tender documents are usually available over the period of interest.

In Table 2, a condensed overview is presented, a complete overview can be found in Appendix B. Procurements are categorized according to publication years, publication months, and contract subjects, following the sample characteristics.

Population	2010 (8 months)	2011	2012	2013	2014 (4 months)	Total
Works	6.895	9.506	7.074	4.159	869	28.503 (42,2%)
Goods	2.592	4.267	3.341	2.513	703	13.416 (20,0%)
Services	5.436	8.014	6.321	4.328	1.214	25.313 (37,7%)
Total	14.923 (22,2%)	21.787 (32,4%)	16.736 (24,9%)	11.000 (16,4%)	2.786 (4,1%)	67.232 (100%)

Table 2: Overview procurement contract in 'all announcements' (NL: May 2010 – April 2014)

In Appendix A (Table III), percentages for all three contract subjects are shown for European procurements. The percentages for European contract subjects are distributed differently from the percentages for both national and European contract subjects.⁴ This difference does not necessarily imply a bias since also national procurements are present (Table 2). The distribution with regard to European contract subjects; works, goods and services within the population (Appendix A – Table III) are compared with the sample to determine bias effects, this is further discussed in paragraph 2.2.

Please note that a disadvantage of the publication type 'all announcements' is the presence of overlapping categories: a single procurement can occur 1-5 times depending on the publication type (RfQ's, award decision, rectifications, publications in Aanbestedingskalender as well as TED and Tendered⁵). However, the distribution between 'all announcements', 'RfQ' and 'award decisions' is approximately the same for European public procurements (as shown in Appendix A and Tables I, II and III).

2.2 Quantitative data (population and sample)

Each selected procurement is analyzed based on the established research format presented in Appendix F. As discussed previously, this research aims at explaining 'problems', with the impact of 10 elements in the award phase by means of an empirical study of award methods of Dutch public procurements. Thus, the empirical study consists of two groups of procurements. The first group consists of a stratified random sample of Dutch public procurements, whereas, the second group exists of a large sample (of the total population) of Dutch public procurements that resulted in problems. In Figure 2, an overview of the population is presented.

⁴ Appendix A table III – Publication type 'all announcements': Works 14,8% / Goods 31% / Services 54,2%. These percentages are different from ABK: Works 42,2% / Goods 20% / Services 37,7%.

Difference between TED and ABK in percentages: Works 27,4% / Goods 11% / Services 16,5%.

⁵ Tendered.nl (TenderNed: the digital procurement system of the Dutch Government. All contracting authorities publish their RfQ's (announcements) on this platform, both national and European Tenders.)

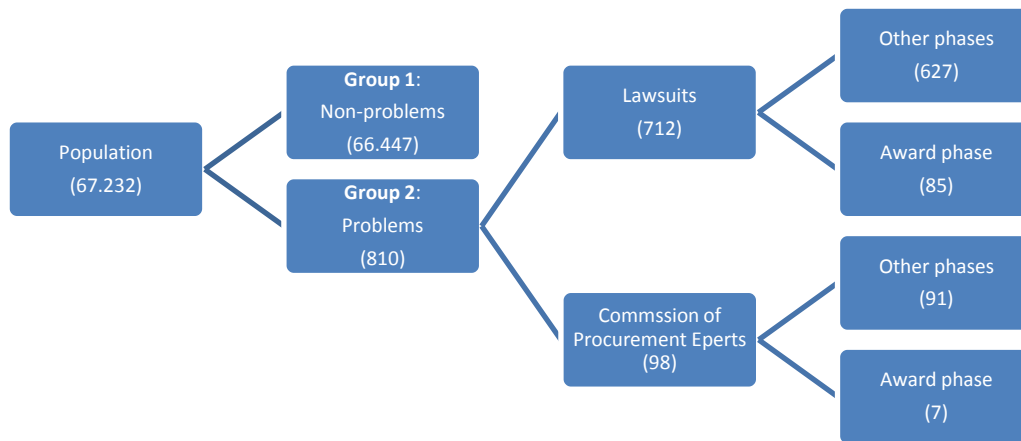


Figure 2: Population – Overview groups (NL: May 2010 – April 2014)

Overall, 810 procurements resulted in problems. More specifically, 712 procurements resulted in a lawsuit and 98 procurements resulted into questions that were addressed by the Commission of Procurement Experts. However, these 712 lawsuits also included problems that did not concern the award method (e.g. motivation of the award decision, a tender bid which offered too low prices, selection criteria, product specifications, etc.).⁶ In the other indicator for ‘problems’, questions that were addressed by the Commission of Procurement Expert, overall 98 procurements were identified. In total 25 questions that have been addressed were available. Similarly to lawsuits, not all of these questions considered the award method. In fact only 7 procurements concerned the award method.⁷

In the sample (Figure 3), group 1 has a size of 402 procurements. For a representative sample as well as for efficiency’s sake, the population is grouped into three main categories (works, goods and services) and divided in groups of publication months. This resulted in 48 groups before sampling (see Appendix C – sample group 1). The populations and sample are divided in these groups to deal with the available selection options of the used data source ABK, which are rather limited. In case a single large sample was drawn, sampling variability was likely to occur.⁸ Besides, the data collection would have resulted into problems due to the available options in ABK. In statistics, this sampling method is referred to as *stratified random sampling*: within each predefined group a random sample is drawn.

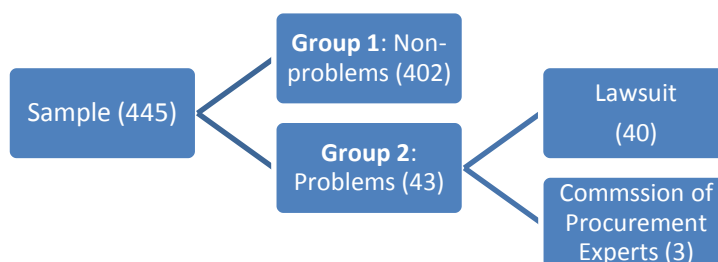


Figure 3: Sample – Overview groups (NL: May 2010 – April 2014)

⁶ Uitspraken.rechtspraak.nl (In the ‘search option’ it is possible to refine searches. In total 712 lawsuits were published in the examined period. Only 85 procurements resulted into a lawsuit with regard to the award method.)

⁷ Commissievanaanbestedingsexperts.nl/behandelde-klachten (A total of 25 questions that were addressed at by Commission of Procurement Experts are published, 7 are concerning the award method.)

⁸ Sampling methods that, by their nature, tend to over- or underemphasize some characteristics of the population are said to be biased (Babbie, 2010).

In the sample (Figure 3), the second group ‘problems’, consisted of 92 procurements (see Appendix D – group 2) of which 85 lawsuits and 7 questions addressed by the Commission of Procurement Experts. After the removal of the procurements with incomplete information, only 43 procurements made it into the sample (40 lawsuits and 3 questions addressed by the CoPE). To verify for bias, all 92 identified procurements are compared to the 43 that made it into the final sample (see Appendix E). The results indicate that the exclusion affected the type of procurement procedures, causing relatively more open procurement procedures to be present in the sample. (e.g. national and European open procedures).

In the second group, ‘problematic procurements’, all relevant and available procurements were selected for analysis. This is not a random sample; all procurements are selected based on their substantive aspects. The overall expectation is that the selected procurements provide a sufficiently representative and reliable idea of the applied award methods, specifically the ten identified elements. Furthermore, all procurements which resulted in ‘problems with regard to the award method’ are selected in the investigated time frame.

Appendix G presents a complete overview of the sample characteristics. The sample is divided into: contract subject, procurement procedure, contracting authority and award criterion (EMAT or lowest price). These descriptive statistics are displayed for the first group (non-problems) and second group (‘problems’ – for both measures). In addition, this information is compared to all descriptive information discussed in paragraph 2.1 and Appendix A (Tables I, II and III).

1. Organizations (European public procurement)

In the first subgroup ‘contacting authorities’, little information was available. The only possible comparison between the population and sample indicated no differences between European procurements and organization types (see Appendix A – Table I and Appendix G – Table VII for a complete overview). The distributions within the groups are approximately the same.⁹ Therefore, the population and sample are not biased towards certain organization types.

2. Procurement procedures (European public procurement)

In the second subgroup ‘procurement procedures’, two comparisons were made. At first, the distribution over national and European procedures is checked. In the entire population, 25.484 (38%) national procurements and 41.748 (62.1%) European procurements are present (paragraph 2.1 – Table 1). In the sample, this is approximately similarly, 169 (42%) national procurements and 233 (58%) European procurements (Appendix G – Table II).

Secondly, the variation within European procurements is checked. In the population 11.858 (75,8%) European open procedures and 3.096 (19,8%) European restricted procedures were identified (Appendix A – Table I). In the sample, this is distributed differently, 219 (94%) European open procedures and 14 (6%) European restricted procedures (Appendix G – Table VIII), which indicates a bias in procurement procedure types. As discussed previously, documents are not always available in restricted procedures. A consequence of this bias (a difference of 18,2% in restricted procedures) is that we are limited in generalization of our findings, when restricted procedures are involved.

⁹ Sample European procurements: State (9,9%) / Provinces (5,2%) + Municipalities (32,6%) + Water Board = (10,3%) = (48,1%) / Public entities (18,9%) / Cooperation (10,7%) / Other (12,4%).

Population European procurements: State (13,3%) / Provinces + Municipalities + Water boards (44,0%) / Public entities (26,8%) / Other (15,9%)

3. Contract subject

In the third subgroup ‘contract subjects’ the population is examined within ABK. The sample is divided over the contract subject types, a *stratified random sample*. Inherently, the population and consequently sample have a similar distribution (Appendix B and C). To check for bias effects in European procurements, the database TED as again applied (Appendix A – Table 3), of which an abstract is presented in Table 3.

Contract subject	Population (TED)	Sample (European procurements)
Work	2.334 (14,9%)	37 (15,9%)
Goods	5.198 (33,2%)	76 (32,6%)
Services	8.112 (51,9%)	120 (51,5%)
Total	15.664 (100%)	233 (100%)

Table 3: Overview European procurement – contract subject in ‘RfQ’ and sample (NL: May 2010 – April 2014)

Bias identification check – Overall conclusion

To correct for bias, three subgroups were taken into account, namely; organizations, procedures and contract subjects. Since there are different options in these subgroups, distributions are examined in order to check for over- or underemphasis of any of the indicated characteristics. In national procedures, little information was available, only general bias checks were allowed. Therefore, we are unable to link the research sample with the accurate number in the research population. However, the general analysis did not indicate a bias. In addition, the research sample is drawn from the publication type ‘all announcements’, including all available subgroups. In European procedures, the research population could be examined in more depth. For all three subgroups distributions were approximately the same, indicating no presence of bias. The only differences between the sample and research population is that the sample is weighed more heavily towards open procedures. The expectation is that this would also be the case for national procedures. Therefore, the only characteristics that are underemphasized in the research sample are restricted and negotiated procedures. It is important that we take these differences into account when generalizing our findings.

2.3.1 Data sources

To collect the required information for this study, three data sources were of importance. The first and most important data source is ‘aanbestedingskalender.nl’. Here, contracting authorities publish their RfQ’s and aligned tender documents. ABK is used since it assures a complete overview of all published procurements in the investigated time frame. In addition, the database is fully in line with national and EU procurement regulations: it includes links with TenderNed and TED. From ABK all information is available for the analysis of award phases, for non-problematic as well as problematic procurements.

The other mentioned data sources are consulted for the selection of the ‘problematic procurements’. In the Dutch database ‘uitspraken.rechtspraak.nl’, all judgments of the Netherlands Judiciary are published. In this database, 85 lawsuits were identified based on their substantive aspects.¹⁰ After a search in ABK, only 40 procurements included the relevant documents and made it into the sample.

¹⁰ A search is performed on the following search terms: beoordelingssystematiek, beoordelingsmethodiek, gunningssystematiek, gunningscriterium, EMVI and berekeningsmethodiek.

In the last data source, 'commissievanaanbestedingsexperts.nl', procurements are selected that resulted in questions addressed by the Commission of Procurement Experts.¹¹ In total 25 questions were addressed by the Commission and resulted in recommendation. Only 7 of the 25 questions were related to the award method, and after a search on information completeness in ABK only 3 appeared to be useful.

2.3 Measures

In the introduction, 'problems' and the ten identified elements which are expected to cause these problems are introduced. Below, these concepts are made measurable by operationalisation.

2.3.1 Measuring 'problems' in procurements (dependent variable)

The dependent variable in this analysis is 'problems in procurements', or more explicitly, problems in the award phase concerning the award method. These problems are measured according to two possibilities of legal protection that suppliers have in public procurement. Specifically: procurements that resulted in lawsuits, and procurements that resulted in questions that were addressed by the Commission of Procurement Experts. Both measures relate to the award criteria and award methods in the award phase. In the first option 'lawsuits', the contracting supplier accuses the contracting authority of insufficient transparency and/or objectivity on the used technique to identify the winner in the award phase. The same holds for questions that are addressed to the Commission of Procurement Experts. In this case, the supplier asked a question regarding procedures or characteristics of the award method.

Thus, problems are measured according to two indicators;

- The Commission of Procurement Experts and;
- The national court (District court).

The above indicators are used since these are serious measures of the supplier to complain about the award method. The used measures indicate certain problems with regard to the award method. Both measures for problems are more thoroughly discussed in chapter 4. In chapter 4, choices are motivated and further information is provided about both measures.

2.3.2 Measuring 10 elements in public procurements (independent variable)

The independent variables in this research consist of the ten identified elements, which are subject to leeway in the award phase and are expected to cause problems. All ten elements are further discussed in chapter 3, which describes, the implication of each element in terms of their impact on problems. Furthermore, within each element a hypothesis is formulated. All procurements are analyzed with regard to the presence of the following ten elements:

- The award criterion 'EMAT' (in particular the Weighted Factor Score in price-to-quality scoring)
- The use of qualitative - constructed performance indicators
- The weights are not published (the impact of award criteria is unclear)
- The score method is not published in the documents (unclear how the score is calculated)
- Form of the score graph for price (hollow)

¹¹ [Commissievanaanbestedingsexperts.nl/behandelde-klachten](http://commissievanaanbestedingsexperts.nl/behandelde-klachten)

- Award criterion contains relative scores
- Scores for an award criterion are awarded on the basis of ranking/hierarchy
- Flat award criteria (resulting in minor distinction between different bids)
- Extra lower boundaries (bids scoring under this lower boundary are excluded)
- Award criterion price consist of various financial sub-criteria.

In case any of these elements are present in the award phase, procurements are prone to become problematic. Inherently, award phases become more susceptible to problems when these 10 elements are increasingly present.

2.4 Data analysis

The objective of the analysis is to examine the effect of the 10 identified elements on problems in the award phase with regard to the award method. Therefore, ten one-sided hypothesis are formulated in chapter three. The analyses of these hypothesis consists of cross-section comparison, that is comparing the differences between the 'non-problematic group' and the 'problematic group'. Thus, each hypothesis involves a single element, which is then examined in both groups, in order to examine if the variable is significantly related to the 'problematic group'. To measure whether or not these differences were significant, a z-test for two proportions (chi-square test) with $\alpha = 0.05$ was used.

2.5 Research Method

The type of research approach is a '*cross-sectional study*', a form of observational research. In this observational study inferences are drawn at one specific point in time on the possible effect of a treatment on subjects. The way in which this research is conducted is both qualitative and quantitative. The quantitative method prevails, since this method is related to the data collection and analysis of both groups of procurements. The following sections describes how the qualitative and quantitative research methods were applied.

Qualitative research method (literature review):

According to Babbie (2010), qualitative research involves the non-numerical examination and interpretation of observations, for the purpose of discovering underlying mechanisms and patterns of relationships (p. 390). In this research, the qualitative research method consists of a literature review in chapter three, answering the first sub-question; 'which 10 elements are selected on theoretical grounds that have a negative impact in the award phase and are expected to result in problems?' The literature review provides the theoretical ground on basis of which all 10 elements are selected. In addition, the consequences of each element are discussed in relation to the result, 'problems in the award phase regarding the award method'. An answer to the first sub-question provides information for the final chapter in which the findings, conclusions and recommendations are presented and discussed.

Quantitative research method (research format):

With quantitative data, the emphasize lies on numerical data. Qualitative research involves, 'the numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect' (Babbie, 2010, p. 414). In this study a standard research format is used, published in Appendix A. This format is created based on the necessary descriptive statistics required for execution of the research (control characteristics). The

10 indicated elements of the theoretical framework, which serve as independent variables, are also included (chapter 3), as well as two measures of problems, the dependent variables (chapter 4). In the published format, the coding process is also indicated. All 10 elements are categorical variables, also called nominal variables. These variables are measured on a nominal scale; when an element is present this is indicated with a zero (coding = 0), when a element is not present, this is indicated with a one (coding = 1).

2.6 Conclusion

This chapter described how the research is conducted. This research is a quantitative study of award phases, used in 445 Dutch public procurements between May 2010 and April 2014 – covering a period of four years. Two groups of procurements are compared on the basis of presence of their dependent variable; ‘problems in the award phase with regard to the award method’. The dependent variable is indicated with the variables ‘lawsuits’ and ‘questions addressed by the Commission of Procurement Experts’. The independent variables serve as a basis for comparison, these variables were specified and made measurable by operationalizing them. Finally, this chapter describes how obtained data is analyzed by the specification of the statistical test and research methods.

3. Elements in the award phase

In this chapter the theoretical framework of the research is outlined. The theoretical framework serves as guidance for this research; it determines the characteristics that will be measured and which statistical relationships are examined.

This chapter highlights expectations with regard to public procurements that have resulted in problems, and, on the contrary, public procurements that should not result in problems. As discussed previously (paragraph 1.1), many various formal decision methods, decision elements, as well as various quantitative and qualitative decision criteria for supplier selection can be taken into account (e.g. De Boer et al., 1998, 2001; Ho et al., 2010), resulting in a range of available options and hence choices. This makes it impossible to cover the effects of all possible supplier evaluation and selection approaches in this research. Therefore, 10 frequently occurring elements which are expected to result in problems are selected from literature. After identifying these elements, it can be determined whether these elements are too a larger extent present in problematic procurements. The degree of presence of these elements then provides a possible explanation for problems. Based on the 10 indicated elements, in the following sections, both groups of procurements are analyzed.

3.1 Characteristics in the award phase

The 10 elements which are expected to affect the result (either problems or non-problems) in procurements are discussed in this section. In order to do so, the publication ‘Gunningsmethodieken voor Uitzendwerk’ of the Public Procurement Research Centre (PPRC) is used. In this publication, eight elements are discussed and analyzed that can result in undesirable or unpredictable outcomes, both from economic and legal perspectives. A legal analysis focused in particular on the legality of the used award method, while an economic analysis focused on the effectiveness of the used award method (Lohman, Manuza and Telgen, 2013, p 2.). The original representation of these eight elements is presented in Table 3.

Elements award method
1. Publication of weights
2. Award criterion price consist of various financial sub-criteria
3. Score method not published in tender documents (unclear how score is calculated)
4. Award criterion contains relative scores
5. Flat award criteria (resulting in minor distinction between different bids)
6. Extra lower boundaries (bids scoring under lower boundaries are excluded)
7. Score graph for an award criterion is hollow
8. Scores for an award criterion are awarded on the basis of ranking/hierarchy

Table 3: Eight elements of award methods - overview (Public Procurement Research Centre, 2013, p 3)

In this research, the eight elements of award methods indicated by the PPRC (2013) serve as a starting point. Especially, the focus in this study is on the economic analysis, since this analysis concentrates on the adequacy and hence effectiveness of the award methods (Lohman, Manuza and Telgen, 2013). Therefore, the economic analysis is directly related to the elements used in the award method and their effect on problems. The eight elements are appended with two additional characteristics of the award phase. These are selected from the article ‘Tender evaluation and supplier selection method in public procurement’ of Bergman and Lundberg (2013), presented in Table 4.

Elements award method

1. Award criteria ('lowest price' or 'EMAT')
2. Sub-award criteria (natural, proxy and constructed performance levels)

Table 4: two elements of award methods – derived of Bergman and Lundberg (2013).

These two elements are selected to determine whether different types of supplier selection criteria and supplier selection methods affect the result 'problems'. By doing so, a new table is created that has a strong theoretical base adapted to this research. In Table 5, an overview is presented of all 10 elements which are selected for this research, and the negative implications of each element are summarized. The table starts with the two selected elements from the article of Bergman and Lundberg (2013).

Elements award method	Effect of used elements on outcome 'problems' (theoretical implications)
1. Award criteria ('lowest price' or 'EMAT') <ul style="list-style-type: none"> - The presence of the WFS 	Contracts can be awarded on the basis of the 'lowest price' and the 'EMAT'. In case of the award criterion 'lowest price', only the aspect price is relevant, whereas in the award criterion 'EMAT', multiple criteria and arrangements are applied to identify the most economically advantageous tender (art. 53 European directive 2004/18/EG). Using both price and quality in supplier selection, enhances the efficiency of public procurement, although it adds complexity to the procedure (Bergman and Lundberg, 2013).
2. Sub-award criteria <ul style="list-style-type: none"> - Performance measure type - number of sub-criteria 	In case of the award criterion 'EMAT', it is obligatory to publish the further award criteria. It is the responsibility of the contracting entities to indicate the criteria for the award of the contract (art. 53 European directive 2004/18/EG). However, no regulations are listed on the type (quantitative or qualitative) and number of these criteria.
3. Weights not published <ul style="list-style-type: none"> - Weights award criteria (EMAT) 	According to the European and Dutch procurement rules, it is the responsibility of the contracting authority to indicate the weights assigned to each of the criteria. In case this is not possible, contracting entities may derogate from indicating weights by ranking the criteria in descending order of importance (art. 53 European directive 2004/18/EG). Note: it is impossible for suppliers to submit an offer that is best for the contracting authority if weights are not published.
4. The score method is not published in the documents (unclear how the score is calculated)	Suppliers that are able to issue multiple bids are not able to calculate the offer preferred by the contracting authority. In this case, it is possible that the best offer never reaches the contracting authority. In addition, the contracting authority has the possibility to manipulate the award after all bids are received, the winning offer depends on the choice of score method (Lohman, Manuza and Telgen, 2013).
5. Form of the score graph for price (questionable)	Contracting authorities are often not aware of the form of the score graph, while this entails important implications for the awarding of the contract. Note: a hollow score graph has – regardless the weighting factor – a strong preference for the award criterion of (lowest) price (Lohman, Manuza and Telgen, 2013). Only the award criterion price is taken in account.
6. Award criterion contains relative scores (price and quality)	The use of relative scoring is not professional. The significance of the differences in performance between two suppliers depends on the bid of a third party. Furthermore, when a non-competitive tender is added or omitted, it is possible that the ranking paradox and/or 'bid rigging' can occur (Lohman, Manuza and Telgen, 2013). Relative scores can relate to price and quality.
7. Scores for an award criterion are awarded	Differences in scores do not provide information on the differences in performance. The contracting authority does not determine how much

on the basis of ranking/hierarchy	money is available for better performance on qualitative award criteria. In addition, this score method includes relative scoring and all its associated disadvantages (Lohman, Manuza and Telgen, 2013).
8. Flat award criteria (resulting in minor distinction between different bids)	Flat award criteria have no effect on the award. The award criterion has a weighting factor with very little impact. This results in other award criteria becoming of greater importance than actually intended. Note: the importance's of different award criteria remain unclear; suppliers do not necessarily submit an offer which is optimal for the contracting authority (Lohman, Manuza and Telgen, 2013).
9. Extra lower boundaries (bids scoring lower are excluded)	By including a lower boundary for a score on an award criterion, the rating scale, and therewith distinctiveness is reduced. Note: other award criteria are given greater importance than actually intended and the best offer might not be submitted by suppliers (Lohman, Manuza and Telgen, 2013).
10. Award criterion price consist of various financial sub-criteria	Numerous sub-criteria within the award criterion of price can result in strategic tendering, especially when sub-criteria contain bid prices for different products. Note: suppliers increase their chances of award without the offering the contracting authority a better bid (Lohman, Manuza and Telgen, 2013).

Table 5: complemented table – ten elements used in award methods

In sub-sections 3.1.1 to 3.1.10 all elements are separately discussed in terms of their effect on the outcome 'problems'. Furthermore, within each element, an hypothesis is formulated regarding the expectations in the 'problematic' and 'non-problematic' procurements. In the first three sub-sections (3.1.1 – 3.1.3), a more detailed explanation is provided, since the three elements are not discussed in the articles from which they are derived. The other sub-sections (3.1.4 – 3.1.10) discusses all elements briefly, since these elements are all addressed in their original publication 'Gunningsmethodieken voor Uitzendwerk' of the PPRC, (2013), which provides, theoretical effects, legal aspects and examples of all seven elements concerned.

3.1.1 Element 1 – Supplier selection methods (Award criteria: 'lowest price' or 'EMAT')

According to the European and Dutch procurement rules, there are two supplier selection methods: the lowest price and the EMAT. The latter requires that further award criteria and their weights are published. In addition, the scoring rule must be specified. According to Bergman and Lundberg (2013), "a scoring rule is a function that assigns a numerical value to different quality levels in a particular dimension or that transforms a value measured on one scale (price or quality) into a measure on another scale (price score or quality score)" (p.75). While a weighing function is defined as, "a function that combines price and quality (or price score and quality score) into a single value so that different bids can be compared and ranked" (p.75).

If both price and quality are used in supplier selection, the efficiency of public procurement can be enhanced, although it adds complexity to the procedure. The supplier selection methods, specifically the scoring rules that are used in practice, are often poorly designed (Bergman and Lundberg, 2013; Telgen and Schotanus, 2010). An overview of possible supplier selection methods in public procurement is presented in Figure 4.

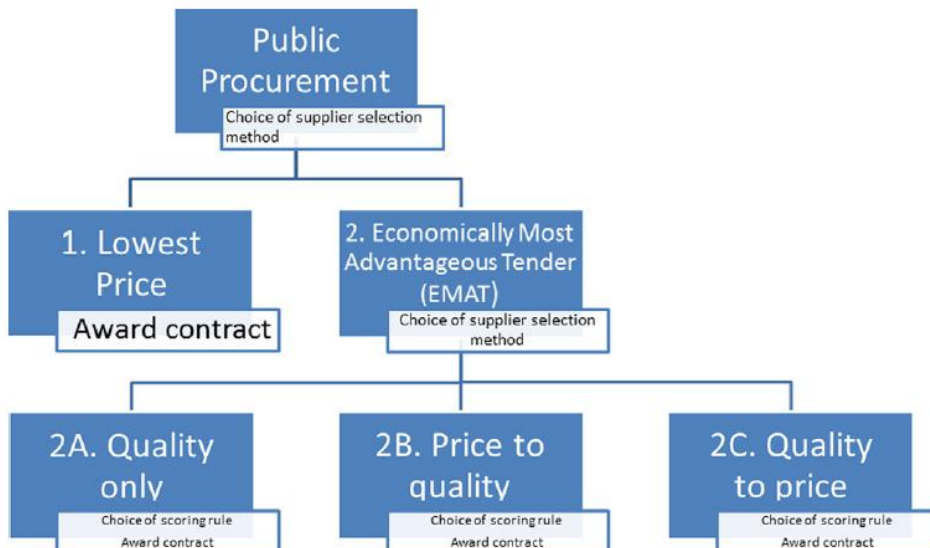


Figure 4: Supplier selection methods (Bergman and Lundberg, 2013)

In the first award criterion Lowest-Price, an evaluation is, theoretically, straight-forward. However, in practice, it may be challenging to define effective and appropriate minimum quality requirements, as well as to weigh multiple prices into a single cost measure (Bergman and Lundberg, 2013). In the second award criterion; EMAT, the options are threefold:

- In option 2A Quality-only (fixed-price), the EMAT can be the bid that offers the highest quality for a given, predetermined price. Quality-only tender evaluations can be complicated. If quality is measured in more than one dimension, the quality measures will have to be combined (weighed) into a single overall score. Contrary to prices, they cannot simply be added together. In this research, this method will be named Fixed-Price.
- In option 2B Price-to-quality, the options are twofold. In this first option, price must be transformed into a score that is added to the quality score, making the tender a price adjusted highest-quality tender. This method will be mentioned as the Weighted Factor Score (WFS). This method is also the most common in public procurement (Bergman and Lundberg, 2013; Lohman, Manuza and Telgen, 2013). In the second option, only the quality aspect is transformed in a score. The score on quality is divided by the price (in euro's). This method is named Value For Money (VFM).
- In the third option 2C Quality-to-price, quality (more than the minimum requirement) is assigned a value that can be subtracted from the price bid (bonus-malus). Alternatively, the value of the quality gap, relative to the maximum quality level can be added to the price bid. For example, the supplier selection method will be quality adjusted lowest price. To this method is referred with Awarding on Value (AOV) "Gunnen op Waarde (GOW)" (Jansen et al., 2007; CROW - publication 253).

According to Bergman and Lundberg (2013), the second and most common option, Price-to-quality (WFS and VFM) is non-transparent. This method makes an accurate representation of the procurer's preferences difficult. Also, the method it is often open to strategic manipulation due to dependence on irrelevant alternatives, and it tends to impose particular and unjustified non-linearity in bid prices. Furthermore, the WFS, as applied in practice, often violates both the transparency principle and the principle of equal treatment (Bergman and Lundberg, 2013).

Based on the referred literature, the following hypothesis is formulated:

Hypothesis 1: The Weighted Factor Score (a form of price-to-quality scoring) occurs more often in problematic procurements than in non-problematic procurements.

3.1.2 Element 2 – Award criteria (Sub-criteria)

As discussed previously (Table 5), it is obligatory to publish further award criteria in the award criterion EMAT. In the EMAT, further award criteria are performance measures. Therefore, the meaning of each criterion must be described and clarified, as objectively as possible. According to Mateus et al. (2010) “a performance measure is an ordered set of plausible performance levels set on a quantitative or qualitative scale” (p. 209). A performance measure may thus be defined as either a quantitative measure (numbers only) or a qualitative measure (verbal or pictorial descriptions). According to Keeney (1992) there are three types of performance measures; natural, proxy and constructed, which are discussed below:

1. Natural (quantitative): Performance levels directly reflect the effects, consequences or ends to be evaluated (e.g. a price criterion measured by total cost in euro’s).
2. Proxy (quantitative): Performance levels mostly reflect causes or means to be evaluated (e.g. a technical quality criterion measured by means of a warranty in years).
3. Constructed (qualitative): Performance levels are defined through a holistic combination of plausible tenders features to be evaluated (such as a work plan measured by various criteria).

According to Mateus et al. (2010), it is preferable to assign a natural performance measure to each criterion, as this type allows for the most directly and objectively defined performance levels. Natural performance measures make criteria less ambiguous; the tender evaluation model will be more easily accepted (less controversial) and effective, thus making the award decision easier to sustain.

Hypothesis 2: Qualitative – constructed performance measures – are to a larger extent present in problematic procurements than in non-problematic procurements.

3.1.3 Element 3 – Publication of overall weights

According to the European and Dutch procurement rules, it is the responsibility of the contracting authority to indicate the criteria weights in sufficient time for tenderers to be aware of them when preparing their tenders (art. 53 European directive 2004/18/EG section 2 – paragraph 1). Contracting entities may derogate from indicating the weighing of the criteria for the award of the contract in duly justified cases (section 2 - paragraph 2). In such cases, they must indicate the importance of the criteria in descending order (section 2 – paragraph 3).

Indicating weights is from a purchasing point of view desired, since publication of weights can lead to better offers (Telgen and Schotanus, 2010). In order to offer the best bid, suppliers need to know the contractors preferences that emerge from the award system. Often, providers have the possibility to offer multiple bids. For example, suppliers can increase quality by offering a higher price or, contrary, less quality at a lower price. Therefore, it is important to know which offer is preferred by contracting authorities (Lohman, Manuza and Telgen, 2013).

From a legal perspective, refraining from publishing the weights is uncommon and explicitly forbidden in procurement regulations. Therefore, the overall expectation is that few cases will be

found in which the weights are not published. However, if the weights are not published in the tender documents, problems are expected to occur. Important to note; if the award method AOV is used, the weights are not directly published. AOV uses monetary values for the comparison of bids, while other award methods compare points. In case of the award method AOV, the relative delta values need to be published, these are the total monetary values which can be earned for a award criterion (specification of the bonus-malus values).

Hypothesis 3: not publishing the weighing in tender documents leads to the outcome 'problems'.

3.1.4 Element 4 – Score method not published in tender documents

In the award phase, performance on award criteria is measured by the contracting authority. When performance of suppliers on award criteria is assessed, all offers on award criteria (such as price or quality) are usually converted to scores (element 1 – definition score method). According to De Boer et al. (2006) and Lohman, Manuza and Telgen (2013) the way to relate criteria to each other and the weights involved are hardly overlooked in procurement documents. However, these steps are still non-trivial in determining scores on each criterion. In order to determine scores, a score method needs to be published. However, in practice this is often not acknowledged and consequently given too little attention. From a purchasing point of view, it is both theoretically and empirically proven that the buyer should fully disclose all details of the awarding mechanism to bidding suppliers (Telgen and Schotanus, 2010; Albano et al., 2008). The omission of the score method leads to worse offers and leaves the possibility open to manipulation.

In addition, not publishing the score method can lead to unexpected and unintended results, because different types of score methods can lead to different winners. This entails a negative effect, since the best offer – according to the contracting authority – is not indicated with the used type of score method. (Lohman, Manuza and Telgen, 2013).

Hypothesis 4: No publication of the score method occurs more often in problematic procurements than in non-problematic procurements.

3.1.5 Element 5 – Form of the score graph (price)

In the most common award method, the offered bid price is transformed into score points (WFS). It is important to determine minimum and maximum scores within the award method with care. However, the scores in between should not be neglected. Intermediate scores and the way they are determined (the scoring curve) may be equally important in determining the winner (De Boer et al., 2006). Figure 5 shows three different scoring curves that could be considered on the same maximum and minimum scores, in no intermediate scores are provided.

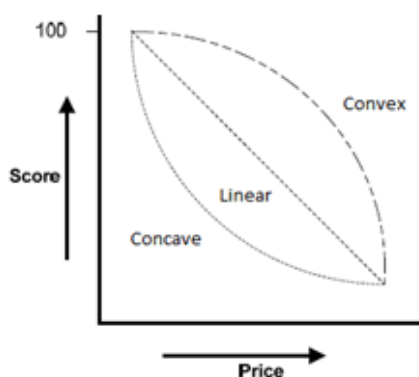


Figure 5: Three types of score graphs (De Boer et al., 2006)

Important to note; the real contribution of an attribute is measured not only by the weight of the criterion, but also by the score of the offer on that criterion (De Boer et al., 2006). Based on the formula of the score, a graph can be obtained (price on the x-axis and score points on the y-axis). The form of the score graph can be different, a straight line (linear relationship) or a concave or convex shape (non-linear relationship). Often, the contracting authority is not aware of the shape of the score graph, while it has important implications for the award (Lohman, Manuza and Telgen, 2013). The most unfavorable, and unintended, effect occurs when a hollow score graph is used for award criterion price. This type of graph has, regardless the weighting factor, a strong preference for the award criterion (lowest) price (Lohman, Manuza and Telgen, 2013). A linear score graph solves the problem of non-linear point distribution, since a price difference in a linear graph delivers an equal point difference.

Hypothesis 5: Non-linear score graphs, especially hollow score graphs for the award criterion price, occur more often in problematic procurements than in non-problematic procurements.

3.1.6 Element 6 – Award criterion contains relative scores

The scores on award criteria can be calculated by means of various scoring methods. According to Telgen and Schotanus (2010), these methods can be classified under as either absolute (i.e., independent) or relative (i.e., interdependent) methods. The calculations based on relative methods depend on the best, worst and/or average supplier bids, while the calculations based on absolute methods are independent of other offers (Telgen and Schotanus, 2010). Therefore, the assessment of a bid in relative methods is impossible without knowledge of other bids, making the scoring rule opaque to suppliers (Bergman and Lundberg, 2013). Bidders may inadvertently submit a non-optimal bid.

Telgen and Schotanus (2010) provide evidence that the buyer should fully disclose all details of the awarding mechanism to suppliers, in order to receive better bids and to avoid subjectivity in supplier selection. When these details are disclosed, a tenderer is able to define his strategy by comparing his monetary cost of improving a 'quality' point against its monetary value. Therefore, the tenderer can optimally allocate his budget among all aspects of the contract (Asker and Cantillon, 2008, 2009). This was also shown by Albano et al. (2008), who proved that absolute scoring methods lead to better price-quality ratios than relative scoring methods. Furthermore, Dini et al. (2010) and Chen (2008) emphasized the favourable properties of scoring rules that are linear in prices. Moreover, the use of relative scores is not professional: By applying relative methods, the contracting authority does not address the importance of differences in performance. Performance of suppliers is then

assessed based on coincidences since the contracting authority does not indicate the importance of differences in performance (Lohman, Manuza and Telgen, 2013).

Hypothesis 6: Relative score methods occur more often in problematic procurements than in non-problematic procurements.

3.1.7 Element 7 – Scores for an award criterion are awarded on the basis of ranking.

In the evaluation of offers, performance is assessed per award criterion. Usually, a score graph is used to indicate performance. However, there are also methods that do not directly assign points on the basis of performance, but indirectly, based on ranking. In this method, the used measurement scale is ordinal. The scale only shows the ranking of the tenders, but is unable to provide information on the degree of differences between different bids. In this method, the differences in scores do not explain the differences in performance (Lohman, Manuza and Telgen, 2013). This means that contracting authority does not determine what an improved performance on the award criteria would be worth.

Furthermore, by using this method, contracting authorities run the risk of not identifying the EMAT. A offer that is marginally more expensive compared to the offer with the lowest price receives significantly less points for the award criterion price. The latter difference is hard to compensate on qualitative award criteria. An additional disadvantage of awarding points on the basis of ranking is that this method is relative (element 6 – paragraph 3.1.6).

Hypothesis 7: Methods which award scores on the basis of ranking occur more often in problematic procurements than in non-problematic procurements.

3.1.8 Element 8 – Flat award criteria

A 'flat' award criterion reduces the contribution of an award criterion and, therefore, leads to a different distribution of weights amongst award criteria. This results in a unintended effect, since other award criteria have greater importance (Mateus et al. 2010; Telgen and Schotanus, 2010). 'Flat' implies that the score method used for the award criterion is unable to differentiate between offers: all bids receive approximately the same score. If a 'flat' award criterion is used, other award criteria (in case these criteria are not 'flat') are determining the award (Lohman, Manuza and Telgen, 2013).

In case 'flat' award criteria are present in the award phase, the entire range of the rating scale is not fully used. This results in a lower contribution of a criterion than originally expected (as discussed previously; a contribution of an attribute is indicated not only by the weight of the criterion, but also by the score of the offer on that criterion (De Boer et al., 2006)). In addition, according to Bergman and Lundberg (2013), the exact weights will not make the bid evaluation mechanism transparent, nor predictable, in case of poorly specified quality scales. Using both flat and non-flat criteria affects the relative importance of criteria. The use of flat criteria can result in a contribution of criteria, that is different from the original intentions of the contracting authority.

Hypothesis 8: Flat award criteria occur more often in problematic procurements than in non-problematic public procurements.

3.1.9 Element 9 – Extra lower boundaries

In the award phase, often lower boundaries are included for one or more award criteria. For instance, by stating that an offer has to obtain a minimum value. Often, contracting authorities do

not take into account the effect of a lower boundary on the contribution of a criterion in the award phase (Lohman, Manuza and Telgen, 2013). The use of a lower boundaries results in unintended higher contributions of other award criteria (such as price), or other criteria that do not include lower boundaries.

Hypothesis 9: Extra lower boundaries occur more often in problematic procurements than in non-problematic procurements.

3.1.10 Element 10 – Award criterion price consist of various financial sub-criteria

The price criterion can be a combination of different financial sub-criteria. For instance, when a contracting authority requests several products that are individually priced. This method is often used when contracting authorities do not exactly know how many examples are needed from each product during the term of the contract (Lohman, Manuza and Telgen, 2013).

A combination of different price sub-criteria invites strategic tendering. This holds especially when contracting authorities estimate the purchasing volume of products differently, or when these purchasing volumes can be a subject to manipulation. In case a purchasing volume is overstated, the suppliers offers low prices, in case a volume is underestimated, the supplier consequently offers a high price for a product. In this case, the supplier ends in a convenient total price and has more chances in winning the tender. In case the estimated purchasing volume are correct according to the supplier, the cost of the contract for the issuing authority are higher than necessary (Lohman, Manuza and Telgen, 2013).

Hypothesis 10: An award criterion price that consist of various financial sub-criteria occurs more often in problematic procurements than in non-problematic procurements.

3.2 Conclusion

This chapter tackled first sub-question: Which 10 elements are selected on theoretical grounds that have a negative impact in the award phase and are expected to result in problems? All 10 elements are shortly discussed in an overview (Table 5 – overview). Based on the extent of occurrence of these elements in 'non-problematic' and 'problematic' procurements, the relationship was studied between the use of these elements and their effect on problems in the award phase. In the next chapter, the indicators of 'problems in the award phase with regard to the award method' are outlined.

4. Problems (Lawsuits and Questions addressed by the Committee)

In the context of public procurement, it occurs that dissatisfaction arises due to certain procedures. This dissatisfaction can result in complaints, filed by involved stakeholders (CA's or suppliers). In this research, complaints serve as indicators for 'problems'. However, not all complaints are reported and are therefore, not available for analysis. Overall, there may be many complaints, ambiguities or problems within procurement procedures that are not indicated by involved parties for a number of reasons (e.g. minor problems; maintain good business relationships; financial resources; financial scope public tender; etc). Therefore, the category 'all complaints' covers more than complaints which are submitted and detectable.

According the Lohman, Manuza and Telgen (2013) there are five possibilities for suppliers to complain about procurement procedures. These are layered and will be indicated below, ranking from most accessible to least accessible. Complaints can be directed to the following authorities:

1. The contracting authority, by means of the memorandum of information (MOI, Nota van Inlichtingen; NvI);
2. The contracting authority, by means of the complaints procedure;
3. The Commission of Procurement Experts;
4. The national court (District court);
5. The European court.

As discussed previously (paragraph 2.3.1), just two indicators for 'problems' were selected. Namely; questions that were addressed by the Commission of Procurement Experts and lawsuits (numbers 3 and 4 in the list above). Number 1 and 2 (the MOI and complaints procedure) are not selected for a number of reasons. First of all, the MOI provides a limited source of information on problems, not all MOI include reports of complaints, and these complaints are not always published in the tender documents (e.g. all questions are addressed in the MOI). In addition, the MOI is not a strong indicator for problems since the threshold to file complaints is low. Secondly, complaints which are submitted in the complaints procedure are not included in the tender documents and therefore hard to obtain. The last enlisted (5), 'the European court' is also not used as an indicator for problems, since, it is beyond the scope of this thesis, only procurements are taken into account by the European court that have resulted in a national lawsuit (list item 4). Thus, these tenders are already included in the indicator 'national court - lawsuits'. In the following sections, both measures for problems used for this research are discussed.

4.1 Commission of Procurement Experts

One of the indicators for 'problems' is a question addressed by the Commission of Procurement Experts. Such questions indicate that the supplier has a complaint about the procurement procedure. In this study, we assume that award methods that resulted into questions are different from award methods that did not. Note; we discuss the situation in the Netherlands.

By the introduction of the Public Procurement Act on 1 April 2013, an independent Committee of Procurement Experts was established. This committee is appointed by the Minister of Economic Affairs to improve the quality of public procurement in the Netherlands (art. 4.27 Aanbestedingswet

2012).¹² The main duties of the Committee are to mediate between suppliers and government as well as to address complaints in connection with a tender. The Committee provides both suppliers and contracting authorities with non-binding advice in response to complaints regarding procurement procedures.

The method of the Committee is regulated in the standard for complaints in public procurement.¹³ Any supplier who participates in a procurement process and finds that procurement rules were not properly applied, or that the contracting authority treated them unfairly, may submit a complaint to the committee. Conversely, contracting authorities may complain about the behavior of participating supplier companies. In this study, the focus is on complaints submitted by suppliers (with regard to the award method).

The Committee only addresses complaints if suppliers complied along the following steps:

1. If a supplier has informed the contracting authority in clear terms about the complaint and;
2. has allowed the contracting authority reasonable time – within the procurement procedure – to respond to the complaint and;
3. the reaction of the contracting authority has not led to the withdrawal of the complaint.¹⁴

If all above requirements are met, the Committee has the opportunity to address the complaint (not mandatory). This indicates that suppliers need to comply with certain requirements and that complaints are not simply all taken into account by the Committee. The Committee publishes their recommendations on their website after the procurement procedure is completed and finally awarded. In addition, the involved parties, but remain anonymous.

The recommendations of the Committee are non-binding, though they are substantial. This means that contracting authorities and judges are not bound to advices of the Committee. A judge can take this advice in consideration before issuing the final judgment.¹⁵ The recommendations included in this research relate all to technical purchasing aspects; elements with regard to the award method. This indicator for ‘problems’ is chosen since it reflects serious complaints with regard to the award method. Furthermore, suppliers need to take comply with the above requirements to eligible for the Committee.

Important to note; the Commission of Procurement Experts was installed on 1 April 2013, while the research time frame covers the period between 1 May 2010 and 31 April 2014. Therefore, ‘problems’ are mainly indicated with the measure ‘lawsuits’. Overall, the differences between lawsuits and question addressed by the Committee are not expected to cause difficulties, since ‘questions addressed by the CoPE’ is solely a measure capturing the dependent variable ‘problems’.

4.2 Lawsuits

The other measure for the dependent variable ‘problems’ are lawsuits. Lawsuits demonstrate that a particular supplier has a claim with regard to a procurement procedure. In the selected lawsuits, the used award method is not complying with the Public Procurement Act and one or more European

¹² “Instellingsbesluit Commissie van Aanbestedingsexperts”

¹³ “Standaard Klachtenafhandeling bij Aanbesteden”

¹⁴ Standaard Klachtenafhandeling bij Aanbesteden – deel 2 – Art. 6.

¹⁵ Standaard Klachtenafhandeling bij Aanbesteden – deel 2 – Art. 1.

procurement principles, according to the supplier. In this case, a supplier accuses the contracting authority of insufficient transparency and/or objectivity in the award phase. This could mean the performance of the supplier lodging for appeal is not properly assessed, or that the award method contained errors. In the RfQ and the tender documents, specific information on deadline(s) and procedures for lodging appeals are usually published.

The judicial system is divided into 11 districts, each with its own court. The District court is the first court where suppliers can lodge for appeal (in Dutch: kort geding bij de voorzieningenrechter). In a lawsuit with regard to the award method, appeal must be lodged before the registration deadline of the procurement. It is also possible to lodge for appeal after the notification of the outcome of the awarding procedure of the provisional award or of the rejection of it (Lohman, Manuza and Telgen, 2013).

Every supplier that feels unjustly sanctioned with the used award methods of the contracting authority, is able to lodge for appeal. However, to successfully lodge for appeal, locus is required. This means that the supplier needs to issue a substantial claim (e.g. often claimed are obtaining the award or a getting a second chance by means of re-tendering). Depending on the grounds for lodging for appeal different claims can be issued. In case a supplier complains about the award method, this usually implies the claim of a retender (Lohman, Manuza and Telgen, 2013).

It is also possible to lodge for appeal against the judgements passed by the district court. In these cases appeals can be lodged at the competent Court of Appeal. In the last possible option, appeals in cassation, public procurement cases can be lodged at the Supreme Court of the Netherlands. However, in the sample of this research no distinction is made in either number or levels of lawsuits, because the subject of this study is the effect of the 10 identified elements.

The lawsuits analyzed for this research relate to technical purchasing aspects. To be more specific, to the award method in the award phase. This indicator for 'problems' is selected as it reflects serious problems with regard to the award method.

4.3 Conclusion

In this chapter both measures for 'problems' are discussed; (1) procurement that resulted in a lawsuit and (2) procurement that resulted in questions addressed by the Commission of Procurement Experts. These two indicators for problems serve as dependent variables and are selected for a number of reasons. Firstly, both measures are serious steps for suppliers to complain and, consequently, to express their disagreement to the used award method. Secondly, suppliers need to complete certain procedures to be taken into consideration (CoPE) or risk financial resources in case of a lawsuit, both indicating serious complaints. Thirdly, because both measures are both visible and publicly accessible. And last, but not least, these measures can be linked to RfQ's in the used database (ABK, TED and Tendered).

5. Results

This chapter provides answers to the second sub-question; ‘Are there differences in the degree to which the 10 elements are present in non-problematic and problematic procurements?’, as well as, to the third sub-question: ‘Is there a causal relationship between the identified elements in the award phase and problems in public procurement?’. Firstly, the occurrence of all involved elements separately, are analyzed and presented in paragraph 5.1. Secondly, the number of jointly applied elements is analyzed for both problematic and non-problematic procurements, in paragraph 5.2. At last, the involved hypotheses are separately addressed in paragraph 5.3. This chapter concludes by answering sub-questions two and three.

5.1 Presence of elements

Here, we present an overview of the number of times the elements are present in the stratified random sample and problematic sample. In Table 6, an overview is presented of the distribution between award criteria (Lowest price and EMAT). Within the EMAT criterion the problematic elements are often present.

Type award method	Lowest price	EMAT
Group 1	218 (54,2%)	184 (45,8%)
Group 2	2 (4,7%)	41 (95,3%)
Total	220 (49,4%)	225 (50,6%)

Table 6: distribution award criterion

Table 7 shows that the certain elements are frequently occurring in the award phase. This information allows to examine a possible causal relationship between the presence of certain elements and problems. For example, if certain elements are frequently occurring in the award phase, and have not resulted in problems, the suggested causal relationships might not be present. However, these elements may still be more susceptible to problems.

Presence of elements (In award criterion EMAT)	Stratified random sample (184 cases)		Problematic procurements (41 cases)	
	Count	Percentage	Count	Percentage
1. Making use of the formal method WFS	141	76,6%	37	90,2%
2. Sub-criteria – performance levels (constructed)	344/551	62,4%	82/128	64,0%
3. No publication of the weights	2	1,1%	3	7,3%
4. No publication of the score method	17	9,2%	8	19,5%
5. Form of the score graph is hollow (price)	53	28,8%	9	22,0%
6. Relative award methods	123	66,8%	34	82,9%
7. Scoring on the basis of ranking	16	8,7%	2	4,9%
8. ‘Flat’ award criteria	86	46,7%	31	75,6%
9. Lower boundaries	22	12,0%	7	17,1%
10. Sub-price criteria for award criterion price	61	33,2%	21	51,2%

Table 7: negative impact of elements

5.2 The number of jointly applied elements

Here, we present an overview on the number of applied elements in the award phase and their susceptibility to problems. As discussed previously (paragraph 2.3.2), we expect procurements to become more prone to problems when any of these elements are present in the award phase.

Inherently, award phases are expected to become more susceptible to problems when elements are increasingly present. This section aims to examine the latter, by doing so, a general analysis is provided in Table 8.¹⁶

Problems	Number of applied elements (in award criterion EMAT)								Total
	0	1	2	3	4	5	6	7	
No	17 (100%)	17 (94,4%)	28 (82,4%)	58 (86,6%)	37 (77,1%)	26 (72,2%)	1 (33,3%)	0 (0,0%)	184 (81,8%)
Yes	0 (0%)	1 (5,6%)	6 (17,6%)	9 (13,4%)	11 (22,9%)	10 (27,8%)	2 (66,7%)	2 (100%)	41 (18,2%)
Total	17 (100%)	18 (100%)	34 (100%)	67 (100%)	48 (100%)	36 (100%)	3 (100%)	2 (100%)	225 (100%)

Table 8: Presence of elements and their impact on problems

Table 8 shows, that the overall number of applied elements varies from 0 to 7 in a single procurement. For non-problematic procurements this varies from 0 to 6 applied elements, for problematic procurements this varies from 1 to 7 applied elements. In addition, the percentages are decreasing for non-problematic procurements while the percentages are increasing for problematic procurements. This indicates that the likelihood of problems increases when the identified elements are increasingly applied. Thus, award phases which contain a high number of elements are more susceptible to problems than those containing less elements.

The result from Table 8 are displayed in Figure 6, to allow a simple interpretation of the results.

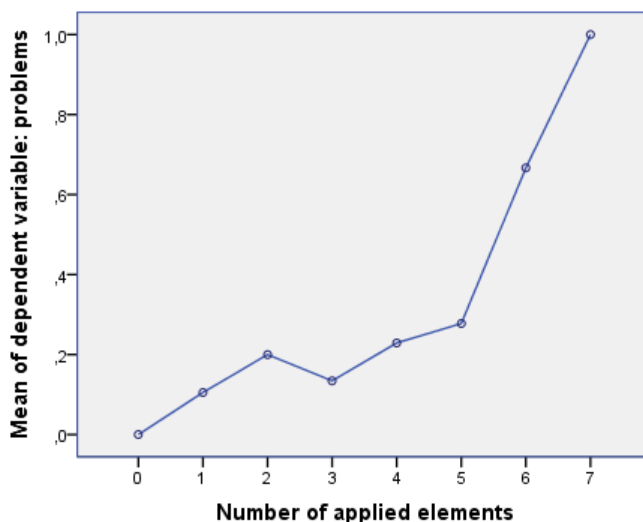


Figure 6: The number of applied elements in their susceptibility to problems

Important to note; the number of procurement including 6 of 7 elements is small, therefore, we need to interpret these findings with caution. Furthermore, this analysis has examined the impact of the presence of multiple elements on their effects on problems in a single procurement. However, it is unknown which of the identified elements are most harmful, this is examined in the following section.

¹⁶ The analysis used for Table 8 includes nine elements. Element 2 – sub-award criteria and their performance levels (paragraph 3.1.2) are not included in this analysis. Element 2 entails multiple variables and categories. Therefore, we are unable to indicate this element with a single variable, making this element not suitable for this analysis. Element 2 is separately addressed in paragraph 5.3.

5.3 Hypothesis

In the prior analysis (Table 8) it became evident that an increased number of jointly applied elements (the presence of multiple elements in a single procurement) are occurring more often in problematic procurements than in non-problematic procurements. However, it is unknown which of the identified elements are most harmful. Therefore, the following analyses are intended to identify which of these 10 elements are significantly related to problematic procurements. The analyses are used to examine whether the identified elements are directed towards problems.

In chapter 3.1.1 – 3.1.10 all hypotheses are formulated. Here, the hypotheses are discussed in terms of their scores on the different analysis. Table 9 offers an overview of all hypotheses with their respective analyses. Important to note; within each hypothesis different analyses are provided, as several subjects are at stake. Furthermore, each hypothesis is analyzed for the ‘contract subjects’ works, goods and services to examine if a negative impact of elements is contingent on contract subject. As discussed previously (paragraph 2.3.2), we expect ‘problems’ when at least one of the 10 identified elements is present in procurements. The latter involves a more qualitative interpretation of the results. In Appendix H, a complete overview is presented of all contingency tables (in Dutch: kruistabellen) and statistical tests used for Table 9.

Hypothesis 1-10	Z-score	N	Missing values	2-sided P-value	1-sided P-value	Significant
Hypothesis 1 – Award criterion						Overall: YES
Analysis 1 – Overall	1,632	219	6	0,103	0,052	NO
Analysis 2 – Overall	1,897	219	6	0,058	0,029	YES
Analysis 3 – Works	0,318	40	0	0,751	0,376	NO
Analysis 4 – Goods	1,352	61	3	0,176	0,088	NO
Analysis 5 – Services	2,014	118	3	0,044	0,022	YES
Hypothesis 2 – Sub-criteria						Overall: NO
Analysis 6 – Frequencies						
Analysis 7 – Overall (2x3 table)	0,161	225	0	0,987	-	NO
Hypothesis 3 – Weights						Overall: YES
Analysis 8 – Overall	2,430	223	2	0,015	0,008	YES
Analysis 9 – Works	2,026	40	0	0,042	0,021	YES
Analysis 10 – Goods	2,492	64	2	0,013	0,007	YES
Analysis 11 – Services	0,575	119	0	0,565	0,283	NO
Hypothesis 4 – Score method						Overall: YES
Analysis 12 – Overall	1,985	220	5	0,047	0,024	YES
Analysis 13 – Works	0,901	40	0	0,368	0,184	NO (direction)
Analysis 14 – Goods	1,422	63	1	0,155	0,078	NO
Analysis 15 – Services	2,525	117	4	0,012	0,006	YES
Hypothesis 5 – Form graph						Overall: NO
Analysis 16 – Overall (2x3 table)	2,262	188	37	0,163	-	NO
Analysis 17 – Overall	0,596	188	37	0,551	0,276	NO
Analysis 18 – Works	0,318	26	14	0,750	0,375	NO
Analysis 19 – Goods	1,817	58	6	0,069	0,035	NO (direction)
Analysis 20 – Services	0,228	104	17	0,820	0,410	NO

Hypothesis 6 – Relativity						Overall: YES
Analysis 21 – Overall (2x3 table)	2.046	210	15	0,123	-	NO
Analysis 22 – (Kendall's tau-c) ¹⁷	0,108	210	15	-	0,039	YES
Analysis 23 – Overall	1.978	210	15	0,048	0,024	YES
Analysis 24 – Works	1,668	39	1	0,095	0,048	YES
Analysis 25 – Goods	1,474	58	6	0,141	0,071	NO
Analysis 26 – Services	0,872	113	8	0,383	0,192	NO
Analysis 27 – 29 (Appendix H 2x3)						
Hypothesis 7 – Ranking						Overall: NO
Analysis 30 – Overall	0,862	214	11	0,389	0,170	NO (direction)
Analysis 31 – Works	0,221	40	0	0,825	0,413	NO
Analysis 32 – Goods	0,221	62	2	0,678	0,339	NO
Analysis 33 – Services	0,989	112	9	0,322	0,161	NO
Hypothesis 8 – ‘Flat’ criteria						Overall: YES
Analysis 34 – Overall	3,723	208	17	0,000	0,000	YES
Analysis 35 – Works	1,919	40	0	0,055	0,028	YES
Analysis 36 – Goods	2,417	59	5	0,045	0,023	YES
Analysis 37 – Services	2,346	109	12	0,019	0,010	YES
Hypothesis 9 – Boundaries						Overall: NO
Analysis 38 – Overall	0,941	211	14	0,355	0,168	NO
Analysis 39 – Works	1,383	40	0	0,167	0,084	NO
Analysis 40 – Goods	1,832	60	4	0,067	0,034	YES
Analysis 41 – Services	1,383	111	10	0,427	0,214	NO (direction)
Hypothesis 10 – Sub-criteria						Overall: YES
Analysis 42 – Overall	2,047	220	5	0,041	0,021	YES
Analysis 43 – Works	0,286	38	2	0,774	0,387	NO (direction)
Analysis 44 – Goods	3,399	63	1	0,002	0,001	YES
Analysis 45 – Services	0,808	119	2	0,566	0,283	NO

Table 9: Overview 10 hypotheses

In Table 9, all hypotheses and their respective analyses are presented. Within each element a one-sided hypothesis is formulated, since we expect that the 10 identified elements are to a larger extent present in problematic procurements than in non-problematic procurement. This means that accepting or rejecting the hypotheses is related to the 1-sided P-value. For statistical evidence of accepting the hypothesis, the P-value needs to be lower than $\alpha = 0.05$. However, the 1-sided P-value can only be used for 2x2 contingency tables; with the use of 2x3 tables, certain assumptions are not fulfilled. Therefore, the 2-sided P-value is used for 2x3 tables. In the following sections, the results are discussed for each hypothesis, separately.

Hypothesis 1: *The Weighted Factor Score (a form of price-to-quality scoring) occurs more often in problematic procurements than in non-problematic procurements.*

In analysis 1, the results indicate a strong direction of price-to-quality scoring towards problems. However, the results are not significant ($z=1,632$, $n=219$, $p=0,052$),¹⁸ indicating that price-to-quality scoring is remotely related to problems. In analysis 2, the data is recoded; the award method

¹⁷ Kendall Tau-c: a statistical technique designed to assess how close the relationship between two variables is to being monotone. A monotone relationship is one that constantly increases or decreases, but not in a linear fashion (De Veaux et al., 2012). This analysis provides information on the direction of relative award methods. More relative award methods are increasingly present in problematic procurements.

¹⁸ Appendix H – Analysis 1: (WFS+VFM) and (Fixed-price+AOV)

'Weighted Factor Score' is compared to the award methods 'Value for Money, Fixed-price and Awarding on Value' and their relations to the dependent variable 'problems'. The second analysis is significant ($z=1,832$, $n=219$, $p=0,029$).¹⁹ Thus the WFS is significantly related to problems. The WFS has a probability of 21% to result in problems, compared to a 10% probability of the other award methods (VFM, F-P and AOV). This 11% difference is significant, therefore, the WFS is more related to problems than other award methods.

To see whether this holds for all three 'contract subjects' (works, goods and services), three additional analysis are provided. Their results show that only the WFS used in services is related to problems ($z=1,76$, $n=118$, $p=0,022$).²⁰ The other two contract subjects (goods and services) are not significantly related to problems. However, contract subject goods is directed towards problems ($z=1,352$, $n=61$, $p=0,088$). In addition, the N of works and goods is small in comparison to that of services. This means that there is a smaller amount of works and goods in the sample compared to services, affecting statistical power. Overall, hypothesis 1 is significant, this means that the WFS is occurs more often in problematic procurements than in non-problematic procurements. Specifically, this only applies to services.

Hypothesis 2: *Qualitative – constructed performance measures – are to a larger extent present in problematic procurements than in non-problematic procurements.*

The frequencies of all performance levels concerning hypothesis 2 are presented in analysis 6. In this analysis, natural performance levels score low, while constructed performance levels score high. Important to note; the number of applied criteria may be different in procurements, which affects the number of applied performance measures and thus, the overall score. This means that procurements with more sub-criteria are likely to score higher. Therefore, natural performance levels are assigned a zero while proxy and constructed performance levels are assigned a one and a two. Furthermore, the original frequencies are recoded into three categories to enable the comparison of overall performance levels.

The results in analysis 7 indicate that the level of performance measure is not directed to problematic procurements. This means that performance levels have are not related to problems in procurements. Consequently, the hypothesis is not significant ($z=0,161$, $n=225$, $p=0,987$)²¹. This hypothesis is also examined over the three contract subjects. However, since they appeared al non-significantly these analyses are not included in Appendix H. Overall, hypothesis 2 does not hold true. The level of performance measures is not related to problems.

Hypothesis 3: *not publishing the weighing in tender documents leads to the outcome 'problems'.*

The results of analysis 8 indicate that not publishing the weights is related to problematic procurements. The data presents that in 60% of the cases where weights were not indicated, problems have occurred. In procurements where weights are announced, the probability of problems is 17%. The difference of 43% is significant ($z=2,430$, $n=233$, $p=0,008$).²² This also applies to the

¹⁹ Appendix H – Analysis 2: (WFS) and (VFM+Fixed-price+AOV)

²⁰ Appendix H – Analysis 3-5: (WFS) and (VFM+Fixed-price+AOV) – Contract subject

²¹ Appendix H – Analysis 7: Performance levels

²² Appendix H – Analysis 8: Publication weights (overall)

contract subjects works ($z=2,026$, $n=40$, $p=0,021$) and goods ($z=2,492$, $n=64$, $p=0,007$). However, no statistical evidence can be found for services ($z=0,575$, $n=119$, $p=0,283$).²³

These tables indicate that the 'stratified random sample' (group 1) contains two cases in which the weights are not published. These cases occurred in the contract subject services. In the problematic procurements (group 2), three cases are present in which the weighing is not published, each contract type contains one procurements without indication of weights. This means that there are very few cases in which weights are not published. In the sample, two services are present where weights are not announced, this affects analysis 11.²⁴ Therefore, the overall hypothesis holds true, not publishing the weights results in problems, regardless the contract subject.

Hypothesis 4: *No publication of the score method occurs more often in problematic procurements than in non-problematic procurements.*

In analysis 12 the results indicate a strong direction towards problems, inherently, the overall analysis is significant ($z=1,985$, $n=220$, $p=0,024$).²⁵ This means that not publishing the score method is related to problematic procurements. In cases where the score method is not announced, 32% of the procurements resulted in problems, this is 16% in cases where the score method is announced. It follows from the analysis that this 16% difference is significant.

To see whether this holds for all three 'contract subjects', additional analysis were conducted. The results show that, not announcing the score method is significantly related to problems only for contract subject services ($z=2,525$, $n=117$, $p=0,006$). Contract subject goods is strongly directed towards problems. In 25% of the cases where the score method is not announced problems are present. Compared to this is 10% in cases where the score methods is announced. However, this difference is not significant ($z=1,422$, $n=63$, $p=0,078$). For contract subject works, the results are not significant ($z=0,901$, $n=40$, $p=0,184$).²⁶ This does not necessarily imply that for works the score method is always published. In the stratified random sample (group 1) three cases are present in which the score method is not indicated. In the problematic sample (group 2), there are no cases present in which the score method is not published, which affects the direction towards problems. Overall, hypothesis 4 holds true for services, this means that no publication of the score methods is significantly more present in problematic procurements than in non-problematic procurements. By means of a more qualitative interpretation of the results of the analyses, hypothesis 4 holds true regardless contract type.

Hypothesis 5: *Non-linear score graphs, especially hollow score graphs for the award criterion price, occur more often in problematic procurements than in non-problematic procurements.*

In analysis 16, the results indicate that different forms of score graphs are not related to problems. Theoretically, linear score graphs are not expected to result in problems, though concave forms are expected to result in problems (paragraph 3.1.5). However, the data indicated that 16% of procurements with the linear score graphs resulted into problems, while this is 14,5% for concave

²³ Appendix H – Analysis 9-11: Publication weights – Contract subjects

²⁴ Data on contract subjects – Works: 1 case without the weights (0% / 100%). Goods: 1 case without weights (0% / 100%). Services: 1 case without weights (66% / 33%).

²⁵ Appendix H – Analysis 12: Publication score method (overall)

²⁶ Appendix H – Analysis 13-15: Publication of the score methods – Contract subjects

score graphs. This means that the results not significant ($z=2,262$, $n188$, $p=0,163$).²⁷ In analysis 17, the data is recoded in two categories (linear and other forms) and compared to each other in their relation to 'problems'. The results are also not significant ($z=0,596$, $n188$, $p=0,276$).²⁸

To examine whether this also applies to all three contract subjects, additional analysis are provided.²⁹ Their results indicate that different forms of the score graphs are not related to problems in either of the three contract subjects (see overview). Overall, hypothesis 5 is not significant, this means that different forms of score graphs are not significantly occurring differently in problematic procurements than in non-problematic procurements. In addition, this is the case for all three contract subjects.

Hypothesis 6: *Relative score methods occur more often in problematic procurements than in non-problematic procurements.*

In analysis 21, the results indicate a strong direction of the level of relativity towards problems. The data indicate that absolute score methods have a probability of 9% to result in problems (i.e. absolute score methods do not contain relative elements). In score methods where a single criterion (price or quality) contains relative score methods, the probability of problems is 21%. In the last option, both price and quality are scored with relative methods, the data indicated that 24% of these cases have resulted in problems. The data indicates that award methods which are to a larger extent relative are occurring more often in problematic procurements. However, analysis 21 is not significant ($z=2,046$, $n210$, $p=0,123$).³⁰ Because more relative score methods are occurring more often in problematic procurements, the direction is examined with a statistic measuring the linearity of the relationship. The direction of the level of relativity is significant, this means that more relative score methods are occurring more frequently in problematic procurements than in non-problematic procurements (Kendall tau-c, $z=0,108$, $n210$, $p=0,039$).³¹

In analysis 23, the data is recoded in two categories, namely; (1) relative score method and (2) absolute score methods. Absolute score methods have a probability of 9% to result in problems while relative score methods have a probability of 22% to result in problems. It follows from the analysis that this 13% difference is significant ($z=1,978$, $n210$, $p=0,024$).³² To examine whether this also applies to all three contract subjects, additional analysis are provided. For all three contract subjects, relative score methods are occur more often in problematic procurements than in non-problematic procurements. However, only for contract subject works this difference is significant ($z=1,668$, $n49$, $p=0,048$).³³ By means of a more qualitative interpretation of the results of analyses 24-29³⁴, it is indicated that problematic procurements contain more relative award methods, regardless contract subject. Overall, hypothesis 6 is significant.

Hypothesis 7: *Methods which award scores on the basis of ranking occur more often in problematic procurements than in non-problematic procurements.*

²⁷ Appendix H – Analysis 16: Linear, convex, concave and not published

²⁸ Appendix H – Analysis 17: Recode (linear) and (convex+concave+not published)

²⁹ Appendix H – Analysis 18-20: Recode (linear) and (convex+concave+not published) – Contract subject

³⁰ Appendix H – Analysis 21: Recode (absolute) / (only price is relative + only quality is relative) / (both relative)

³¹ Appendix H – Analysis 22: Kendall's tau-c (direction relativity)

³² Appendix H – Analysis 23: Recode (absolute (both relative+ 1 of 2 relative)

³³ Appendix H – Analysis 24-26 and 27-29: Absolute and relative – Contract subject

³⁴ Appendix H – Analysis 24-26 and 27-29: Absolute and relative – Contract subject

In analysis 30, the results indicate that scores on based on ranking are not related to problems. Scores based on ranking do not occur differently in problematic procurements. The first analysis is not significant ($z=0,862$, $n214$, $p=0,170$).³⁵ In addition, the results show the opposite of the expected direction. In problematic procurements, 19% of all non-ranking scores have resulted in problems, while this is 11% for scores based on ranking. This is also the case for all three contract subjects. The additional analysis, provide evidence that scores bases on ranking are not significantly related to problems in either works, goods or services (see overview Table 9).³⁶

By means of a more qualitative interpretation of the results, it becomes clear that scores on basis of ranking are not frequently occurring in public procurement. Furthermore, all analysis provide evidence that scores on the basis of ranking is not significantly related to problems. This provides evidence for rejecting hypothesis 7.

Hypothesis 8: *Flat award criteria occur more often in problematic procurements than in non-problematic public procurements.*

In analysis 34, the results indicate that ‘flat’ award criteria are more present in problematic procurements than in non-problematic procurements. This does not necessarily imply that flat award criteria will result in problems. In the stratified random sample, 86 procurements contain flat award criteria, these have not resulted in problems. However, the distribution of flat award criteria is significantly different in problematic procurements. This means that the presence of flat award criteria is directed towards problems. In problematic procurements, 7%% of all non-flat scores have resulted in problems, this is 27% for flat scores. It follows from the analysis that this 20% difference is significant ($z=3,723$, $n208$, $p=0,000$)³⁷.

The similar holds for the three contract subjects; works ($z=1,919$, $n40$, $p=0,028$), goods ($z=2,417$, $n59$, $p=0,023$) and services ($z=2,346$, $n109$, $p=0,010$).³⁸ For contract subject works, 6% of all procurements with non-flat criteria resulted in problems, this is 30% for procurement with flat award criteria. For goods (0% non-flat and 21% flat) and services (10% non-flat and 28% flat) this is similar. Overall, this means that procurements with flat award criteria are three to four times more likely to result in problems. Overall, hypothesis 8 holds true, this means that flat award criteria are occurring more frequently in problematic procurements than in non-problematic procurements.

Hypothesis 9: *Extra lower boundaries occur more often in problematic procurements than in non-problematic procurements.*

In analysis 38, the results indicate a strong direction towards problems. This means that the use of lower boundaries in more present in problematic procurements than in non-problematic procurements. The data indicates that procurements with lower boundaries have a probability of 24% to result in problems, this is 17% for procurements which make use of the total rating scale. However, this 7% difference is not significant ($z=0,941$, $n211$, $p=0,168$).³⁹ In the contract subjects works and goods the data indicates a similar trend, the results show that lower boundaries are more

³⁵ Appendix H – Analysis 30: Scoring on basis of ranking

³⁶ Appendix H – Analysis 31-33: Scores on the basis of ranking – Contract subject

³⁷ Appendix H – Analysis 34: Flat award criteria

³⁸ Appendix H – Analysis 35-37: Flat award criteria – Contract subject

³⁹ Appendix H – Analysis 38: Extra lower boundaries

occurring more often in problematic procurements. However, this is only significant for contract subject goods ($z=1,832$, $n60$, $p=0,034$).⁴⁰ For contract subject services, the use lower boundaries are not directed towards problems in procurements.

Overall, hypothesis 9 does not hold true, this means that the use of lower boundaries is not significantly present in problematic procurements. However, when lower boundaries are present in the award phase, procurements are directed towards problems in the overall analysis and for contract subjects works and goods.

Hypothesis 10: *An award criterion price that consist of various financial sub-criteria occurs more often in problematic procurements than in non-problematic procurements.*

In analysis 42, the results show that procurements including financial sub-criteria are strongly directed towards problems. Consequently this analysis is significant ($z=2,047$, $n220$, $p=0,021$).⁴¹ This means that the use of financial sub-criteria is significantly occurring more often in problematic procurement than in non-problematic procurements.

In contract subjects goods and services the data indicates a similar trend, the results show that the use of financial sub-criteria are directed towards problems. However, the results only show that financial sub-criteria used for goods are significantly related to problems ($z=3,399$, $n63$, $p=0,001$). Financial sub-criteria in works and services are not significantly related to problems. However, contract subjects services is slightly directed towards problems ($z=0,808$, $n119$, $p=0,283$).⁴² Overall, hypothesis 10 holds true, this means that financial sub-criteria are occurring more often in problematic procurements than in non-problematic procurements. Specifically, this only applies to contract subject goods.

5.4 Conclusion

In this chapter, the hypotheses that we are trying to support are tested. The results of each formulated hypotheses are separately addressed, in certain elements also qualitative information from the contingency tables is used. By means of an overview of all tables and analyses (see Tables 6-9) we are able to answer the second and third sub-questions.

First, the second sub-question is answered; *'Are there differences in the degree to which the 10 elements are present in non-problematic and problematic procurements?'*. An answer to this question is not uniform since certain elements are distributed differently in problematic and non-problematic procurements. This means, that only the significant elements in overview Table 9 are more present in problematic procurements. Important to note; the elements are not always significant for all subdivided contract subjects. However, the result are interpreted more qualitative. The following elements are more present in problematic procurements; the WFS, no publication of the weights or score method, relative award methods, 'flat' award criteria and the use of financial sub-criteria. The presence of lower boundaries in the award phase is also more present in problematic procurements, hence, this difference is not significant. The other indicated elements; constructed performance levels, the form of the score graph and scoring on basis of ranking are not significantly more present in problematic procurements. These are also not directed towards problematic procurements.

⁴⁰ Appendix H – Analysis 39-41: Extra lower boundaries – Contract subject

⁴¹ Appendix H – Analysis 42: Price sub-criteria

⁴² Appendix H – Analysis 43-45: Price sub-criteria – Contract subject

Sub-question two addressed the distribution of the identified elements between the two sub-groups. It is concluded that certain elements are more present in problematic procurements. However, this question did not involve the causal relationship between the negative impact of elements and problems. This is addressed in sub-question three.

Sub question three gives answers to the following question; *'Is there a causal relationship between the identified elements in the award phase and problems in public procurement?'.* To be able to answer this question the stratified random sample and problematic sample are compared on basis of percentages in the problematic categories (i.e. presence or exclusion of certain elements, negative impact of an element). This is examined since certain elements are frequently occurring in procurement, often these do not result in problems (see Table 7). For this reason we are unable to conclude that there is a causal relationship between elements and problems.

In two identified elements the percentages are strongly directed towards problems; element 3 (weights) and element 4 (score method). In case the weights are not indicated the probability of problems is 60%. This means that in 40% of the procurements not indicating the weights did not result in problems. In case the score method is not indicated the probability of problems is 32%. In the other significant elements the probability of problems varies from 21% to 27%.

6. Conclusion and discussion

In this final chapter, the overall conclusions of this study are presented: the answers to the main research question are provided. In addition, suggestions for future research and limitations are presented in section 6.2 and 6.3. This chapter closes with recommendations and implications for policy makers.

6.1 Conclusion

The rationale for this research stemmed from the available leeway which contracting authorities have to give their interpretation to the award phase. This leeway exists as there is few legislation in European and Dutch public procurement law with regard to the award phase. A consequence of this leeway is that contracting authorities have the right to determine which additional decision-making criteria apply to a public tender and how these criteria are combined, weighed, scored and measured. As a result, many contracting authorities, especially when trying to identify the EMAT, struggle with the pressure to explain their choices and, therefore, rely on some form of formal method. However, the practical use of formal decision methods is not without problems, since there are many various formal methods and elements which can be considered, and the effects of these aspects and methods are not always known. In some cases, the values and preferences of the awarding mechanism are not transparent or objectively measured, while in other cases, the award method results in undesired and/or unforeseen effects. These occurrences have the possibility to make the scoring mechanism opaque for suppliers, and, consequently, problems can arise.

In order to make a contribution to the process and practices within the available leeway of the award phase, 10 frequently occurring elements are selected, which are susceptible to problems. Theoretically, these 10 elements are expected to result in problems, from both legal and economic perspectives. In this study, it is determined whether the negative impact (i.e. presence or exclusion) of these 10 elements is empirically related to problems. The overall expectation is that problems when a single or more of these 10 elements are present in the award phase. To examine this two groups of procurements were studied: (1) a random sample of 402 Dutch public procurements and (2) a sample of 43 Dutch public procurements which all resulted in problems with regard to the award method. These groups are compared to each other with regard to their degree of identified elements and regarding the presence of 'problems' to be able to identify which elements are related to problematic procurements.

The empirical findings in this study indicate that six identified elements are significantly more present in problematic procurements than non-problematic procurements. These significant elements are shortly discussed;

Element 1: The Weighted Factor Score

The WFS makes an accurate representation of the procurer's preferences difficult. Consequently, it is often open to strategic manipulation and tends to impose non-linearity in bid prices. The results indicate the WFS is significantly more present in problematic procurements. Moreover, the WFS has a probability of 21% to result in problems, while for other indicated award methods (i.e. ; VFM, GOW and Fixed-price) this is 10%. This implies that the WFS, a form of price-to-quality scoring used for identifying the EMAT, is empirically more susceptible to problems than other award methods.

Element 3: No publication of overall weights

Not indicating the weights results in uncertainty on the impact of further award criteria (price-quality ratio). Suppliers often have the possibility to offer multiple bids. In cases where the procurer's preferences are unclear, the best offer might then not be submitted. The results indicate that not publishing the weights is significantly related to problematic procurements. In procurements where the weights have not been indicated, the probability of problems is 60%, while this is 17% for procurements where the weights are known.

Element 4: No publication of the score method

In order to determine scores on an award criterion, a score method is used. The omission of a score method can lead to unexpected and unintended results, since different methods can result in different winners. The findings indicate that no publication of the score method is significantly more present in problematic procurements. Furthermore, in cases where the score method is not indicated, the probability of problems is 32%. This is 16% in cases where the score method is indicated. This means that award methods which lack a score method are more prone to problems than methods in which the score method is provided and, consequently, can be identified by suppliers.

Element 6: Relative award methods

In award phases where an assessment of a involves the best, worst and/or average supplier bids, the scoring rule becomes opaque to suppliers. By using relative award methods, the procurer's preferences are not clear, because the importance of differences in performance are dependent on other suppliers. The findings support this theory by indicating that a high level of relativity is directed towards problems. In cases where a scoring rule becomes more relative, the likelihood of problems increases. Relative score methods have a probability of 22% to result in problems, while absolute score methods have a probability of 9% to result in problems.

Element 8: 'Flat' award criteria

A 'flat' award criterion reduces the contribution of an award criterion and, therefore, leads to different distributions of weights between award criteria. This results in an unintended effect, since other award criteria gain greater importance. Moreover, the weights will not make the bid evaluation transparent, nor predictable, in case of poorly specified rating scales. The empirical findings support theory; 'flat' award criteria are significantly more present in problematic procurements. In case 'flat' award criteria are used the probability of problems is 27%, this is 7% for non-flat award criteria. General, this means that procurements with 'flat' award criteria are three to four times more likely to result in problems, and thus more susceptible to problems.

Element 10: Sub-criteria for award criterion price

In case of financial sub-criteria, multiple criteria are included in the awarding mechanism. All of these need to be weighed in a single cost measure. In case award criterion price is subdivided in subcategories, strategic tendering is evoked. The results indicate that financial sub-criteria for award criterion price to a significant extent present in problematic procurements. Moreover, by using financial sub-criteria the probability of problems is 26%, while the probability is 15% when financial sub-criteria are not present. Important to note; the above only applies for contract subject goods.

The other four identified elements are not significantly present in problematic procurements. However, the presence of lower boundaries (element 9) is directed towards problems. This implies

that lower boundaries appear to be more present in problematic procurements, only not significantly. The other three elements are not directed towards problems. Based on the empirical findings, the following elements are not distributed differently in problematic procurements. These concern; sub-criteria performance levels (element 2), form of the score graph (element 5) and scoring on the basis of ranking (element 7). It is however, important to realize that these elements, in theory, are expected to result in undesirable and unpredictable outcomes, although we have been unable to provide empirical evidence of a relation towards problems.

The overall results indicate that the elements are distributed differently across procurements. The six significant elements are more present in problematic procurement than non-problematic procurements. However, this does not necessarily imply a causal relationship between the identified elements and problems in public procurement. As discussed previously, this concerns the number of times an element has occurred in the stratified random sample (section 5.1 – Table 7). This implies that the statistical test can be significant, while these elements are commonly present in cases where no problems have arisen. Therefore, we are unable to provide empirical evidence on a causal relationship between these elements and problems, but we do signal an increased risk to problems.

With these overall findings we are able to answer the main question:

To what extent do award methods (in the award phase of Dutch public tender between May 2010 and April 2014) include the 10 selected elements, and to what extent do these elements result in ‘problems’?

The first part of the main question; to what extent do award methods include the 10 selected elements, is extensively discussed in section 5.1 and 5.2. We conclude that certain elements are often applied in the award phase. In addition, the simultaneous use of 3, 4 or 5 of these elements in a single procurement is not uncommon. In many cases, the use of these elements did not result in (measurable) problems. However, the data indicates, that the likelihood of problems increases when the elements are increasingly applied. Thus, award phases which contain a combination of the selected elements are more susceptible to problems than others.

The second part of the main question; to what extent do these elements result in problems, is discussed in section 5.3. We conclude that the six significant elements have an increased likelihood of problems. With the use or exclusion of these elements (negative impact) the chance of problems is twice to four times higher, varying in their susceptibility to problems from 21% to 60%. In addition, two outliers were indicated, element 3 (weights – 60%) and element 4 (score method – 32%). This indicates that there are few cases in the stratified random sample. In case these two elements are not indicated in RfQ's, the outcome problems is expected. These two elements are uncommon from a legal perspective and explicitly forbidden in procurement regulations. Furthermore, not indicating these elements is in violation with the principle of transparency and in violation the European directives and Dutch Public Procurement Act 2012.

6.2 Recommendations and implications for practitioners

In this section the findings are more broadly interpreted. We discuss findings with respect to their implication for practitioners.

As discussed previously, in public procurement law there has been little regulation on award methods.⁴³ However, the European directives and the Dutch Public Procurement Act base legislation on the five procurement principles (non-discrimination, equal treatment, mutual recognition, proportionality and transparency). This means that these principles have an essential role in the interpretation and application of procurement law. Consequently, these principles play a decisive role in determining whether an award method is regarded as legally permissible.

In this study, it is concluded that contracting authorities often meet the procurement principles in their award methods, while these award methods entail 'problematic' elements from a economic and purchasing technical perspective. For example, the procurement principles have been satisfied, while suppliers still cope with uncertainties, questions or complaints in preparing their offers. This insight indicates the negative theoretical and/or empirical consequences of certain elements. Practitioners can apply award methods in such a way that the identified elements are withhold, or applied in non-problematic ways.

In practice, there may be many formal methods, decision-criteria, decision makers and stakeholders with different perspectives, which all complicate decision-making. This results in lots of options and hence choices to be made. Overall, more transparency by means of a 'full, open and truthful exchange' of values and preferences increases public accountability, its level of interest, equal opportunities and promotes an environment of genuine market competition (Raiffa et al., 2007; Mateus et al. 2010). This leads in the final notion, that award phases, which include all parameters of the awarding mechanism and measure objectively, will be (most importantly) effective in determining the 'best' offer, more easily accepted and easier to sustain. This is beneficial for both contracting authorities, suppliers as well as for other stakeholders.

6.3 Limitations

The results of our research indicate that six elements more susceptible to problems, because they are empirically related to problematic procurements. However, our finding needs to be interpreted with caution, as there are some limitations of the to our study. Our analysis covered a particular group of public procurements in a specific time period. Furthermore, a specific method was used to compare the groups of procurements. These consideration resulted in limitation over sample, location and method.

Firstly, the sample is biased in procurement procedures, since contracting authorities are not obligated to publish RfQ's in case of negotiated procedures and single call for proposals. In addition, RfQ's are not always publicly available for restricted procedures. This means that our findings are to a lesser extent generalizable for these types of procurement procedures.

⁴³ Public procurement law: only general rules with regard to the award method are included in legislation (European directives 2004/18/EG and 2004/17/EG, Dutch Public Procurement Act 2012, Guide of proportionality, General Administrative Law Act (AWB). Furthermore, judgments of the Supreme Court of the Netherlands are also of influence.)

Secondly, this study compares Dutch problematic and non-problematic procurements. We were able to correct for function, task and institutional setting. However, questions remain on whether the results can be extrapolated to private procurement and other countries. That is, we cannot tell from this study what results would appear in other settings, in, for example, the United States. An expectation is that these findings may be even more pronounced in other countries, including different procurement regulations and guidelines and/or more leeway which can make the award phase more susceptible to problems. In general, these findings can be generalized to all EU countries, in which European directives are implemented. However, the directives can be implemented in several ways. For example, Portugal (EU Member State) prohibited relative scoring in their national law. Thus, the award cannot depend on features submitted in other tenders, given the substantial undesirable consequences that may occur by using this method (Mateus et al, 2010). This means that the findings in element 6 (relative award methods), do not relate to Portuguese award methods. Overall, the analysis focused on purchasing technical aspects, these can be generalized towards any setting.

Thirdly, this study has selected problematic procurements based on their substantive aspects. Consequently, there is a huge potential methodological problem when selecting on the dependent variable. It would have been preferable to select on the theoretical grounds expected to cause problems, and, subsequently measure which of these actually resulted to problems. However, in practice this is not feasible, since measurable problems (i.e. lawsuits and questions addressed by the commission of procurement experts) occur infrequently. This brings us to the last limitation with regard to method, not all problems can be identified. Hence, they cannot be captured with the measures used for the dependent variable 'problems'.

Finally, only 10 elements out of many which are subject to leeway in the award phase are analyzed. Other elements and multivariate analysis can provide additional information on elements and their susceptibility to problems. To conclude, there are some limitations to the generalizability and completeness of this study. Offering starting points for future research.

6.4 Further (empirical) research

The preliminary literature search, analysis and results indicate a number of possibilities for further research. In addition to the limitations of this study, three further research recommendations are assessed briefly below.

Firstly, more research is needed to shed light on the theoretical and empirical effects of the elements used in the award phase. It would be useful to identify which elements (besides the 10 discussed in this study), are susceptible to problems. This could be achieved by examining a number of problematic tenders in depth by means of qualitative research.

Secondly, in this study we adopted the following principle: if a supplier complains there are problems. However, there are many other ways to identify problems. For example, it is possible to make use of unstructured or open ended questionnaires to identify which elements in the award phase resulted in unintended or unforeseen effects according to stakeholders.

Thirdly, in practice, formal methods are often applied (see Appendix H – element 1: WFS, VFM and AOV). In many cases, these are applied correctly and practitioners who used these methods will gain accurate information of differences in performance. However, there are also situations in which

these methods are applied incorrectly, including problematic elements which result in undesired and unforeseen effects. As discussed previously (paragraph 6.2), award methods often comply with the European procurement principles, while in practice economical and purchasing technical aspects are the ones causing problems. Thus, when it comes to supplier selection, further research can bridge the gap between the complying with European principles and complying with purchasing technical perspectives. This can be achieved by developing a model or integral approach for establishing the award method. In current literature, no model could be identified which provided guidance for developing the award method with regard to technical aspects.

References

- Albano, G.L., Dini, F., Zampino, R., Suppliers' Behavior in Competitive Tendering: Evidence from the Italian Ministry of Economy and Finance's Acquisitions of IT Services, Proceedings of the 3rd IPPC Conference, Amsterdam (The Netherlands), August 28-30, 2008.
- Asker, J., Cantillon, E., 2008. Properties of scoring auctions. *Rand Journal of Economics* 39 (1), 69–85.
- Asker, J., Cantillon, E., 2009. Procurement when price and quality matter forthcoming in *RAND. Journal of Economics*.
- Babbie, E. (2010). *The Practice of Social Research*. Belmont: Wadsworth.
- Bergman, M.A., Lundberg, S., 2013. Tender evaluation and supplier selection methods in public procurement. *Journal of Purchasing and Supply Management*, 19 (2), pp. 73-83. doi: 10.1016/j.pursup.2013.02.003
- Boer L. de, Labro E., Morlacchi P., A review of methods supporting supplier selection, *European Journal of Purchasing and Supply Management*, Volume 7, Number 2, pp. 75-89(15), 2001.
- Boer, L. de, Linthorst, M.M., Schotanus, F., Telgen, J., An analysis of some mistakes, miracles and myths in supplier selection, Proceedings of the 15th IPSERA Conference, San Diego (United States), April 6-8, 2006.
- Boer L. de, van der Wegen L., Telgen J., Outranking methods in support of supplier selection, *European Journal of Purchasing and Supply Management*, Volume 4, Number 2, pp. 109-118(10), 1998.
- Chen, T.H., 2008. An economic approach to public procurement. *Journal of Public Procurement* 8, 407–430.
- Dini, F., Pacini, R., Valletti, T., 2006. Scoring rules. Nicola Dimitri, Gustavo Piga, Giancarlo Spagnolo (Eds), *Handbook of Procurement*, Cambridge University Press.
- European Union, 2004. Directive 2004/18/EC of the European parliament and of the council of 31 march 2004 on the coordination of procedures for the award of public works contracts, public supply contracts and public service contracts. *Official Journal*, L 134, 30.04.2004, pp. 114–240.
- Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research*, 202(1), 16–24. doi:10.1016/j.ejor.2009.05.009
- Jansen A., Kolkman, S., Kuijpers, P., Pries F., Reeuwijk, T. van, Witteveen, B., & Zijlstra, J.O. (2007). *Gunnen op waarde: hoe doe je dat?* Praktische handreiking voor bouwopdrachten. (1st ed. publicatie 253). The Netherlands: CROW.

- Keeney, R.L., 1992. Value-Focused Thinking: a Path to Creative Decision-making. Harvard University Press, Cambridge.
- Mateus, R., Ferreira, J.A., Carreira, J., 2010. Full disclosure of tender evaluation models: background and application in Portuguese public procurement. *Journal of Purchasing & Supply Management* 16, 206–215
- Public Procurement Research Center (2013). Hoe wordt uitzendwerk gegund? Onderzoek naar gunningsmethodieken voor uitzendwerk: een economische en juridische analyse. Utrecht/Twente: W, Lohamn., E, Manuza and J. Telgen.
- Raiffa, H., Richardson, J., Metcalfe, D., 2007. *Negotiation Analysis: The Science and Art of Collaborative Decision Making*, Cambridge, MA.. Harvard University Press.
- Telgen, J., Schotanus, F., 2010. The effects of full transparency in supplier selection on subjectivity and bid quality. In: *Proceedings of the Fourth International Public Procurement Conference*, Seoul.
- Veaux, R. De, Velleman, P. & Bock, D. (2012). *Stats: Data and Models*. Third edition. Pearsons Education.

Appendix A: The setting - TED Europe descriptive information

The population is divided in three sub-groups; contracting authorities, procurement procedures and contract subject. The distribution of these three subgroups is presented during the research time frame (May 2010 – April 2014). All information within the following three tables is based on information in the database ted.europe.eu.

1. Contracting authorities (derived from ted.europe.eu)

Organization type – All /RfQ/Award	All	Percent	RfQ	Percent	Award	Percent
Ministerie of elke andere nationale of federale instantie	5.527	14,6%	2.076	13,3%	2.281	15,3%
Regionale/lokale overheid	16.813	44,5%	6.876	44,0%	6.796	45,7%
Nutssectoren	2.712	7,2%	1.108	7,1%	911	6,1%
Europese instelling/Europees agentschap of internationale organisatie	531	1,4%	217	1,4%	214	1,4%
Andere	1.727	4,6%	752	4,8%	593	4,0%
Niet van toepassing	41	0,1%	0	0,0%	0	0,0%
Publiekrechtelijke instelling	9.530	25,2%	4.200	26,8%	3.773	25,4%
Nationaal of federaal agentschap/bureau	554	1,5%	220	1,4%	206	1,4%
Regionaal of lokaal agentschap/bureau	131	0,3%	50	0,3%	61	0,4%
Niet gespecificeerd	249	0,7%	145	0,9%	26	0,2%
Total	37.815	100%	15.644	100%	15.499	100%

Appendix A – Table I – TED descriptive information **organization type**: All announcements, RfQ and Award of procurements

2. Procurement procedures (derived from ted.europe.eu)

Procedure type – All/RfQ/Award	All	Percent	RfQ	Percent	Award	Percent
Openbare procedure	25.825	69,4%	11.858	75,8%	10.160	68,4%
Niet-openbare procedure	6.508	17,4%	3.096	19,8%	2.524	17,0%
Versnelde, niet-openbare procedure	27	0,1%	10	0,1%	12	0,1%
Procedure van gunning via onderhandelingen	1.200	3,2%	563	3,6%	447	3,0%
Versnelde procedure van gunning via onderhandelingen	14	0,0%	5	0,0%	8	0,1%
Concurrentiegericht dialog	201	0,5%	112	0,7%	55	0,4%
Onderhandelingsprocedure zonder een oproep tot mededinging	386	1,0%	0	0,0%	281	1,9%
Gunning van een opdracht zonder voorafgaande bekendmaking van een aankondiging van de opdracht	1.493	4,0%	0	0,0%	1.363	9,2%
Andere	0	0,0%	0	0,0%	0	0,0%
Niet van toepassing	1.584	4,2%	0	0,0%	5	0,0%
Niet gespecificeerd	10	0,0%	0	0,0%	6	0,0%
Onderhandse gunning	0	0,0%	0	0,0%	0	0,0%
Total	37.340	100%	15.644	100%	14.861	100%

Appendix A – Table II: TED descriptive information **procurement procedures**: All documents, RfQ and award of procurements.

3. Contract subject (derived from ted.europe.eu)

Procedure type – All/RfQ/Award	All	Percent	RfQ	Percent	Award	Percent	Average Percentage
Works	5.581	14,8%	2.334	14,9%	1.721	11,6%	13,8%
Goods	11.715	31,0%	5.198	33,2%	4.353	29,3%	31,2%
Services	20.478	54,2%	8.112	51,9%	8.787	59,1%	55,0%
Niet van toepassing	39	0,1%	0	0,0%	0	0,0%	0,0%
Niet gespecificeerd	0	0.0%	0	0,0%	0	0,0%	0,0%
Total	37.813	100%	15.644	100%	14.861	100%	100%

Appendix A – Table III: TED descriptive information **contract subject**: All documents, RfQ and award of procurements.

Appendix B: Population – All announcements ABK

Population 2010													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	-	-	-	-	629	1.317	1.028	658	947	803	867	676	6.895
Goods	-	-	-	-	228	312	388	302	378	332	356	296	2.592
Services	-	-	-	-	473	679	694	576	764	821	780	649	5.436
Total	-	-	-	-	1.330	2.308	2.110	1.536	2.089	1.956	1.973	1.621	14.923

Population 2011													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	699	722	894	942	933	1.020	852	583	790	679	876	606	9.506
Goods	266	321	358	414	402	339	366	318	431	394	357	301	4.267
Services	668	721	883	743	607	529	567	606	753	693	659	585	8.014
Total	1.633	1.764	2.135	2.099	1.942	1.888	1.795	1.523	1.858	1.766	1.892	1.492	21.787

Population 2012													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	541	664	802	725	677	774	546	466	443	486	518	432	7.074
Goods	271	297	340	277	273	304	277	230	270	321	241	240	3.341
Services	634	726	641	513	408	446	399	423	548	646	520	417	6.321
Total	1.446	1.687	1.783	1.515	1.358	1.524	1.222	1.119	1.261	1.453	1.279	1.089	16.736

Population 2013													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	448	457	675	440	324	300	265	262	294	299	197	198	4.159
Goods	183	206	355	156	134	129	183	228	234	344	172	189	2.513
Services	390	425	602	301	257	195	285	342	373	508	324	326	4.328
Total	1.021	1.088	1.632	897	715	624	733	832	901	1.151	693	713	11.000

Population 2014													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	152	166	219	332	-	-	-	-	-	-	-	-	869
Goods	120	142	156	285	-	-	-	-	-	-	-	-	703
Services	250	240	251	473	-	-	-	-	-	-	-	-	1214
Total	522	548	626	1.090	-	-	-	-	-	-	-	-	2786

Population	2010 (8 months)	2011	2012	2013	2014 (4 months)	Total
Works	6.895	9.506	7.074	4.159	869	28.503 (42,2%)
Goods	2.592	4.267	3.341	2.513	703	13.416 (20,0%)
Services	5.436	8.014	6.321	4.328	1.214	25.313 (37,7%)
Total	14.923 (22,2%)	21.787 (32,4%)	16.736 (24,9%)	11.000 (16,4%)	2.786 (4,1%)	67.232 (100%)

Appendix C: Sample Group 1 – non-problems

Sample 2010													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	-	-	-	-	4	8	6	4	6	5	5	4	42
Goods	-	-	-	-	1	2	2	2	2	2	2	2	15
Services	-	-	-	-	3	4	4	3	5	5	5	4	33
Total	-	-	-	-	8	14	12	9	13	12	12	10	90

Sample 2011													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	4	4	5	6	6	6	5	3	4	4	5	4	56
Goods	2	2	2	2	2	2	2	2	3	2	2	2	25
Services	4	4	5	4	4	3	3	4	4	4	4	3	46
Total	10	10	12	12	12	11	10	9	11	10	11	9	127

Sample 2012													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	3	4	5	4	4	5	3	3	3	3	3	3	43
Goods	2	2	2	2	2	2	2	1	2	2	1	1	21
Services	4	4	4	3	2	3	2	3	3	4	3	2	37
Total	9	10	11	9	8	10	7	7	8	9	7	6	101

Sample 2013													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Works	3	3	4	3	2	2	2	2	2	2	1	1	27
Goods	1	1	2	1	1	1	1	1	1	2	1	1	14
Services	2	3	4	2	2	1	2	2	2	3	2	2	27
Total	6	7	10	6	5	4	5	5	5	7	4	4	68

Sample 2014													
	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Work	1	1	1	2	-	-	-	-	-	-	-	-	5
Goods	1	1	1	2	-	-	-	-	-	-	-	-	5
Service	1	1	1	3	-	-	-	-	-	-	-	-	6
Total	3	3	3	7	-	-	-	-	-	-	-	-	16

Sample	2010 (8 months)	2011	2012	2013	2014 (4 months)	Total
Works	42	56	43	27	5	173 (43,0%)
Goods	15	25	21	14	5	80 (19,9%)
Services	33	46	37	27	6	149 (37,1%)
Total	90 (22,4%)	127 (31,6%)	101 (25,1%)	68 (16,9%)	16 (4,0%)	402 (100%)

Appendix D: Sample Group 2 – problems

Procurements which resulted in lawsuits						
	Rechtbank	Uitspraak	Publicatie	Zaaknummer	Type aanbesteding	Sample
1	Zeeland-West-Brabant	9-10-2013	29-10-2013	127261 - KG ZA 11-377	Europees openbaar	JA
2	Dordrecht	20-11-2012	20-11-2012	100320 - KG ZA 12-184	Europees openbaar	NEE
3	Leeuwarden	6-7-2011	19-7-2011	112632 - KG ZA 11-146	Europees openbaar	JA
4	s-Gravenhage	13-5-2011	17-8-2011	391406 - KG ZA 11-383	Europees openbaar	NEE
5	s-Gravenhage	27-3-2013	27-3-2013	407781 - KG ZA 11-1374	Europees openbaar	NEE
6	Arnhem	13-12-2012	15-1-2013	235586 - KG ZA 12-574	Europees openbaar	JA
7	Amsterdam	1-4-2011	3-5-2011	483304 - KG ZA 11-285	Europees openbaar	JA
8	Den Haag	25-2-2014	21-5-2014	455756 - KG ZA 13-1382	Europees openbaar	JA
9	Zutphen	1-3-2012	7-5-2012	127261 - KG ZA 11-377	Europees openbaar	NEE
10	's-Gravenhage	12-9-2012	28-9-2012	424025 - KG ZA 12-778	Europees openbaar	JA
11	's-Gravenhage	21-3-2011	2-12-2011	387675 - KG ZA 11-188	Europees openbaar	JA
12	Arnhem	20-4-2011	10-5-2011	213256 - KG ZA 11-130	Europees openbaar	JA
13	Rotterdam	7-6-2011	22-7-2011	375160 - KG ZA 11-252	Europees openbaar	NEE
14	's-Gravenhage	19-8-2010	30-9-2010	371165 - KG ZA 10-888	Europees openbaar	NEE
15	Almelo	4-6-2012	5-6-2012	128137 - KG ZA 12-78	Nationaal niet-openbaar	NEE
16	Rotterdam	2-8-2012	7-8-2012	403627 - KG ZA 12-469	Europees openbaar	NEE
17	Amsterdam	27-1-2012	7-2-2012	506266 - KG ZA 11-1982	Europees openbaar	NEE
18	Utrecht	24-1-2014	9-5-2014	309740 - KG ZA 11-636	Europees openbaar	JA
19	's-Gravenhage	5-3-2012	7-5-2012	412264 - KG ZA 12-106	Europees openbaar	JA
20	Leeuwarden	15-12-2010	17-12-2010	107368 - KG ZA 10-284	Europees openbaar	JA
21	Arnhem	11-6-2010	6-7-2010	199851 - KG ZA 10-289	Europees openbaar	NEE
22	Noord-Nederland	11-1-2013	11-1-2013	137574 - KG ZA 12-345	Europees openbaar	NEE
23	s-Gravenhage	19-1-2011	7-2-2011	381746 - KG ZA 10-1510	Europees openbaar	JA
24	s-Gravenhage	19-1-2011	7-2-2011	383075 - KG ZA 10-1609	Europees openbaar	JA
25	's-Gravenhage	6-9-2011	2-12-2011	398830 - KG ZA 11-849	Nationaal niet-openbaar	NEE
26	Arnhem	19-11-2010	29-11-2010	206107 - KG ZA 10-642	Nationaal openbaar	NEE
27	's-Gravenhage	29-3-2011	7-4-2011	387696 - KG ZA 11-189	Europees openbaar	JA
28	's-Gravenhage	27-10-2010	29-10-2010	376032 - KG ZA 10-1150	Europees niet-openbaar	JA
29	Zwolle-Lelystad	16-7-2010	26-7-2010	171534 - KG ZA 10-244	Nationaal openbaar	JA
30	's-Gravenhage	20-12-2011	23-12-2011	406218 - KG ZA 11-1289	Nationaal openbaar	JA
31	Gelderland	24-1-2014	28-1-2014	254822 - KG ZA 13-648	Europees niet-openbaar	JA
32	Limburg	22-8-2013	5-9-2013	182343 - KG ZA 13-298	Europees openbaar	JA
33	Noord-Nederland	29-5-2013	30-5-2013	125826 - KG ZA 13-70	Europees niet-openbaar	NEE
34	Arnhem	6-12-2010	28-12-2010	208039 - KG ZA 10-730	Meervoudig onderhands	JA
35	Amsterdam	4-2-2011	3-5-2011	481691 - KG ZA 11-156	Europees openbaar	NEE
36	Arnhem-Leeuwarden	17-9-2013	30-9-2013	HB - 200.128.145	Europees niet-openbaar	NEE
37	Amsterdam	5-12-2011	10-1-2012	501650 - KG ZA 11-1612	Europees openbaar	JA
38	Den Haag	24-9-2013	25-11-2013	446705 - KG ZA 13-828	Europees openbaar	NEE

39	Amsterdam	9-7-2013	21-10-2013	HB - 200.123.074	Europees openbaar	NEE
40	's-Hertogenbosch	12-2-2013	20-2-2013	HB - 200.099.589	Nationaal openbaar	NEE
41	Utrecht	16-11-2012	19-11-2012	331347 - KG ZA 12-743	Europees openbaar	NEE
42	's-Hertogenbosch	1-10-2010	4-10-2010	217422 - KG ZA 10-596	Europees openbaar	JA
43	Zwolle-Lelystad	12-11-2012	14-11-2012	203139 - KG ZA 12-191	Meervoudig onderhands	JA
44	Noord-Nederland	12-2-2014	31-02-2014	130561 - KG ZA 13-317	Meervoudig onderhands	NEE
45	Noord-Nederland	29-11-2013	1-12-2013	143360 - KG ZA 13-267	Europees openbaar	NEE
46	Amsterdam	22-1-2013	6-2-2013	530168 - KG ZA 12-1567	Europees openbaar	NEE
47	's-Gravenhage	9-5-2012	25-5-2012	416277 - KG ZA 12-344	Europees niet-openbaar	NEE
48	Midden-Nederland	1-2-2013	4-2-2013	332043 - KG ZA 12-774	Europees openbaar	NEE
49	s-Hertogenbosch	16-5-2011	19-5-2011	228782 - KG ZA 11-229	Europees openbaar	JA
50	s-Hertogenbosch	15-4-2011	21-4-2011	227457 - KG ZA 11-171	Europees openbaar	JA
51	Noord-Nederland	25-10-2013	20-11-2013	140213 - KG ZA 13-85	Europees openbaar	JA
52	Den Haag	10-10-2013	4-12-2013	447983 - KG ZA 13-892	Europees openbaar	JA
53	Amsterdam	24-3-2011	3-5-2011	480801 - KG ZA 11-86	Europees openbaar	JA
54	Zeeland-west-Brabant	9-10-2013	29-10-2013	269605 - KG ZA 13-533	Europees openbaar	JA
55	's-Gravenhage	25-6-2012	29-6-2012	418456 - KG ZA 12-456	Europees niet-openbaar	NEE
56	Den Haag	20-11-2013	4-12-2013	451680 - KG ZA 13-1112	Europees niet-openbaar	NEE
57	Den Haag	21-6-2013	2-8-2013	442443 - KG ZA 13-511	Europese aanbesteding	NEE
58	Arnhem	9-10-2012	10-10-2012	HB - 200.096.019	Europees openbaar	JA
59	Utrecht	16-9-2011	19-9-2011	309740 - KG ZA 11-636	Europees openbaar	JA
60	Noord-Nederland	23-1-2014	28-1-2014	102330 - KG ZA 13-256	Europees openbaar	JA
61	Rotterdam	20-12-2010	18-1-2011	363081 - KG ZA 10-941	Europees openbaar	JA
62	Den Haag	29-10-2013	4-12-2013	449395 - KG ZA 13-968	Europees openbaar	NEE
63	's-Hertogenbosch	29-11-2010	29-11-2010	221060 - KG ZA 10-758	Europees openbaar	JA
64	Den Haag	2-5-2013	17-6-2013	439520 - KG ZA 13-306	Europees openbaar	NEE
65	Arnhem	29-6-2011	12-7-2011	216697 - KG ZA 11-281	Europees openbaar	NEE
66	Amsterdam	21-12-2010	26-1-2011	475727 - KG ZA 10-2130	Europees niet-openbaar	JA
67	Leeuwarden	24-11-2010	17-12-2010	107702 - KG ZA 10-305	Meervoudig onderhands	NEE
68	Almelo	4-6-2012	4-6-2012	128158 - KG ZA 12-80	Europees niet-openbaar	NEE
69	Middelburg	9-7-2010	20-5-2011	73410 - KG ZA 10-89	Europees niet-openbaar	NEE
70	Den Haag	25-6-2013	18-7-2013	HB - 200.126.228-01	Europees openbaar	JA
71	Arnhem	22-11-2011	30-12-2011	221840 - KG ZA 11-548	Europees openbaar	JA
72	Amsterdam	9-12-2010	24-12-2010	475131 - KG ZA 10-2089	Europees openbaar	NEE
73	Den Haag	23-12-2013	27-1-2014	454713 - KG ZA 13-1305	Europees openbaar	NEE
74	Zutphen	27-4-2012	23-7-2012	129075 - KG ZA 12-92	Europees niet-openbaar	NEE
75	Den Haag	2-4-2013	10-7-2013	436273 - KG ZA 13-106	Europees niet-openbaar	NEE
76	's-Gravenhage	19-6-2012	29-6-2012	418048 - KG ZA 12-440	Meervoudig onderhands	NEE
77	's-Gravenhage	13-2-2012	5-3-2012	409532 - KG ZA 11-1493	Europees openbaar	NEE
78	Arnhem	19-1-2011	19-1-2011	208280 - KG ZA 10-737	Meervoudig onderhands	NEE
79	Den Haag	25-2-2014	21-5-2014	455756 - KG ZA 13-1382	Europees openbaar.	NEE
80	Den Haag	25-2-2014	21-5-2014	455836 - KG ZA 13-1388	Europees openbaar	NEE

81	Gelderland	20-1-2014	1-5-2014	255773 - KG ZA 13-681	Europees niet-openbaar	JA
82	Noord-Nederland	9-4-2014	22-4-2014	133115 - KG ZA 14-87	Europees openbaar	JA
83	Noord-Nederland	19-3-2014	19-3-2014	132103 - KG ZA 14-25	Nationaal niet-openbaar	NEE
84	Oost-Brabant	5-3-2014	11-3-2014	273283 - KG ZA 14-21	Meervoudig onderhands	NEE
85	Den Haag	23-12-2013	27-1-2014	454713 - KG ZA 13-1305	Europees openbaar	JA
Procurement which resulted in questions that were addressed by the Commission of Procurement Experts.						
	Nr. advies	Datum	Type aanbesteding			Sample
86	nr. 22	13-8-2013	Europees openbaar			JA
87	nr. 23	8-8-2013	Nationaal openbaar			JA
88	nr. 30	25-10-2013	Meervoudig onderhands			NEE
89	nr. 48	27-1-2014	Europees niet-openbaar			NEE
90	nr. 59	24-12-2013	Meervoudig onderhands			NEE
91	nr. 70	24-4-2014	Meervoudig onderhands			NEE
92	nr. 79	12-5-2014	Europees openbaar			JA

Appendix E: Sample Group 2 ‘problems’ – bias check removal of cases with insufficient information

Lawsuit	Population	Percent	Sample	Percent
Nationaal openbaar	4	4,7%	2	5,0%
Nationaal niet-openbaar	3	3,5%	0	0,0%
Meervoudig onderhands	7	8,2%	2	5,0%
Europees openbaar	57	67,1%	31	77,5%
Europees niet-openbaar	13	15,3%	5	12,5%
Europese aanbesteding	1	1,2%	0	0,0%
Total	85	100%	40	100%
Commission Procurement Experts				
Nationaal openbaar	1	14,3%	1	33,3%
Nationaal niet-openbaar	0	0,0%	0	0,0%
Meervoudig onderhands	3	42,9%	0	0,0%
Europees openbaar	2	28,6%	2	66,7%
Europees niet-openbaar	1	14,3%	0	0,0%
Europese aanbesteding	0	0,0%	0	0,0%
Total	7	100%	3	100%

Both measures (lawsuit + Commission)	Population	Percent	Sample	Percent
Nationaal openbaar	5	5,4%	3	7,0%
Nationaal niet-openbaar	3	3,3%	0	0,0%
Meervoudig onderhands	10	10,9%	2	4,7%
Europees openbaar	59	64,1%	33	76,7%
Europees niet-openbaar	14	15,2%	5	11,6%
Europese aanbesteding	1	1,1%	0	0,0%
Total	92	100%	43	100%

Appendix F: Research format

	Analysis	Coding
Part 1: Control Characteristics		
- Number Procurement	Number	Numeric
- Name Contracting Authority	Name	String
- Name procurement	Name/description	Date
- Publication date	Date	Numeric
- Type of contracting authority	State / Province / Municipality / Water Board / Public institution / Cooperation of the previous governments of public institutions / Other: such as Education and Port companies	Numeric
- Type of procurement contract	Works / Goods / Services	Numeric
- Type of procurement procedure	National open procedure / European open procedure / National restricted procedure / European restricted procedure / Negotiated procedure / Negotiated procedure without publication	Numeric
- Province of contracting authority	Province	Numeric
Part 2: 10 Elements		
- Award criterion	EMAT / Lowest price	Numeric
- Award criteria 1 st line	Name / Type	String
- Number of award criteria 1 st line	Number	Numeric
- Number of award criteria total	Number	Numeric
- Award criteria	Natural / proxy / constructed	Numeric
- Type of score Method (Award method)	Fixed-price / WFS / VFM / GOW	Numeric
- Relative scores (not absolute)	Absolute / relative	Numeric
- Form of the score graph	Hollow graph / Convex graph / Linear graph / Score method not published	Numeric
- Criterion price contains financial sub-criteria	Yes / No	Numeric
- Weights not published	Yes / No	Numeric
- Score method not published	Yes / No	Numeric
- Flat award criteria	Yes / No	Numeric
- Extra lower boundaries	Yes / No	Numeric
- Awarding points on the basis of ranking	Yes / No	Numeric
- Weight award criteria (EMAT)	Price-Quality ratio	Numeric

Part 3: Problems		
- Lawsuit	Yes / No	Numeric
- Lawsuit (winner and losers)	Suppliers lose = lawsuit is rejected / Suppliers win = claim is assigned)	Numeric
- Commission of Procurement Experts	Yes / No	Numeric
- Commission of Procurement Experts (founded and unfounded)	Suppliers complaints are founded / Suppliers complaints are unfounded	Numeric

Appendix G: Population and Sample (bias check)

Group 1: Non-problems

1. Organization type – Contracting authority

Contracting authority	Award criterion		Total
	EMAT	Lowest price	
State	13 7,1%	10 4,6%	23 5,7%
Province	10 5,4%	12 5,5%	22 5,5%
Municipality	79 42,9%	117 53,7%	196 48,8%
Water board	19 10,3%	23 10,6%	42 10,4%
Public entities	29 15,8%	26 11,9%	55 13,7%
Cooperation of previous governments or public institutions	15 8,2%	11 5,0%	26 6,5%
Other	19 10,3%	19 8,7%	38 9,5%
Total	184 100,0%	218 100,0%	402 100,0%

Appendix G – Table I: Sample non-problems organization types

2. Procurement procedure

Procedure type	Award criterion		Total
	EMAT	Lowest price	
National open	45 24,5%	121 55,5%	166 41,3%
European open	134 72,8%	85 39,0%	219 54,5%
National restricted	0 0,0%	2 0,9%	2 0,5%
European restricted	4 2,2%	10 4,6%	14 3,5%
Competitive dialogue	1 0,5%	0 0,0%	1 0,2%
Total	184 100,0%	218 100,0%	402 100,0%

Appendix G – Table II: Sample non-problems procurement procedures

3. Contract subject

Contract subject	Award criterion		Total
	EMAT	Lowest price	
Work	32 17,4%	141 64,7%	173 43,0%
Delivery	55 29,9%	25 11,5%	80 19,9%
Service	97 52,7%	52 23,9%	149 37,1%
Total	184 100,0%	218 100,0%	402 100,0%

Appendix G – Table III: Sample non problems contract subjects

Group 2: Problems (lawsuits and questions addressed by the CoPE)

1. Organization type – Contracting authority

Contracting authority	Award criterion		Total
	EMAT	Lowest price	
State	3 7,3%	1 50,0%	4 9,3%
Province	5 12,2%	0 0,0%	5 11,6%
Municipality	15 36,6%	1 50,0%	16 37,2%
Water board	2 4,9%	0 0,0%	2 4,7%
Public institution	2 4,9%	0 0,0%	2 4,7%
Cooperation of previous governments or public institutions	8 19,5%	0 0,0%	8 18,6%
Other	6 14,6%	0 0,0%	6 14,0%
Total	41 100,0%	2 100,0%	43 100,0%

Appendix G – Table IV: Sample problems organization types

2. Procurement procedure

Procedure type	Award criterion		Total
	EMAT	Lowest price	
National open	2 (4,9%)	1 (50,0%)	3 (7,0%)
European open	32 78,0%	1 50,0%	33 76,7%
European restricted	5 12,2%	0 0,0%	5 11,6%
Negotiated procedure	2 4,9%	0 0,0%	2 4,7%
Total	41 100,0%	2 100,0%	43 100,0%

Appendix G – Table V: Sample problems procurement procedures

3. Contract subject

Contract subject	Award criterion		Total
	EMAT	Lowest price	
Works	8 19,5%	2 100,0%	10 23,3%
Goods	9 22,0%	0 0,0%	9 20,9%
Services	24 58,5%	0 0,0%	24 55,8%
Total	41 100,0%	2 100,0%	43 100,0%

Appendix G – Table VI: Sample problems contract subject

Bias check

Bias check (1. Organization) - European procedures and contracting authority type

Contracting authority	Lawsuit		Total
	No	Yes	
State	23 9,9%	3 7,9%	26 9,6%
Province	12 5,2%	4 10,5%	16 5,9%
Municipality	76 32,6%	13 34,2%	89 32,8%
Water board	24 10,3%	2 5,3%	26 9,6%
Public entities	44 18,9%	2 5,3%	46 17,0%
Cooperation of previous governments or public institutions	25 10,7%	8 21,1%	33 12,2%
Other	29 12,4%	6 15,8%	35 12,9%
Total	233 100,0%	38 100,0%	271 100,0%

Appendix G – Table VII: Sample non-problems (selected European procurement) distribution by contracting authority

Bias check (2. Procurement procedure) - European procedures

Procedure type	Lawsuit		Total
	No	Yes	
European open	219 94,0%	33 86,8%	252 93,0%
European restricted	14 6,0%	5 13,2%	19 7,0%
Total	233 100,0%	38 100,0%	271 100,0%

Appendix G – Table VIII: Sample non-problems (selected European procurement) distribution by contracting authority

Bias check (3. Contract subjects) - European procedures

Contract subject	Lawsuit		Total
	No	Yes	
Work	37 15,9%	6 15,8%	43 15,9%
Delivery	76 32,6%	9 23,7%	85 31,4%
Service	120 51,5%	23 60,5%	143 52,8%
Total	233 100,0%	38 100,0%	271 100,0%

Appendix G – Table IX: Sample non-problems (selected European procurement) distribution by contract subject

Appendix H: Data analysis – 10 hypothesis

Element 1: The Weighted Factor Score (a form of price-to-quality scoring)

Analysis 1: (WFS + VFM) and (Fixed-price + AoV)

Type award method: recode (WFS+VFM) and (Fixed-price+AoV)

Problems	Type award method		Total
	WFS / VFM	Fixed-price / AoV	
NO	141 79,2%	37 90,2%	178 81,3%
YES	37 20,8%	4 9,8%	41 18,7%
Total	178 100,0%	41 100,0%	219 100,0%

Chi-Square Tests – analysis 1

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,665 ^a	1	,103
N of Valid Cases	219		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,68.

Analysis 2: (WFS) and (VFM + Fixed-price + GOW)

Type award method: recode (WFS) and (VFM+Fixed-price+GOW)

Problems	Type award method		Total
	WFS	VFM / GOW / Fixed-price	
NO	137 78,7%	41 91,1%	178 81,3%
YES	37 21,3%	4 8,9%	41 18,7%
Total	174 100,0%	45 100,0%	219 100,0%

Chi-Square Tests – analysis 2

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,599 ^a	1	,058
N of Valid Cases	219		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 8,42.

Analysis 3, 4 and 5: (WFS) and (VFM + Fixed-price + GOW) – Contract subject

Overview – award method WFS versus (VFM + Fixed-price and GOW)

Problems	Works			Goods			Services		
	WFS	Other	Total	WFS	Other	Total	WFS	Other	Total
NO	14 (77,8%)	18 (81,8%)	32 (80%)	43 (82,7%)	9 (100%)	52 (85,2%)	80 (76,9%)	14 (100%)	94 (79,7%)
YES	4 (22,2%)	4 (18,2%)	8 (20%)	9 (17,3%)	0 (0,0%)	9 (14,8%)	24 (23,1%)	0 (0,0%)	24 (20,3%)
Total	18 (100%)	22 (100%)	40 (100%)	52 (100%)	9 (100%)	61 (100%)	104 (100%)	14 (100%)	118 (100%)

Chi-square test (Works, Goods and Services)

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	,101 ^A	1	,751
Pearson Chi-Square (Goods)	1,827 ^B	1	,176
Pearson Chi-Square (Services)	4,056 ^C	1	,044
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	61		
N of Valid Cases (Services)	118		

- 2 cells (50,0%) have expected count less than 5. The minimum expected count is 3,60
- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,33.
- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 2,85.

Element 2: Qualitative – constructed performance levels

Analysis 6: Frequencies performance levels

Sum performance levels

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ,00	4	1,8	1,8	1,8
1,00	16	7,1	7,1	8,9
2,00	75	33,3	33,3	42,2
3,00	7	3,1	3,1	45,3
4,00	52	23,1	23,1	68,4
5,00	10	4,4	4,4	72,9
6,00	53	23,6	23,6	96,4
7,00	1	,4	,4	96,9
8,00	7	3,1	3,1	100,0
Total	225	100,0	100,0	

Sum performance levels recoded in to three categories

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Low	95	42,2	42,2	42,2
Medium	69	30,7	30,7	72,9

High	61	27,1	27,1	100,0
Total	225	100,0	100,0	

Analysis 7: performance levels

Performance levels and problems

Problems	Natural, proxy and constructed			Total
	Low	Medium	High	
NO	78 82,1%	56 81,2%	50 82,0%	184 81,8%
YES	17 17,9%	13 18,8%	11 18,0%	41 18,2%
Total	95 100,0%	69 100,0%	61 100,0%	225 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,026 ^a	2	,987
N of Valid Cases	225		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,12.

Element 3: Weighing not published

Analysis 8: publication weights (overall)

Publication weights (YES / NO)

Problems	Publication weights		Total
	NO	YES	
NO	2 40,0%	180 82,6%	182 81,6%
YES	3 60,0%	38 17,4%	41 18,4%
Total	5 100,0%	218 100,0%	223 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,903 ^a	1	,015
N of Valid Cases	223		

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,92.

Analysis 9, 10 and 11: Publication of weights – Contract subject

Overview – Publication of weights (Works / Goods and Services)

Problems	Works			Goods			Services		
	NO	YES	Total	NO	YES	Total	NO	YES	Total
NO	0 (0,0%)	32 (82,1%)	32 (80%)	0 (0,0%)	55 (87,3%)	55 (85,9%)	2 (66,7%)	93 (80,2%)	95 (79,8%)
YES	1 (100%)	7 (17,9%)	8 (20%)	1 (100%)	8 (12,7%)	9 (14,1%)	1 (33,3%)	23 (19,8%)	24 (20,2%)
Total	1 (100%)	39 (100%)	40 (100%)	1 (100%)	63 (100%)	64 (100%)	3 (100%)	116 (100%)	119 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	4,103 ^A	1	,042
Pearson Chi-Square (Goods)	6,208 ^B	1	,013
Pearson Chi-Square (Services)	,331 ^C	1	,565
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	64		
N of Valid Cases (Services)	118		

- 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,20.
- 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,14.
- 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,61.

Element 4: *No publication of the score method*

Analysis 12: publication score method (overall)

Publication score method (YES / NO)

Problems	Publication method		Total
	NO	YES	
NO	17 68,0%	164 84,1%	181 82,3%
YES	8 32,0%	31 15,9%	39 17,7%
Total	25 100,0%	195 100,0%	220 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,940 ^a	1	,047
N of Valid Cases	220		

- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 4,43.

Analysis 13, 14 and 15: Publication of the score methods – Contract subject

Overview – Publication of weights (Works / Goods and Services)

Problems	Works			Goods			Services		
	NO	YES	Total	NO	YES	Total	NO	YES	Total
NO	3 (100%)	29 (78,4%)	32 (80%)	9 (75,0%)	46 (90,2%)	55 (87,3%)	5 (50,0%)	89 (83,2%)	94 (80,3%)
YES	0 (0,0%)	8 (21,6%)	8 (20%)	3 (25,0%)	5 (9,8%)	8 (12,7%)	5 (50,0%)	18 (16,8%)	23 (19,7%)
Total	3 (100%)	37 (100%)	40 (100%)	12 (100%)	51 (100%)	63 (100%)	10 (100%)	107 (100%)	117 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	,881 ^A	1	,368
Pearson Chi-Square (Goods)	2,023 ^B	1	,347
Pearson Chi-Square (Services)	6,374 ^C	1	,012
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	63		
N of Valid Cases (Services)	117		

- 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,60.
- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,52.
- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,97.

Element 5: Non-linear score graphs – hollow score graphs for award criterion price.

Analysis 16: linear, convex, concave and not published

Form score graph

Problems	Form Score Graph				Total
	Linear	Convex	Hollow	Not published	
NO	84 84,0%	2 100,0%	53 85,5%	16 66,7%	155 82,4%
YES	16 16,0%	0 0,0%	9 14,5%	8 33,3%	33 17,6%
Total	100 100,0%	2 100,0%	62 100,0%	24 100,0%	188 100,0%

Chi-Square Tests – analysis 1

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,117 ^a	3	,163
N of Valid Cases	188		

- 3 cells (37,5%) have expected count less than 5. The minimum expected count is ,35.

Analysis 17: recode (linear) and (convex + concave + not published)

Form score graph: recode (linear) and (convex+concave+not published)

Problems	Form score graph		Total
	Linear	Concave, convex and not published	
NO	84 84,0%	71 80,7%	155 82,4%
YES	16 16,0%	17 19,3%	33 17,6%
Total	100 100,0%	88 100,0%	188 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,356 ^a	1	,551
N of Valid Cases	188		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 15,45.

Analysis 18, 19 and 20: recode (linear) and (convex + concave + not published) – Contract subject

Overview – Form score graph – (Works / Goods and Services)

Problems	Works			Goods			Services		
	Linear	Other	Total	Linear	Other	Total	Linear	Other	Total
NO	18 (85,7%)	4 (80,0%)	22 (84,6%)	34 (82,9%)	17 (100,0%)	51 (87,9%)	50 (78,1%)	32 (80,0%)	82 (78,8%)
YES	3 (14,3%)	1 (20,0%)	4 (15,4%)	7 (17,1%)	0 (0,0%)	7 (12,1%)	14 (21,9%)	8 (20,0%)	22 (21,2%)
Total	21 (100%)	5 (100%)	26 (100%)	41 (100%)	17 (100%)	58 (100%)	64 (100%)	40 (100%)	104 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	,101 ^A	1	,750
Pearson Chi-Square (Goods)	3,301 ^B	1	,069
Pearson Chi-Square (Services)	,052 ^C	1	,820
N of Valid Cases (works)	26		
N of Valid Cases (Goods)	58		
N of Valid Cases (Services)	104		

- a. 3 cells (75,0%) have expected count less than 5. The minimum expected count is ,77.
 b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,05.
 c. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 8,46.

Element 6: Relative score methods

Analysis 21: Recode: (absolute) / (only price is relative + only quality is relative) / (both relative)

Relative score methods

Problems	Relative methods			Total
	Absolute	1 of 2 is relative	Both relative	
NO	48 90,6%	92 79,3%	31 75,6%	171 81,4%
YES	5 9,4%	24 20,7%	10 24,4%	39 18,6%
Total	53 100,0%	116 100,0%	41 100,0%	210 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4,188 ^a	2	,123
N of Valid Cases	210		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,61.

Analysis 22: Kendall's tau-c (direction relativity, more relative scoring, more problems)

Kendall's tau-c

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal Kendall's tau-c	,108	,053	2,060	,039
N of Valid Cases	210			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Analysis 23: Recode: (absolute) and (both relative+ 1 of 2 relative)

Relative award methods

Problems	Relative methods		Total
	Absolute	Relative	
NO	48 90,6%	123 78,3%	171 81,4%
YES	5 9,4%	34 21,7%	39 18,6%
Total	53 100,0%	157 100,0%	210 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,914 ^a	1	,048
N of Valid Cases	210		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 9,84.

Analysis 24, 25 and 26: Absolute and Relative – Contract subject

Overview – Relative award methods – (Works / Goods and Services)

Problems Absolute or Relative	Works			Goods			Services		
	Absol.	Rel.	Total	Absol.	Rel.	Total	Absol.	Rel.	Total
NO	18 (90,0%)	13 (68,4%)	31 (79,5%)	11 (100,0%)	39 (83,3%)	50 (86,2%)	19 (86,4%)	71 (78,0%)	90 (79,6%)
YES	2 (10,0%)	6 (31,6%)	8 (20,5%)	0 (0,0%)	8 (17,0%)	8 (13,8%)	3 (13,6%)	20 (22,0%)	23 (20,4%)
Total	20 (100%)	19 (100%)	39 (100%)	11 (100%)	47 (100%)	58 (100%)	22 (100%)	91 (100%)	113 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	2,783 ^A	1	,095
Pearson Chi-Square (Goods)	2,172 ^B	1	,141
Pearson Chi-Square (Services)	,760 ^C	1	,383
N of Valid Cases (works)	39		
N of Valid Cases (Goods)	58		
N of Valid Cases (Services)	113		

- a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 3,90.
 b. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,52.
 c. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 4,48.

Analysis 27, 28 and 29: Absolute / 1 Relative / Both relative – Contract subject

Overview – Relative award methods (2x3 table) – (Works / Goods and Services)

Problems Absolute or Relative	Works				Goods			
	Absol.	1 Rel.	Both	Total	Absol.	1 Rel.	Both	Total
NO	18 (90,0%)	9 (60,0%)	4 (100%)	31 (79,5%)	11 (100,0%)	30 (90,9%)	9 (64,3%)	50 (86,2%)
YES	2 (10,0%)	6 (40,0%)	0 (0,0%)	8 (20,5%)	0 (0,0%)	3 (9,1%)	5 (35,7%)	8 (13,8%)
Total	20 (100%)	15 (100%)	4 (100%)	39 (100%)	11 (100%)	33 (100%)	14 (100,0%)	58 (100,0%)

Problems Absolute or Relative	Works			
	Absol.	1 Rel.	Both	Total
NO	19 (86,4%)	53 (77,9%)	18 (78,3%)	90 (79,6%)
YES	3 (13,6%)	15 (22,1%)	5 (21,7%)	23 (20,4%)
Total	22 (100%)	68 (100%)	23 (100%)	113 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	5,881 ^A	1	,053
Pearson Chi-Square (Goods)	8,181 ^B	1	,018
Pearson Chi-Square (Services)	,762 ^C	1	,508
N of Valid Cases (works)	39		
N of Valid Cases (Goods)	58		
N of Valid Cases (Services)	113		

- 4 cells (66,7%) have expected count less than 5. The minimum expected count is ,82.
- 3 cells (50,0%) have expected count less than 5. The minimum expected count is 1,52.
- 2 cells (33,3%) have expected count less than 5. The minimum expected count is 4,48.

Element 7: Scores on the basis of ranking**Analysis 30: Scoring on basis of ranking****Scores awarded on the basis of ranking**

Problems	Ranking		Total
	NO	YES	
NO	158 80,6%	16 88,9%	174 81,3%
YES	38 19,4%	2 11,1%	40 18,7%
Total	196 100,0%	18 100,0%	214 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,743 ^a	1	,389
N of Valid Cases	214		

- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,36.

Analysis 31, 32 and 33: Scoring on basis of ranking – Contract subject**Overview – scoring on basis of ranking – (Works / Goods and Services)**

Problems	Works			Goods			Services		
	Ranking >	NO	YES	Total	NO	YES	Total	NO	YES
NO	27 (79,4%)	5 (83,3%)	32 (80,0%)	52 (85,2%)	1 (100,0%)	50 (86,2%)	79 (78,2%)	10 (90,0%)	90 (79,5%)
YES	7 (20,6%)	1 (16,7%)	8 (20,0%)	9 (14,8%)	0 (0,0%)	8 (14,5%)	22 (21,8%)	1 (9,1%)	23 (20,5%)
Total	34 (100%)	6 (100%)	40 (100%)	61 (100%)	1 (100%)	62 (100%)	101 (100%)	11 (100%)	112 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	,049 ^A	1	,825
Pearson Chi-Square (Goods)	,173 ^B	1	,678
Pearson Chi-Square (Services)	,979 ^C	1	,322
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	62		
N of Valid Cases (Services)	112		

- 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,20.
- 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,15.
- 1 cells (25,0%) have expected count less than 5. The minimum expected count is 2,26.

Element 8: Flat award criteria**Analysis 34: 'Flat' award criteria****Lower boundaries**

Problems	Flat award criteria		Total
	NO	YES	
NO	85 93,4%	86 73,5%	171 82,2%
YES	6 6,6%	31 26,5%	37 17,8%
Total	91 100,0%	117 100,0%	208 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	13,864 ^a	1	,000
N of Valid Cases	208		

- 0 cells (0,0%) have expected count less than 5. The minimum expected count is 16,19.

Analysis 35, 36 and 37: Flat award criteria – Contract subject**Overview – Flat criteria – (Works / Goods and Services)**

Problems	Works			Goods			Services		
	'Flat' >	NO	YES	Total	NO	YES	Total	NO	YES
NO	16 (94,1%)	16 (69,6%)	32 (80,0%)	25 (100,0%)	27 (79,4%)	52 (88,1%)	44 (89,8%)	43 (71,7%)	87 (79,8%)
YES	1 (5,9%)	7 (30,4%)	8 (20,0%)	0 (0,0%)	7 (20,6%)	7 (11,9%)	5 (10,2%)	17 (28,3%)	22 (20,2%)
Total	17 (100%)	23 (100%)	40 (100%)	25 (100%)	34 (100%)	59 (100%)	49 (100%)	60 (100%)	109 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	3,683 ^A	1	,055
Pearson Chi-Square (Goods)	5,840 ^B	1	,045
Pearson Chi-Square (Services)	5,503 ^C	1	,019
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	59		
N of Valid Cases (Services)	109		

- 2 cells (50,0%) have expected count less than 5. The minimum expected count is 3,40.
- 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,97.
- 0 cells (0,0%) have expected count less than 5. The minimum expected count is 9,89.

Element 9: Extra lower boundaries**Analysis 38: Extra lower boundaries****Presence of lower boundaries**

Problems	Lower boudaries		Total
	NO	YES	
NO	151 83,0%	22 75,9%	173 82,0%
YES	31 17,0%	7 24,1%	38 18,0%
Total	182 100,0%	29 100,0%	211 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,855 ^a	1	,355
N of Valid Cases	211		

- 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,22.

Analysis 39, 40 and 41: Extra lower boundaries – Contract subject**Overview – Lower boundaries – (Works / Goods and Services)**

Problems	Works			Goods			Services		
	NO	YES	Total	NO	YES	Total	NO	YES	Total
Lower bound. >									
NO	27 (84,4%)	5 (62,5%)	32 (80,0%)	49 (89,1%)	3 (60,0%)	52 (86,7%)	75 (78,9%)	14 (87,5%)	89 (79,5%)
YES	5 (15,6%)	3 (37,5%)	8 (20,0%)	6 (10,9%)	2 (40,0%)	8 (13,3%)	20 (21,1%)	2 (12,5%)	22 (20,5%)
Total	32 (100%)	8 (100%)	40 (100%)	55 (100%)	5 (100%)	60 (100%)	95 (100%)	16 (100%)	111 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	1,914 ^A	1	,167
Pearson Chi-Square (Goods)	3,357 ^B	1	,067
Pearson Chi-Square (Services)	,630 ^C	1	,427
N of Valid Cases (works)	40		
N of Valid Cases (Goods)	60		
N of Valid Cases (Services)	111		

- d. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,60.
 e. 2 cells (50,0%) have expected count less than 5. The minimum expected count is ,67.
 f. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,17.

Element 10: Award criterion price that consist of various financial sub-criteria**Analysis 42: Price sub-criteria****Price sub-criteria**

Problems	Price sub-criteria		Total
	NO	YES	
NO	118 85,5%	61 74,4%	179 81,4%
YES	20 14,5%	21 25,6%	41 18,6%
Total	138 100,0%	82 100,0%	220 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4,192 ^a	1	,041
N of Valid Cases	220		

- a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 15,28.
 b. Computed only for a 2x2 table

Analysis 43, 44 and 45: Price sub-criteria – Contract subject**Overview – price sub – (Works / Goods and Services)**

Problems	Works			Goods			Services		
	NO	YES	Total	NO	YES	Total	NO	YES	Total
NO	25 (78,1%)	5 (83,3%)	30 (78,9%)	33 (100,0%)	21 (70,0%)	52 (85,7%)	60 (82,2%)	35 (76,1%)	95 (79,8%)
YES	7 (21,9%)	1 (16,7%)	8 (21,1%)	0 (0,0%)	9 (30,0%)	9 (14,3%)	13 (17,8%)	11 (23,9%)	24 (20,2%)
Total	32 (100%)	6 (100%)	38 (100%)	33 (100%)	30 (100%)	63 (100%)	73 (100%)	46 (100%)	119 (100%)

Chi-square test

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square (works)	,082 ^A	1	,774
Pearson Chi-Square (Goods)	11,550 ^B	1	,002
Pearson Chi-Square (Services)	,653 ^C	1	,566
N of Valid Cases (works)	38		
N of Valid Cases (Goods)	63		
N of Valid Cases (Services)	119		

- a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,26.
- b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 4,29.
- c. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 9,28.