Associate Professor Milos MILOSAVLJEVIC, PhD
E-mail: milos.milosavljevic@fon.bg.ac.rs
Department of Financial Management and Accounting
Faculty of Organizational Sciences
University of Belgrade
Researcher Sandro RADONOVANOVIC, PhD Candidate
E-mail: milos.milosavljevic@fon.bg.ac.rs
Department of Business Systems Organization
Faculty of Organizational Sciences
University of Belgrade
Professor Boris DELIBASIC, PhD
E-mail: boris.delibasic@fon.bg.ac.rs
Department of Business Systems Organization
Faculty of Organizational Sciences
University of Belgrade

EVALUATION OF PUBLIC PROCUREMENT EFFICIENCY OF THE EU COUNTRIES USING PREFERENCE LEARNING TOPSIS METHOD

Abstract. *Holding governments accountable for public procurement efficiency has been high on the agenda of public finance practitioners in the last few decades. The European Commission has developed a set of value-for-money indicators for public procurements within the Single Market Scoreboard. Although this matrix is actively used to rank countries, a number of downsides have been hitherto reported. This paper proposes preference learning (a machine learning method) for criteria weight estimation in combination with Technique for Order Performance by Similarity to Ideal Solution (as a multi criteria decision making technique) to re-evaluate the public procurement performance of the EU countries. This approach can be used for unbiased ex-post evaluations and focus of efforts and resources on critically important public procurement policies.*
Keywords*: Public Procurement; Ranking; Preference Learning; TOPSIS.*

1. Introduction

Holding governments accountable for public procurement efficiency has been high on the scholarly agenda in the last few decades. Recent studies have shown that public procurement is steadily transforming from a specialized

187

profession responsible for routine processes to deliver goods and services to the one delivering important governance and policymaking outputs (Trammell et al., 2019). This is by no means odd, as contracting authorities in the European Union only consume somewhere between 10 and 15% of the total Gross Domestic Product and nearly a third of the total governmental expenditures. Having the unrelenting public budget constraints in mind, lawmakers around the continent have been propounding a number of value-for-money performance measures for the efficient procurement of goods, works and services (Milosavljevic et al., 2016).

The European Commission has been actively working on developing functional performance measures and indicators aimed at comparing the relative performance of member countries with regards to their public procurement (Flynn, 2018). The systematization of this methodology is given as part of the Single Market Scoreboard. However, this Scoreboard has been widely criticized for various inconsistencies. For instance, a fundamental downside of the Single Market Scoreboard is the subjective use of weights for different value-for-money measures, which in turn creates a myriad of biases in the rankings (Milosavljević et al., 2019). Naturally, none of the proposed metrics would ultimately provide a comprehensive, unified and unbiased measure for public procurement efficiency, but both academia and practice are in search of improvements.

The benefits of using preference learning for weights estimation is the ability to adjust weights in such a manner that they express the opinion of an expert. More specifically, initial weights can be provided, as well as partial belief in ranking of alternatives (i.e. one alternative should be ranked higher than another alternative) for a small number of observations. Preference learning methodology will take into account experts' opinion and adjust weights in such a manner that they express the belief of the decision-maker. Since initial weights are optional, decision-maker can, using preference learning, obtain weights without specifying relations between criteria (or without even observing criteria). Additionally, usage of TOPSIS methodology is more suitable for the problem at hand (compared to simple additive weighting that is used by Single Market Scoreboard) since alternatives are compared to both positive and negative ideal solution. By comparing countries with the best performing and worst performing ones, one can interpret the results as closeness to the best possible value and the worst possible value, or how much room for improvement exists. The aim of our study is to utilize preference learning for criteria weights estimation for Technique for Order of Preference by Similarity to Ideal Solution [preference learning TOPSIS] method to rank European countries based on their public procurement performance. Specific goals of the study are to 1) develop data-driven weights and 2) to compare the findings with the 'raw' ranks provided by the EU Single Market Scoreboard.

To the best of our knowledge, a study of this kind has never been conducted before. Most of the concurrent studies are focused on a single, rather than on a holistic view of public procurement performance measurement (Plaček et al., 2020). The idiosyncratic nature of such studies might propel inadequate policy

188

implications. Finally, artificial intelligence and multi-criteria decision making (MCDM) techniques have been sparsely used in the realm of public procurements to date (Dotoli et al., 2020). The use of any MCDM technique is limited to the procurement process, rather than to the nation-wide relative performance ranking.

The remainder of the paper is organized in the following order. Section 2 reviews the literature related to public procurement value-for-money measurement. Section 3 thoroughly elaborates on the methods and data used for the analysis. Section 4 presents the ranking of European countries based on their public procurement efficiency based on machine learning (preference learning) and MCDM (TOPSIS) framework. Section 5 contextualizes the results. The final part of this section is reserved for concluding remarks.

2. Literature review

An important output of the European Single Market is the Single Market Scoreboard. The Single Market Scoreboard is envisioned as a tool that generates peer pressure to enhance regulatory compliance (Egan, 2010). This Scoreboard provides a wide spectrum of useful rankings. For the purpose of this study, the one that attracts particular attention is the scoreboard that addresses public procurements among the countries of the Old Continent. This specific part of the overall scoreboard was used for the first time in 2014 and immediately provided important insights into the procurement performance at the national level (Flynn, 2018). Each year, EEA countries (EU countries plus Switzerland, Norway, Iceland and Liechtenstein) are categorized in three loosely constructed strata based on 12 weighted performance measures.

The aim of the Scoreboard is to reflect on various value-for-money and other important performance measures related to public procurements. Specifically, attention should be drawn to (1) competition, (2) bureaucracy, (3) transparency, (4) openness of the market, (5) economies of scale, (6) quality of procured goods, services or works, (7) local economic growth and development of SME, and (8) reporting quality, all of which are described in more detail below.

Competition. The contracting authority is supposed to create a stimulating environment and level the field for all players so that none of the suppliers has an unfair advantage. The open competition is one of the core principles in Europe on both national and supranational – the EU level. Without the competition, there is no assurance that the price of the lowest bidder will be fair and reasonable. It is, accordingly, believed that the competition leads to financial savings (Hanák & Serrat, 2018). As inferred by Broms et al. (2019), 'single bidding invites an unhealthy relationship between entrenched [political] parties and certain firms, and, therefore, increases the risk of elite collusion'. Following the aforementioned, a number of bidders, as an indicator, has been noted as the best proxy for the open competition. Unfortunately, many European countries still cope with having more than one bidder for numerous calls for bids. For instance, in 2018 only two EEC

_____

countries (in specific, Sweden and Iceland) exceeded in this criterion by having equal or less than 10% of calls for bids with a single bidder.

Bureaucracy. Unlike the private sector procurement, their public counterpart is rather bureaucratic, at least for two reasons – there is a towering potential for political influence and the sector is highly regulated (Schapper et al., 2006). Public procurement is anecdotally synonymous with bureaucracy. A number of indicators are related to the degree of bureaucracy in public procurements, such as the quality of regulatory framework, the managerial capacity of the contracting authority, and post-award behavior (Plaček et al., 2020). However, the most comprehensive and compelling measure of efficiency is the 'decision speed', as it encompasses all bureaucracy-related inefficiencies by assessing 'the period of time between the deadline for receipt of tenders and the actual awarding of the contract' (Flynn, 2018).

Transparency. As implied by Halonen (2016) '[t]he principle of transparency is one of the fundamental and underlying principles of EU public procurement rules.' Moreover, transparency in the public sector is strongly associated with accountability and corruption as depicted in the bibliometric study conducted by Lyrio et al. (2018). The current body of knowledge does not unequivocally support transparency as an important dimension of public procurement efficiency, as the evidence still remains scant and mixed. Whilst Atkinson (2019) states that 'a willingness to entertain bids does not alone ensure that additional bids will be forthcoming', Bauhr et al. (2019) argue that overall tender transparency reduces corruption risks substantially, but this effect is mainly driven by the possible horizontal monitoring.

Openness of public procurement market. As aforementioned, public procurements make a large proportion of governmental expenditures. Carboni et al. (2017) infer that '[p]ublic procurement is a politically sensitive lever in the governments' hands, often used for rising political consensus rather than achieving value for money for public purchases.' There is no better way to proliferate contracting authorities' distortive agenda but unpublishing the tender. The measure often referred to as 'publishing rate' seems to be an obvious indicator of the openness of the market. It should be noted that the European contracting agencies are not obliged to publish the call for bids if the value of the bid falls behind the legally set threshold.

Economies of scale. Economies of scale refer to aggregate purchasing where two or more contracting authorities combine their purchases and thus approach the supply side. For more than a decade, a number of countries propounded contracting authorities to either pool or share purchasing volumes, information, and/or resources (Schotanus et al., 2008). In a way, contracting authorities form buyer to buyer alliances which assumably creates economies of scale, reduce the overall purchasing price and improve efficiency. However, combining purchasing powers might have negative overall effects – a number of distortions that can offset the economies of scale (Decarolis, 2018).

190

Quality of procured goods, services or works. Value-for-money in public procurement does not imply that the purchase of goods, works or services has to be made solely on the basis of the lowest price criterion. Directive 2004/18/EC sets the application for the most economically advantageous tender as well. The lowest price as a criterion is preferable when the goods, services or works are highly standardized. However, when the goods required by the contracting authority are complex, need to create new value or might even be unknown at the time the tender is announced, the award criterion seems to be the most economically advantageous (Detelj et al., 2015). Whenever the quality of goods is considered as an utterly important characteristic of procurement procedure, contracting authorities should employ the most economic advantageous tendering.

Local economic growth and development of SMEs. Public procurements are not simply meant to optimize only the price or quality for the contracting authority. The European public procurement system is envisaged as a tool for the implementation of different national, regional and local sustainability policies, particularly those related to the growth of SMEs (Loader & Norton, 2015). Large companies are still predominant in tendering procedures across the EU. A number of factors still hinder SMEs involvement in public procurements, such as onerous and lengthy bidding processes, information about procurement opportunities failing to reach SMEs, bundled contracts, and lack of a critical mass of internal resources' (Ancarani et al., 2019). Accordingly, the public procurement performance matrix measures indicators such as 'SME contractor', 'SME bids', and 'Procedures divided in lots.'

Reporting quality. Sharing the relevant information on public procurement procedures is essential in pre-tender, tender and post-tender phases (Holma et al., 2020). For the purpose of performance measurement, however, only awarded contracts and post-tender reporting quality make a useful case. Publishing data on awarded contracts is a useful source of information for both the extent and any future competitions. After the tender is completed, the contracting authority should provide sufficient amount of data for general audience (i.e. competitors, regulators, taxpayers, and analysts). Flynn (2018) emphasizes that both buyers and suppliers benefit from the disclosed information on awards containing all details on the call for bids, and whom the bid was awarded.

3. Methods

The general analytical framework encompasses three broad steps: 1) Data driven calculation of weights for the public procurement value-for-money indicators, 2) TOPSIS-based country rankings for three subsequent years, and 3) Quasi-clustering based on two dimensions – TOPSIS ranking and the size of public procurements, has been displayed in Figure 1.

As for the first step, we utilized a preference learning and preference elicitation technique for weight derivation (Lepird et al., 2015). In many decision-

_____

making setups, one observes multiple conflicting criteria. For example, higher performance means higher costs. Handling conflicting criteria and making trade-offs are at the core of decision-making techniques though estimation of weights of criteria. However, in presence of many criteria and many alternatives setting weights appropriately is a very challenging task. Therefore, we use preference learning or series of pairwise comparisons between alternatives where the decision-maker, or in this research experts in the area of public procurement, states whether one alternative is preferred, not preferred, or indifferent compared to another one. Preference learning methodology will take into account the preferences of the decision-maker and calculate weights of the criteria in such a manner that the error (the difference between expressed preferences and real ranking) is minimized. Usage of preference learning and preference elicitation has shown an advantage over traditional criteria weight estimation, for example using domain knowledge for weights estimation or usage of group decision-making techniques, because an individual can compare a small number of alternatives and the preference elicitation methodology will optimally yield in criteria weights (Astudillo & Frazier, 2020).

Mathematically, weights are obtained by finding the maximum a-posteriori of distribution of weights and covariance matrix, or:

$$p(\mu, \Sigma | D) \propto P(D|\mu, \Sigma)P(\mu, \Sigma) \tag{1}$$

where μ present mean values of weights, Σ covariance matrix, and D preferences of the expert. Steps for solving this problem are as follows (Lepird et al., 2015).

Sub-step 1.1. Setting prior weights: Although this step is optional, a domain expert can express beliefs about the weights of criteria. Adding prior ($P(\mu, \Sigma)$) means that one has some domain-knowledge or wanted values for the weight for some criteria and it is common in Bayesian optimization to add priors as beliefs of the underlying phenomena. The idea of the prior weights is to provide a starting point for the multivariate distribution of weights. If the decision-maker has good prior weights, then the data will express the same behavior and, as a result, the final weights will be similar. If priors are not given, they are assumed to have a normal distribution with an expected value equal to zero.

It is important to take into account for multi-collinearity in data. Therefore, the covariance matrix Σ needs to be calculated. Most commonly, covariance matrix Σ is deduced from data at hand (from matrix M, explained in the text below).

Having the above-mentioned in mind, an expert provides a vector of expected weights of criteria.

$$\mu = [\mu_1, \dots, \mu_n] \tag{2}$$

where each criterion C_j has its corresponding expected weight μ_j that presents the belief of the decision-maker about the weights of the criteria.

Sub-step 1.2. Adding preferences about alternatives: Since the idea of preference learning is to help deriving criteria weights based on the preferences

192

between alternatives, the decision-maker needs to express belief in form of preferences between alternatives. The decision-maker can make three types of preferences. Namely, the decision-maker can prefer one alternative to another ($A_i \succ A_j$), do not prefer one alternative to another ($A_i \prec A_j$), and be indifferent between two alternatives ($A_i \sim A_j$). By expressing a preference for one alternative to another, the decision-maker expresses that one alternative should be ranked higher to the other one. However, if indifference is expressed then the decision-maker is indifferent than one alternative is higher ranked than the other one, or vice versa.

Preference learning methodology allows decision-maker to express a small number of preferences, and based on them calculate the likelihood of the preferences given distribution of criteria weights and covariance matrix. More specifically, decision-maker need to express (a small number of) preferences D over alternatives A where:

$$D = D_{\succ} \cup D_{\sim}$$
$$D_{\succ} = \{A_k \succ A_q, \quad k \neq q; k = 1, \ldots, m; q = 1, \ldots, m\} \qquad (3)$$
$$D_{\sim} = \{A_k \sim A_q, \quad k \neq q; k = 1, \ldots, m; q = 1, \ldots, m\}$$

Sub-step 1.3. Calculating the likelihood of the preferences and new weights:

Having defined prior distribution of weights ($P(\mu, \Sigma)$) and expressed preferences D likelihood of the preferences given the distribution of weights $P(D|\mu, \Sigma)$ is calculated. It is calculated from the data at hand (matrix M). After the calculation of likelihood, a-posteriori weights are obtained using gradient descent to solve maximum a-posteriori estimation. Since computation is mathematically very complex and we used an existing approach we refer the interested reader to the paper (Lepird et al., 2015) where a detail explanation can be found.

As an additional value of preference elicitation methodology, one can ask for a comparison between two alternatives that will help in better weight estimation. This feature of the methodology was used during the weight estimation phase.

As for the second step, we incorporate previously generated weights into the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). The standard version of the TOPSIS was developed by Hwang & Yoon (1981). Nowadays, it is considered to be a standard multi-criteria decision-making method used for the ranking of alternatives. The method is based on the calculation of the positive ideal solution $[S_i^+]$ and negative ideal solution $[S_i^-]$, whereby the optimal alternative is geometrically closest to the positive ideal and farthest from the negative ideal solution. The general TOPSIS process consists of several activities used as steps for the purpose of this study.

Sub-step 2.1. Establishing a decision matrix for the ranking. The structure of the matrix is as follows:

193

_____

$$M = \begin{matrix} M \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{matrix} \begin{matrix} C_1 & C_2 & \dots & C_n \\ \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \\ x_{n1} & x_{n2} & & x_{nm} \end{bmatrix} \end{matrix} \qquad (4)$$

where A stands for countries, and C stands for public procurement performance measure.

Sub-step 2.2. Conducting a vector normalization of the decision matrix using the following formula:

$$\bar{x}_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^{m} X_{ij}^2}} \qquad (5)$$

Sub-step 2.3. Calculating weighted normalized matrix, as shown in formula below:

$$v_{ij} = \omega_j \times \bar{x}_{ij} \qquad (6)$$

where: v_{ij}– weighted normalized matrix; and ω_j– weight for criterion C_i. The weights were imported from the step one described in Step 1.

Sub-step 2.4. Setting the positive ideal solution and negative ideal solution where the most favorable (Max-Min) value is considered as the positive ideal solution for the given column. The ideal negative is the Min-Max counterpart. For the positive ideal solution, the calculus is given below:

$$IPS(v^+) = \{v_1^+, v_2^+, \dots, v_n^+\} \qquad (7)$$

where:

$$v_j^+ = \{(\max \wedge \min v_{ij} \leftrightarrow j \in J); (\min \wedge \max v_{ij} \leftrightarrow j \in J^*) \qquad (8)$$

for J being the most favorable and J^* being the least favourable value for column j. The calculus for the negative ideal solution is the opposite.

Sub-step 2.5. Calculating separation measures of each alternative from ideal positive and ideal negative values:

$$S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, i = 1,2,\dots,m \qquad (9)$$

194

$$S_i^- = \sqrt{\sum\nolimits_{j=1}^{n} \left(v_{ij} - v_j^-\right)^2}, i = 1,2,\dots,m \tag{10}$$

where S_i^+, S_i^- are distances from the positive ideal and negative ideal solution, respectively.

Sub-step 2.6. Calculating the closeness coefficients for the given alternatives CC_i, as follows:

$$CC_i = \frac{S_i^-}{S_i^- + S_i^+} \tag{11}$$

Sub-step 2.7. Ranking the alternatives in descending order based on CC_i. Intuitively, the best possible value would have $CC_i = 1$ and that would be a positive ideal solution v_j^+ (in that case $S_i^+ = 0$). Similarly, the worst value of CC_i would be 0, or an alternative equal to the negative ideal solution.

As for step three, the ranking generated following the framework described above itself provides a basis for a vivid discussion. Nonetheless, more detailed and valuable information can be obtained when public procurement efficiency ranking is compared against the size of public procurements per country. The following sub-steps were used for the quasi-clustering:

Sub-step 3.1. Calculation of mean for both dimensions (the size of public procurement for observed countries, and their TOPSIS generated CC_i).

The data analyzed in this paper come from two secondary sources. The first one being the Public Procurement Single Market Scoreboard (European Commission, 2020) for the performance indicators. Only the data for three consecutive years (FY2017, FY2018 and FY2019) were used since the indicators used in years prior to 2017 have slightly changed over time, which particularly refers to SME inclusion and reporting quality. Some countries have been excluded from the study for various reasons (Slovenia for having incomplete data and Norway, Iceland, Switzerland and Liechtenstein for missing the values on publication rates in Tenders Electronic Daily - TED). The second source of data are OECD statistics for the value of public procurement by country in Government at a Glance publication series. Hitherto, no data reliability issues have been reported for OECD statistics on public procurement size by country. Weights are obtained by implementing preference learning and preference elicitation methodology in Python using library PyPbl. Based on weights we performed TOPSIS method, in Python as well.

_____

3. Results

For the purpose of the analysis, data taken into account include 12 public procurement indicators from 28 European countries. Some countries have been excluded from the analysis due to incomplete data (Slovenia, Norway, Iceland and Liechtenstein).

First, we describe the obtained weight factors presented in Table 1. As it can be observed, due to the inserted priors, values are close to the priors, but with differences. Criteria Single bidder and criteria No calls for bids have its importance increased compared to the proposed prior value, while criteria Publication rate is valued less. This is an indicator that, for a given preference, these criteria are of more importance compared to the Publication rate. Similarly, criteria Cooperative procurement and Decision speed have shown greater value for the decision-making than expected (expected value was 1). Criteria that were considered as of less importance are also of less importance in preference elicitation methodology, especially criteria Missing calls for bids, which is considered as almost unimportant (in 2019 its weight was 0.01).

Table 1. Original weights and weights obtained using preference elicitation methodology

Criteria	Original	2017	2018	2019
Single bidder	3	3.147	3.316	3.304
No calls for bids	3	3.139	3.138	3.142
Publication rate	3	2.937	2.486	2.409
Cooperative procurement	1	1.528	1.500	1.524
Award criteria	1	1.084	1.071	0.970
Decision speed	1	1.251	1.600	1.379
SME contractor	1/3	0.274	0.052	0.213
SME bids	1/3	0.319	0.676	0.325
Procedures divided in lots	1/3	0.198	0.007	0.309
Missing calls for bids	1/3	0.085	0.080	0.010
Missing seller reg. No	1/3	0.204	0.218	0.192
Missing buyer reg. No	1/3	0.254	0.455	0.478

Using obtained weights, we performed the TOPSIS method as previously explained. The results are presented in Table 2 sorted by rank obtained in 2018. The best performing countries are in the north of Europe. The best ones are Denmark (DK), Finland (FI), United Kingdom (UK), Sweden (SE), and Ireland (IE). Their scores are continuously in the top five performing countries.

On the other hand, southeast European countries have the worst performing procurement practice. Cyprus (CY) and Bulgaria (BG) are by far the worst-performing countries having a TOPSIS closeness coefficient around 0.2.

This indicates that Cyprus and Bulgaria are very close to the negative ideal solution
(worst values for every criterion). Romania (RO) is also present in the worst-
performing countries. However, their score is closer to the average TOPSIS
closeness coefficient score than the worst countries.

Table 2. Results of the TOPSIS method

Country	TOPSIS 2017	Rank 2017	TOPSIS 2018	Rank 2018	TOPSIS 2019	Rank 2019
FI	0.780	2	0.782	1	0.747	2
UK	0.770	3	0.757	2	0.712	3
DK	0.785	1	0.751	3	0.748	1
SE	0.741	5	0.743	4	0.708	4
IE	0.717	6	0.738	5	0.694	5
EE	0.693	7	0.731	6	0.641	10
BE	0.767	4	0.704	7	0.673	7
FR	0.689	8	0.698	8	0.684	6
LT	0.682	9	0.691	9	0.642	9
DE	0.641	14	0.665	10	0.628	13
MT	0.682	10	0.664	11	0.635	11
HR	0.610	20	0.658	12	0.621	15
AT	0.671	11	0.648	13	0.631	12
PT	0.637	16	0.638	14	0.565	20
NL	0.621	18	0.634	15	0.613	18
HU	0.597	22	0.630	16	0.648	8
SK	0.638	15	0.626	17	0.614	17
LU	0.664	12	0.618	18	0.614	16
IT	0.609	21	0.595	19	0.563	21
EL	0.613	19	0.576	20	0.571	19
ES	0.560	23	0.559	21	0.562	22
LV	0.662	13	0.549	22	0.626	14
PL	0.624	17	0.546	23	0.558	23
CZ	0.505	24	0.492	24	0.482	24
BG	0.315	26	0.304	25	0.269	27
RO	0.411	25	0.293	26	0.452	25
CY	0.179	27	0.200	27	0.276	26

_____

When compared against the actual rankings for three consecutive years, overall public procurement performance denotes a number of interesting findings. For instance, Croatia has the largest negative difference. This is to some extent expected, as this country has slightly manipulated the real data. For instance, the legal obligation in Croatia is to use most economically advantageous tendering which stipulates the rise in Award criteria indicator, without the real effects on public procurement efficiency. On the other side, the largest positive difference can be attributed to Ireland and France.

The ranking itself provides a basis for a vivid discussion. Nonetheless, more detailed and valuable information can be obtained when public procurement efficiency ranking is compared against the size of public procurements per country, as displayed in Figure 1. We present results for year 2018 due to data availability.
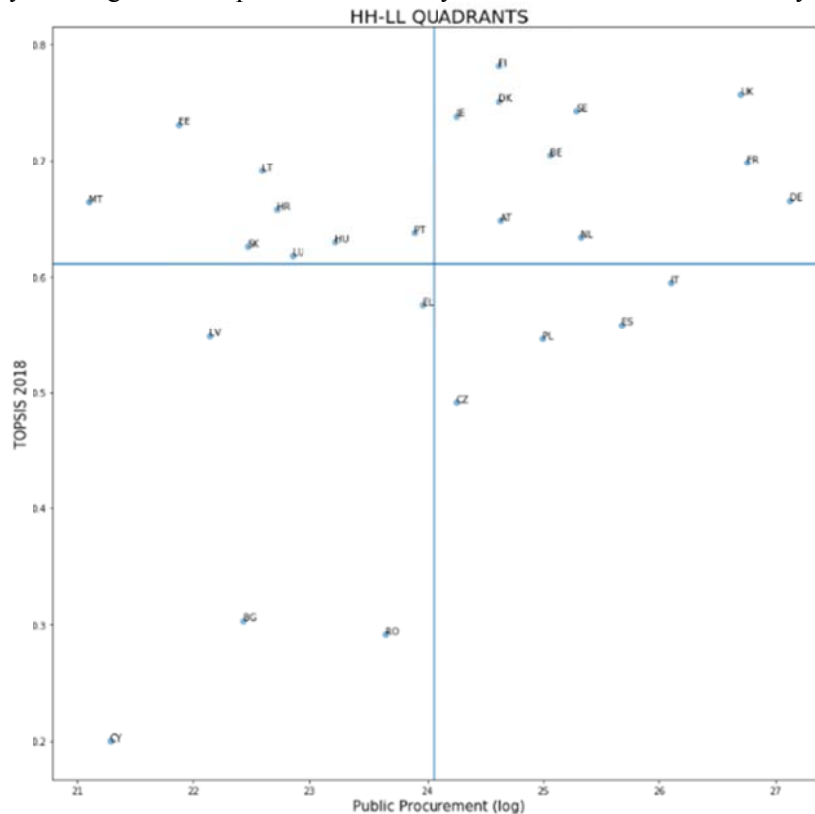

Figure 1. Scatter plot diagram for size and overall public procurement efficiency

The scatter plot diagram clusters the countries in four groups (HH, HL, LH, and LL) using public procurement size (natural logarithm) and ranks as

198

dimensions. The top tier countries or 'superstars' are in the HH quadrant. These countries might serve as role models, and the majority of them come from Western and Northern Europe. The countries clustered in the HL quadrant are 'small tigers', and they predominantly encompass Eastern European countries. Five countries (Greece, Latvia, Romania, Bulgaria, and Cyprus) are in the bottom-left (LL) or the 'snail quadrant'. Finally, only most dangerous L-H quadrant (with Italy, Spain, Poland and Czech Republic).

4. Discussion

Key findings and contributions. The aim of our study is to create a comprehensive matrix for the performance measurement of public procurements for 28 European countries. For this purpose, we used the preference learning TOPSIS method to rank the countries based on 12 Single Market Scoreboard indicators. The contribution is usage of preference learning for criteria weight estimation. This approach allows decision-maker to utilize small number of preference relations between alternatives to derive weights for each criterion, as well as, use preference elicitation for suggestion of preferences that will improve criteria weight estimation.

From a broader point of view, our study contributes to the New Public Management stream of suppositions by adding to the growing body of knowledge on performance measurement in the public sector. Given the multidimensionality of governmental functions, and the plurality of the objectives supposed to be met by contracting authorities, this study emphasizes a critical perspective on current models of public procurement performance measurement and adopts a methodologically grounded approach for the re-evaluation of European national public procurement systems. This fertile research area still requires novel approaches, techniques, methods and instruments to capture the holistic view on whether the public funds meet the desired outcomes.

Policy implications. The first practical implication of the study is that it provides a much-needed comprehensive measure of public procurement performance at the national level. The ranking given in Table 2 provides a reliable basis for policyholders and decision-makers on both supranational and national level to set major objectives that each public procurement system needs to address and fulfil in order to improve the performance of public procurements and increase value-for-(tax payers')-money accordingly. This ranking is important given that the only comparable performance matrix is given in the European Commission's Public Procurement Single Market Scoreboard.

Policy holders and decision-makers of the European countries can use the findings of this paper to: a) set the role models in public procurement efficiency (FI, DK, UK, SE, IE, BE, FR, DE, AT and NL), b) loosely inspect 'small tigers' (EE, LT, MT, HR, PT, HU, SK and LU), c) pay particular attention and set specific country-related targets to countries that have gigantic consumption of public funds

but under-perform in terms of public procurement efficiency (IT, ES, PL and CZ), and d) reinforce countries in the 'snail quadrant' by setting general EU-based targets (EL, LV, BG, RO and CY).

By using the same methodological steps, policyholders can re-run the analysis for every year and examine future over- and under-performers. Given the nature of the weights, particularly the fact that they are data-driven, the rankings generated in this study allow for an unbiased formulation of the best and worst performers. Thus, ex-post analyses of public procurement systems can be used for adequate comparisons.

Limitations and further recommendations. As is often the case with quantitative studies, this study has a number of flaws, some being fundamental some being technical by nature. As for the fundamental ones, the rankings generated using the data-driven TOPSIS suffer from indices used in the decision matrix. First, each and every indicator is subject to a thorough judicious judgement of whether they even represent the value-for-money measure. If we only take a look at the decision speed as an indicator – an interrogative remains if the longer decision process reveals inefficiencies or simply draws the data on complexity of procured goods and procedures used for procurement. By analogy, we might question all the other indicators used in the Scoreboard and accordingly in this study. Second, the set of indicators is merely composed of those requiring low effort to acquire information. A number of scholarly and practice-related calls were made to introduce measures of corruption and fraud within public procurement (Fazekas & Kocsis, 2017). These study flaws open an avenue for further research. A detailed examination and refinement of extent indicators and an introduction of new ones (preferably those measuring corruption, green initiatives, innovations etc.) would complete the decision matrix for the improved TOPSIS analyses.

As for the technical limitations, the most noteworthy one is that data-driven weights were used to determine the importance of indicators. As the data included in calculus change, the year-over-year changes in weights will eventually occur. Although the major strength of this approach is the objective nature of weights, it simultaneously means that only ex-post policy analyses are possible. Further research, however challenging it might be, could focus on the ultimate scales for all value-for-money public procurement indicators. Another important technical downside of this study is the use of data from three years span (from 2017 to 2019). It would be highly beneficial for future studies to conduct an analysis of time series, thus capturing the evolutionary characteristics of nation-wide procurement efficiency.

Concluding remarks. An extensive body of public management knowledge has impressively been developing around the output performance of public entities. Even though public procurement drains colossal tax-payers' funds, and both directly and indirectly affects the socio-economic development of nations, the scholarly and practical studies (quantitative and qualitative, single-country and comparative) in this field are still in rather short supply. This paper is an attempt to

200

shed light on European country-level efficiency by using preference learning
TOPSIS as a multi-criteria decision-making technique. Given the parsimony of
concurrent approaches to the measurement of public procurement performance, the
results of this study can serve as an indication of the adequacy of public
procurement systems of European countries.

REFERENCES

[1] **Ancarani, A., Di Mauro, C., Hartley, T., Tátrai, T. (2019),** *A Comparative Analysis of SME Friendly Public Procurement: Results from Canada, Hungary and Italy*. *International Journal of Public Administration*, 42(13): 1106–1121;

[2] **Astudillo, R., Frazier, P. (2020, June),** *Multi-Attribute Bayesian Optimization with Interactive Preference Learning*. *International Conference on Artificial Intelligence and Statistics* (pp. 4496-4507);

[3] **Atkinson, C. L. (2019),** *Full and Open Competition in Public Procurement: Values and Ethics in Contracting Opportunity*. *International Journal of Public Administration,* 43(13): 1169–1182;

[4] **Bauhr, M., Czibik, Á., Licht, J., Fazekas, M. (2019),** *Lights on the Shadows of Public Procurement: Transparency as an Antidote to Corruption*. *Governance*, 33(3): 495–523;

[5] **Broms, R., Dahlström, C., Fazekas, M. (2019),** *Political Competition and Public Procurement Outcomes*. *Comparative Political Studies*, 52(9): 1259–1292;

[6] **Carboni, C., Iossa, E., Mattera, G. (2017),** *Barriers to Public Procurement: A Review and Recent Patterns in the EU* (No. 92). *IEFE, Center for Research on Energy and Environmental Economics and Policy*, Universita' Bocconi, Milano, Italy;

[7] **Decarolis, F. (2018),** *Procurement Centralization in the EU: The Case of Italy*. *CEPR Discussion Paper* No. DP12567;

[8] **Detelj, K., Jagric, T., Markovic-Hribernik, T. (2016),** E*xploration of the Effectiveness of Public Procurement for Innovation: Panel Analysis of EU Countries' Data*. Lex Localis - *Journal of Local Self-Government*, 14(1): 93-114;

[9] **Dotoli, M., Epicoco, N., Falagario, M. (2020),** *Multi-Criteria Decision Making Techniques for the Management of Public Procurement Tenders: A Case Study*. *Applied Soft Computing*, 88: 106064;

[10] **Egan, M. (2010),** *The Single Market*. In *M. Cini, & N.-S. Borragán,* (Eds.) European Union Politics (fifth edition) (pp. 258-274). Oxford, UK: Oxford University Press;

[11] **European Commission (2020), Public Procurement Single Market Scoreboard**, Brussels: European Commission;

[12] **Fazekas, M., Kocsis, G. (2017),** *Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data*. *British Journal of Political Science*, 50(1): 155–164;

_____

[13]**Flynn, A. (2018),** *Measuring Procurement Performance in Europe*. *Journal of Public Procurement*, 18(1): 2–13;

[14]**Halonen, K.-M. (2016),** *Disclosure Rules in EU Public Procurement: Balancing between Competition and Transparency*. *Journal of Public Procurement*, 16(4): 528–553;

[15]**Hanák, T. & Serrat, C. (2018),** *Analysis of Construction Auctions Data in Slovak Public Procurement*. *Advances in Civil Engineering*, 2018: 1–13;

[16]**Holma, A.-M., Vesalainen, J., Söderman, A., Sammalmaa, J. (2020),** *Service Specification in Pre-Tender Phase of Public Procurement - A Triadic Model of Meaningful Involvement*. *Journal of Purchasing and Supply Management*, 26(1): 100580;

[17]**Hwang, C.L., Yoon, K., (1981),** *Multiple Attributes Decision Making Methods and Applications*. *Berlin: Springer;*

[18]**Lepird, J. R., Owen, M. P., Kochenderfer, M. J. (2015),** *Bayesian Preference Elicitation for Multiobjective Engineering Design Optimization*. *Journal of Aerospace Information Systems*, 12(10): 634-645;

[19]**Loader, K., Norton, S. (2015),** *SME Access to Public Procurement: An Analysis of the Experiences of SMEs Supplying the Publicly Funded UK Heritage Sector*. *Journal of Purchasing and Supply Management*, 21(4): 241–250;

[20]**Lyrio, M. V. L., Lunkes, R. J., Taliani, E. T. C. (2018),** *Thirty Years of Studies on Transparency, Accountability, and Corruption in the Public Sector: The State of the Art and Opportunities for Future Research*. *Public Integrity*, 20(5): 512–533;

[21]**Milosavljevic, M., Milanović, N., Benković, S. (2016),** *Politics, Policies and Public Procurement Efficiency: A Quantitative Study of 25 European Countries*. *Lex Localis - Journal of Local Self-Government*, 14(3): 537–558;

[22]**Milosavljević, M., Dobrota, M., Milanović, N. (2019),** *A New Approach to the Evaluation of Public Procurement Efficiency Among European Countries*. *European Review*, 27(2): 246–259;

[23]**Plaček, M., Ochrana, F., Schmidt, M., Nemec, J., Půček, M. (2020),** *The Factors Causing Delays in Public Procurement: The Czech Republic versus the UK*. *Public Money & Management*, 40(2): 131–139;

[24]**Schapper, P. R., Veiga Malta, J. N., Gilbert, D. L. (2006),** *An Analytical Framework for the Management and Reform of Public Procurement*. *Journal of Public Procurement*, 6(1/2): 1–26;

[25]**Trammell, E., Abutabenjeh, S., Dimand, A.-M. (2019),** *A Review of Public Administration Research: Where Does Public Procurement Fit in?* *International Journal of Public Administration*, 43(8): 655–667.