On-field Performance Evaluation in Soccer based on Network Data Envelopment Analysis

Thanasis Bouzidis [1]

Department of Economics, University of Macedonia, 156 Egnatia Str., 540 06, Thessaloniki, Greece; thabou@uom.edu.gr; +30 2310 891 781

October 2019

[1] ***Acknowledgements*:** I would like to thank Prof. Dr. G. Karagiannis, Dr. S. A. Kourtzidis, and Dr. I. N. Pinopoulos for their helpful comments on an earlier draft of this paper. Special thanks to two anonymous referees for their constructive suggestions. Finally, special thanks to *Galanis Sports Data* for generously supplying the game statistics used in this paper.

On-field Performance Evaluation in Soccer based on Network Data Envelopment Analysis

Abstract

In this paper, we present for the first time the on-field production process of soccer teams as a *mixed* (serial and parallel) structure two-stage network system. According to this, the *first* stage consists of two distinct sub-processes (offense and defense) that operate in *parallel*. These respectively use players' offensive and defensive actions as inputs to produce two different intermediate measures, namely, goals scored and prevention of goals conceded. These, in turn, are the inputs of the *second* stage (points' accumulation sub-process) that produces accumulated points. Furthermore, based on a two-stage network Data Envelopment Analysis (DEA) model, we estimate the offensive, defensive, and athletic efficiency of soccer teams during a league season. According to our proposed framework, these three different efficiency scores are provided (for each soccer team under evaluation) by a *single* linear programming problem. For this purpose, aggregate-over-games statistics from the 2013-14 Greek premier soccer league are used.

Keywords

Two-stage Network DEA; On-field Production Process; Sub-process; Soccer Teams; Offensive / Defensive / Athletic Efficiency; League Season

On-field Performance Evaluation in Soccer based on Network Data Envelopment Analysis

1 Introduction

Since the study of Carmichael *et al*. (2001), on-field performance evaluation in soccer seeks to answer whether athletic efficiency leads teams to sporting success. To this end, Espitia-Escuer and García-Cebrián (2004, 2006) proposed to estimate the athletic efficiency of soccer teams during a league season based on a single-stage Data Envelopment Analysis (DEA) model. According to this model, the on-field actions of soccer teams are considered as inputs and their accumulated points as a single output. This variable is preferred over the number of teams' victories (or their winning percentage) during a league season (Dawson *et al*., 2000, p. 407). This is due to the fact that (in each soccer game) there are *three* probable results, each of which is rewarded differently. In particular, victories (losses) receive three (zero) points and ties one point for each soccer team.

Some years later, the studies of Boscá *et al*. (2009) and Sala-Garrido *et al*. (2009) took a step further in the on-field performance evaluation of soccer teams. Specifically, these studies proposed to *separately* estimate the offensive and defensive efficiency of soccer teams during a league season based on *two* different single-stage DEA models. According to these models, teams' goals scored and the inverse of teams' goals conceded are independently used as a single output. In this way, the above studies sought to clarify which sub-process (offense or defense) was more important in each league season under consideration. For this purpose, they suggested relating estimated (offensive and defensive) efficiencies with the accumulated points of soccer teams. However, to *a priori* determine a direct relationship between teams' (offensive and defensive) efficiencies and accumulated points seems rather difficult. This is because different soccer teams generally achieve different efficiency scores in offense and defense.

This fact led García-Sánchez (2007) and (some years later) Rossi *et al*. (2018) to employ a multiple-stage DEA model for evaluating the performance of soccer teams during a league season. According to this model, the different offensive and defensive efficiencies of soccer teams are estimated separately at the *first* stage; then, at the *second* stage, the estimated (offensive and defensive) efficiencies are used as inputs and teams' accumulated points as a single output in order their athletic effectiveness to be measured. By this means, managers are provided with sub-process-specific guidance to improve the on-field performance of their soccer team. Furthermore, it is discovered whether it would have been more effective for soccer teams under evaluation to be efficient in offense or defense.

Alternatively, Kern *et al*. (2012) employed a two-stage network DEA model for evaluating the performance of soccer teams during a league season. [2] According to this model, there is only a single intermediate measure that is simultaneously used as both the output of the *first* and the input of the *second* stage. In addition, the two stages of the production process in soccer are *independently* evaluated (for each team) by *two* different linear programming problems. This is why the specific network DEA model fails to address any conflict (regarding the optimal level of the

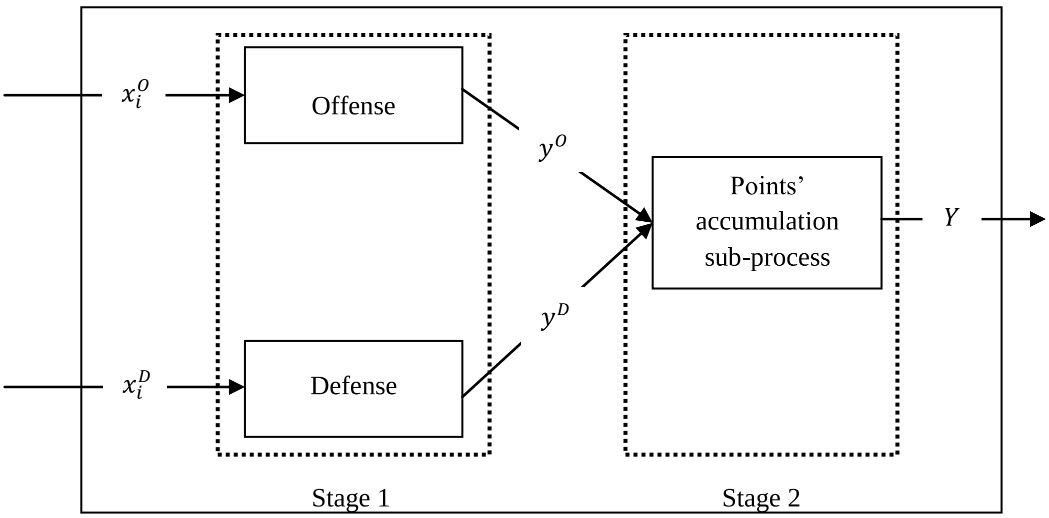
[2] For a complete survey of two-stage network DEA models see Halkos *et al*. (2014).

intermediate measure) between its two stages. To better understand this issue, consider that the *second* stage may have to shrink its input to become efficient. Such an action, however, would also shrink the output and, thus, the efficiency of the *first* stage.

Our objective, in this paper, is to add detail (by following the suggestion of Kern *et al*., 2012, p. 192) to the representation of the on-field production process in soccer. To this end, we present it for the first time as the *mixed* (serial and parallel) structure two-stage network system of Figure 1. According to this system, the *first* stage consists of two distinct sub-processes (offense and defense) that operate in *parallel*. These respectively use players' offensive and defensive actions as inputs to produce two different intermediate measures, namely, goals scored and prevention of goals conceded. These, in turn, are the inputs of the *second* stage (points' accumulation sub-process) that produces accumulated points.

Figure 1: On-field Production Process in Soccer

Mixed (serial and parallel) Structure Network



Note: This representation is seemingly similar to that used by Lewis *et al*. (2009) and by Lewis and Sexton (2004a) for presenting the on-field production process in baseball.

Furthermore, based on a two-stage network DEA model, we aim to estimate the offensive, defensive, and athletic efficiency of soccer teams during a league season. According to our proposed framework, these three different efficiency scores are obtained (for each soccer team under evaluation) by a *single* linear programming problem. This is based on the linear programming problem solved in Chen *et al*. (2010) that is able to: (*i*) define the efficiency of each sub-process on its specific production possibility sub-set; (*ii*) connect these sub-sets via intermediate measures to form the overall production possibility set; (*iii*) address any conflict (regarding the optimal level of intermediate measures) between the two stages of each production process under consideration. In this way, we provide (in a more direct way) more information (than the studies previously mentioned in this section) on the on-field performance of each soccer team under evaluation.

Despite that the deterministic DEA model ignores any measurement or data entry errors, we use it in this paper for two reasons. The *first* is that this model does not require the specification of a function for the production process under consideration. The *second* is that technology in soccer (i.e., sports equipment, action plans, training methods, eating plans, physical and mental preparation, employed tactics, etc.) is homogeneous and familiar to all managerial, coaching, and technical staffs (Boscá *et al*., 2009, p. 65).

We will demonstrate the usefulness of our proposed framework by employing aggregate-over-games statistics from the 2013-14 Greek premier soccer league. Consideration of an entire season (instead of a single game) ensures that the on-field performance of soccer teams is evaluated on an equal basis. This is because their estimated (offensive, defensive, and athletic) efficiencies are not influenced by chance and/or the quality of a particular opponent. The reason is that, during a league season, each team plays twice (on and away from its home ground) against each other team. As a result, stronger teams can be more easily differentiated from weaker teams since any effect of chance on the on-field performances declines as the number of games played increases.

2 Theoretical Framework

According to Espitia-Escuer and García-Cebrián (2004, p. 338), the production process in soccer consists of two distinct sub-processes. The *first* uses players' skills (i.e., their sporting talent, experience, physical condition, form, etc.) and the work of their coach (to conceive during trainings the best opponent-specific tactics, starting lineup, and substitutions) as inputs to generate the on-field (offensive and defensive) actions of soccer teams during games. These, in turn, are the inputs of the *second* sub-process that generates the athletic outputs (i.e., goals scored, prevention of goals conceded, and accumulated points) of soccer teams.

In this paper, by following earlier literature (e.g., Espitia-Escuer and García-Cebrián, 2004, 2006; Boscá *et al.*, 2009; García-Sánchez, 2007; Sala-Garrido *et al.*, 2009; Rossi *et al.*, 2018), we concentrate on the *second* of the above sub-processes, into which we delve deeper. This is because we seek to determine which soccer teams perform efficiently on the field and which soccer teams should perform better considering their effort during games. Specifically, we concentrate on the attempt of soccer teams to: (*i*) score goals with offensive actions; (*ii*) prevent their opponents from scoring with defensive actions; (*iii*) cumulate points by performing well in offense and/or defense.

It should be noted here that teams': (*i*) goals scored are positively influenced by the quantity and quality of their offensive actions; (*ii*) goals conceded are negatively influenced by the quantity and quality of their defensive actions; (*iii*) accumulated points are positively influenced by the quality of their performance in offense and defense. Therefore, soccer teams enhance their: (*i*) offensive efficiency by scoring more goals without missing lots of opportunities; (*ii*) defensive efficiency by conceding fewer goals without spending much energy to prevent their opponents from scoring (i.e., by possessing the ball longer than them); [3] (*iii*) athletic efficiency by *either* winning more games (even by a single goal) *or* (at least) not losing against stronger opponents.

In this section, the offensive, defensive, and athletic efficiency of soccer teams during a league season is estimated by a two-stage network DEA model. This assumes an output orientation because soccer teams prefer to expand their athletic outputs (i.e., goals scored, prevention of goals conceded, and accumulated points) rather than to shrink their on-field (offensive and defensive) actions (García-Sánchez, 2007, p. 31). In addition, this model assumes Variable Returns-to-Scale (VRS) because soccer teams might differ considerably in terms of: *first*, their available budget that might affect the quality and quantity of their players; *second*, whether or not they are among the usually strong league members; *third*, the experience, personality, and know-how of their managerial, coaching, and technical staff; *fourth*, their specific infrastructure, fan base, organization and ownership structure, history, and culture. Consider, for example, that some soccer teams generally prefer to primarily use inexperienced players with a view to advance and, then, trade them to wealthier soccer teams. Alternatively, some soccer teams generally prefer to primarily use players born in the area where their home ground is located in order to

[3] On the other hand, possessing the ball longer than opponents does not necessarily enhance the offensive efficiency of soccer teams.

promote (and gain by) the local society. Naturally, such decisions may affect the on-field performance of these soccer teams during a league season.

Given the above assumptions, the following linear programming problem is solved:

$$
\begin{array}{lll}
\underset{h_k^A, \tilde{y}_k^O, \tilde{y}_k^D, \lambda_j^k, \mu_j^k, \nu_j^k}{Max} & h_k^A = 1/A_k & \\
s.t. & \sum_{j=1}^{N} \lambda_j^k x_{ij}^O \leq x_{ik}^O, \quad i = 1, \dots, I; & \left.\right\} \text{Offense} \\
 & \sum_{j=1}^{N} \lambda_j^k y_j^O \geq \tilde{y}_k^O, & \\
 & \sum_{j=1}^{N} \mu_j^k x_{ij}^D \leq x_{ik}^D, \quad i = 1, \dots, I; & \left.\right\} \text{Defense} \\
 & \sum_{j=1}^{N} \mu_j^k y_j^D \leq \tilde{y}_k^D, & \\
 & \sum_{j=1}^{N} \nu_j^k y_j^O \leq \tilde{y}_k^O, & \\
 & \sum_{j=1}^{N} \nu_j^k y_j^D \geq \tilde{y}_k^D, & \left.\right\} \text{Points' accumulation sub-process} \\
 & \sum_{j=1}^{N} \nu_j^k Y_j \geq h_k^A Y_k, & \\
 & \sum_{j=1}^{N} \lambda_j^k = 1, & \\
 & \sum_{j=1}^{N} \mu_j^k = 1, & \\
 & \sum_{j=1}^{N} \nu_j^k = 1, & \\
 & \lambda_j^k \geq 0, \quad j = 1, \dots, k, \dots, N; & \\
 & \mu_j^k \geq 0, \quad j = 1, \dots, k, \dots, N; & \\
 & \nu_j^k \geq 0, \quad j = 1, \dots, k, \dots, N; &
\end{array} \tag{1}
$$

where $0 < A_k \leq 1$ refers to the athletic efficiency of the k^{th} soccer team under evaluation, $1 \leq h_k^A < \infty$ to its distance from the athletically efficient frontier evaluated by (1), x^O to the offensive inputs, x^D to the defensive inputs, y^O to the offensive intermediate measure, $\tilde{y}^O$ to the optimal level of y^O to be estimated, y^D to the *reverse* defensive intermediate measure, $\tilde{y}^D$ to the optimal level of y^D to be estimated, Y to the single output, λ_j^k, μ_j^k, and ν_j^k refer to the intensity variables, i is used to index both the offensive and defensive inputs, and j to index soccer teams. Consequently, if $A_k = h_k^A = 1$, then the k^{th} soccer team is athletically efficient since it could not have expanded its single output without increasing (decreasing) its offensive (*reverse* defensive) intermediate measure.

From the above, it follows that: *first*, (1) is seemingly similar to the linear programming problem solved in Chen *et al*. (2010). This is because (1): (*i*) provides each sub-process (namely, offense, defense, and points' accumulation) with its specific production possibility sub-set that is given by a particular group of intensity variables; (*ii*) connects these sub-sets via intermediate measures to form the overall production possibility set; (*iii*) benchmarks the performance of each sub-process of the k^{th} soccer team against the performance of the corresponding sub-process of all other soccer teams.

Second, (1) slightly differs from the linear programming problem solved in Chen *et al*. (2010). This is because (1): (*i*) represents a production process that consists of three (instead of two) sub-processes; (*ii*) makes use of a *reverse* intermediate measure, larger values for which mean a poorer defensive performance for soccer teams; to achieve this, we had to change (by following Lewis and Sexton, 2004b) the greater (less)-than-or-equal-to sign in the fourth (sixth) constraint of (1) for a less (greater)-than-or-equal-to sign; (*iii*) assumes VRS instead of Constant Returns-to-Scale (CRS). This is possible since the CRS two-stage network DEA model of Chen *et al*. (2010), on which (1) is based, is grounded on the two-stage network DEA model of Chen and Zhu (2004) that may assume *either* CRS *or* VRS (Chen *et al*., 2009, p. 602).

Third, (1) can also estimate the offensive efficiency of the k^{th} soccer team as $0 < O_k = 1/h_k^O = y_k^O/\tilde{y}_k^O \leq 1$, where $1 \leq h_k^O < \infty$ refers to its distance from the offensively efficient frontier evaluated by (1), and its defensive efficiency as $0 < D_k = \tilde{y}_k^D/y_k^D \leq 1$. Consequently, if $O_k = h_k^O = 1$ ($D_k = 1$), then the k^{th} soccer team is offensively (defensively) efficient since it could not have expanded (shrunk) its offensive (*reverse* defensive) intermediate measure without increasing its offensive (defensive) inputs. Note, however, that the two-stage network DEA model of Chen *et al*. (2010) may yield multiple optimal intermediate measures. Thus, (1) that is based on the above model may, in turn, estimate multiple offensive and defensive efficiencies for the k^{th} soccer team.

Before proceeding, it should be also noted that A_k, O_k, and D_k disregard the interdependency of the athletic output levels. Specifically, athletic efficiencies produced by our proposed framework disregard the fact that total points, which can be awarded to soccer teams during a league season, are (according to Bi *et al*., 2015) bounded from above. Similarly, offensive and defensive efficiencies produced by our proposed framework disregard the fact that total goals scored (conceded) during a league season cannot differ from total goals conceded (scored). Consequently, A_k, O_k, and D_k may be biased downward and, for this reason, we have to adjust them by following Bouzidis and Karagiannis (2019). For robustness purposes, this adjustment is achieved by means of both the *proportional* output reduction strategy of Lins *et al.* (2003) and the *equal* output reduction strategy of Collier *et al.* (2011).

3 Data Description

The relevant data used for our on-field performance evaluation are aggregate-over-games statistics (presented in Table 1) from the 2013-14 Greek premier soccer league. The reasons for this data choice are the following: *first*, game statistics can accurately describe the (offensive and defensive) skills as well as the playing style and structure of soccer teams (Boscá *et al*., 2009, p. 66); *second*, there are no direct (qualitative or quantitative) data regarding the (offensive and defensive) skills, playing style, and structure of soccer teams; *third*, the available budget of each soccer team does not necessarily indicate its (offensive and defensive) skills (Boscá *et al.*, 2009, fn. 8); *fourth*, the efficiencies of soccer teams under evaluation are almost the same irrespective of whether financial or game statistics are used as inputs in the DEA model (Zambom-Ferraresi *et al.*, 2019).

Table 1: Aggregate-over-games Statistics, 2013-14 Greek Premier Soccer League

Final Ranking	Offense				Defense				Points' accumulation sub-process
Teams	x_{1k}^O	x_{2k}^O	x_{3k}^O	y_k^O	x_{1k}^D	x_{2k}^D	x_{3k}^D	y_k^D	Y_k
1. OSFP	496	982	79	88	68	308	566	19	86
2. PAOK	426	996	68	68	108	395	797	37	69
3. ATROMITOS	386	786	66	54	139	571	731	25	66
4. PAO	423	854	57	57	134	493	699	28	66
5. ASTERAS TRIPOLIS	425	880	50	46	150	500	628	35	58
6. OFI	364	871	42	30	109	506	745	39	44
7. ERGOTELIS	334	688	39	39	163	565	752	40	44
8. LEVADIAKOS	268	590	37	42	158	576	658	61	42
9. PANAITOLIKOS	390	712	30	32	133	519	596	33	42
10. PANTHRAKIKOS	298	659	27	39	168	624	848	52	41
11. PAS GIANNINA	299	627	40	34	141	499	733	43	41
12. KALLONI	297	461	29	31	172	597	647	62	39
13. PANIONIOS	279	619	27	33	152	588	613	42	39
14. PLATANIAS	344	575	35	39	123	480	581	48	38
15. VERIA	348	688	29	31	119	463	817	51	38
16. SKODA XANTHI	280	696	40	44	120	479	789	54	38
17. APOLLON ATHENS	297	876	34	43	187	537	598	54	36
18. ARIS	321	750	25	26	129	513	825	53	22
Max	496	996	79	88	187	624	848	62	86
Min	268	461	25	26	68	308	566	19	22
Average	348.6	739.4	41.9	43.1	137.4	511.8	701.3	43.1	47.2
Standard Deviation	64.1	146.4	15.9	15.5	28.2	75.6	92.7	12.2	15.5

Notes: 1) x_1^O refers to the total *Shots* and *Headers*, x_2^O to the total *Crosses*, and x_3^O to the total *Assists*,

2) x_1^D refers to the *Saves*, x_2^D to the *Clearances*, and x_3^D to the *Steals*,

3) y^O refers to the *Goals Scored*, y^D to the *Goals Conceded*, and Y to the *Points* cumulated.

Source: *Galanis Sports Data* (www.galanissportsdata.com)

In particular, the aggregate-over-games statistics (from the 2013-14 Greek premier soccer league) used for our on-field performance evaluation are teams': (*i*) on-field actions, such as total *Shots* and *Headers* (outside and inside penalty area, off and on goal), total *Crosses* (including those that did not reach a teammate), and total *Assists* (including those that failed to turn into a goal) that are considered as offensive inputs; (*ii*) on-field actions, such as *Saves*, *Clearances* (both coming from goalkeeper and other players), and *Steals* (which are required for a soccer team to regain the ball) that are considered as defensive inputs; (*iii*) *Goals Scored* that are considered as the offensive intermediate measure; (*iv*) *Goals Conceded* that are considered as the *reverse* defensive intermediate measure; (*v*) accumulated *Points* that are considered as the single output.

4 Empirical Results

The offensive, defensive, and athletic efficiency of soccer teams in respectively transforming (during the league season under consideration) their: (*i*) offensive actions into goals scored, (*ii*) defensive actions into prevention of goals conceded, and (*iii*) goals (scored and conceded) into accumulated points is presented in Tables 2 and 3. From these tables, it can be seen that: *first*, soccer teams under evaluation could potentially have scored approximately 8-16% more goals on average keeping constant their offensive actions. This could have been accomplished, for example, by forcing their opponents to commit more errors. On the other hand, they could potentially have conceded approximately 9-51.5% fewer goals on average keeping constant their defensive actions. This could have been accomplished, for example, by resisting their opponents' pressure in a better way. Thus, soccer teams under evaluation generally scored goals more efficiently than prevented their opponents from scoring.

Second, soccer teams under evaluation could potentially have cumulated approximately 12-28% more points on average keeping constant their goals (scored and conceded). To understand how this is possible, consider that a soccer team would be actually benefited by exchanging, for example, two of its (1-1) ties for a (2-0) victory and a (0-2) loss. *Third*, had it been athletically efficient, Apollon Athens (#17) could have been placed tenth (according to the results provided by (1) and the *equal* output reduction strategy of Collier *et al*., 2011). In this case, this soccer team would have avoided relegation to the secondary league, in which Kalloni (#12) would have ended up. *Fourth*, OSFP (#1) and Aris (#18) proved to be athletically efficient irrespective of whether they were evaluated by (1), the *proportional* output reduction strategy of Lins *et al.* (2003) or the *equal* output reduction strategy of Collier *et al*. (2011). This is because OSFP (#1) cumulated the most points in the league season under consideration and Aris (#18) some points despite scoring the fewest goals.

Table 2: Offensive Inefficiency/Defensive Efficiency, 2013-14 Greek Premier Soccer League

Final Ranking	Offense				Defense			
Teams	$\tilde{y}_k^O$	h_k^O	$\hat{h}_k^O$	$\check{h}_k^O$	$\tilde{y}_k^D$	D_k	$\hat{D}_k$	$\check{D}_k$
1. OSFP	88	1	1	1	19	1	1	1
2. PAOK	74	1.087	1.079	1	19	0.514	0.537	1
3. ATROMITOS	65	1.206	1.189	1.096	19	0.760	0.768	1
4. PAO	67	1.181	1.165	1.076	19	0.679	0.690	1
5. ASTERAS TRIPOLIS	61	1.319	1.295	1.190	19	0.543	0.563	1
6. OFI	53	1.770	1.719	1.571	19	0.487	0.513	1
7. ERGOTELIS	48	1.235	1.220	1.082	19	0.475	0.502	1
8. LEVADIAKOS	42	1	1	1	19	0.311	0.366	0.707
9. PANAITOLIKOS	42	1.307	1.291	1.121	19	0.576	0.594	1
10. PANTHRAKIKOS	39	1	1	1	19	0.365	0.408	0.829
11. PAS GIANNINA	46	1.346	1.326	1.170	19	0.442	0.473	1
12. KALLONI	31	1	1	1	19	0.306	0.362	0.695
13. PANIONIOS	33	1	1	1	19	0.452	0.482	1
14. PLATANIAS	40	1.032	1.030	1	19	0.396	0.433	0.898
15. VERIA	41	1.319	1.303	1.127	19	0.373	0.414	0.845
16. SKODA XANTHI	44	1.009	1.009	1	19	0.352	0.397	0.798
17. APOLLON ATHENS	43	1	1	1	19	0.352	0.397	0.798
18. ARIS	26	1	1	1	19	0.358	0.402	0.813
Max	88	1.770	1.719	1.571	19	1	1	1
Min	26	1	1	1	19	0.306	0.362	0.695
Average	49.1	1.156	1.146	1.080	19	0.486	0.517	0.910
Standard Deviation	16.1	0.204	0.191	0.139	-	0.179	0.164	0.112

Notes: 1) $\tilde{y}_k^O$, $\tilde{y}_k^D$, h_k^O, and D_k were produced by (1),

2) $\hat{h}_k^O$ and $\hat{D}_k$ were provided by the *proportional* output reduction strategy of Lins *et al*. (2003),

3) $\check{h}_k^O$ and $\check{D}_k$ were provided by the *equal* output reduction strategy of Collier *et al*. (2011).

Table 3: Athletic Inefficiency, 2013-14 Greek Premier Soccer League

Final Ranking			Alternate Ranking			Alternate Ranking			Alternate Ranking
Teams	h_k^A	$h_k^A Y_k$	Teams	$\hat{h}_k^A$	$\hat{h}_k^A Y_k$	Teams	$\check{h}_k^A$	$\check{h}_k^A Y_k$	Teams
1. OSFP	1	86	1	1	85	1	1	78	1
2. PAOK	1.126	78	2	1.115	69	2	1.011	70	2
3. ATROMITOS	1.099	73	4	1.092	66	4	1	65	4
4. PAO	1.118	74	3	1.109	66	3	1	66	3
5. ASTERAS TRIPOLIS	1.206	70	5	1.190	58	5	1.069	62	5
6. OFI	1.482	65	6	1.449	45	6	1.302	57	6
7. ERGOTELIS	1.378	61	7	1.353	44	7	1.198	53	7
8. LEVADIAKOS	1.310	55	11	1.291	42	8	1.121	47	11
9. PANAITOLIKOS	1.306	55	16	1.288	42	9	1.117	47	16
10. PANTHRAKIKOS	1.274	52	17	1.259	41	11	1.081	44	17
11. PAS GIANNINA	1.426	58	8	1.399	41	10	1.233	51	8
12. KALLONI	1.152	45	9	1.144	39	13	1	37	9
13. PANIONIOS	1.199	47	15	1.189	39	12	1	39	15
14. PLATANIAS	1.405	53	14	1.382	38	16	1.197	45	14
15. VERIA	1.420	54	10	1.396	38	15	1.212	46	10
16. SKODA XANTHI	1.506	57	13	1.475	39	14	1.297	49	13
17. APOLLON ATHENS	1.553	56	12	1.520	37	17	1.333	48	12
18. ARIS	1	22	18	1	22	18	1	14	18
Max	1.553	86		1.520	85		1.333	78	
Min	1	22		1	22		1	14	
Average	1.276	58.9		1.258	47.2		1.121	51	
Standard Deviation	0.172	14.4		0.161	15.2		0.120	14.4	

Notes: 1) h_k^A was produced by (1),

2) $\hat{h}_k^A$ was provided by the *proportional* output reduction strategy of Lins *et al.* (2003),

3) $\check{h}_k^A$ was provided by the *equal* output reduction strategy of Collier *et al.* (2011),

4) The top five teams in the final ranking were entitled to participate in the next season's UEFA tournaments,

5) The bottom two teams in the final ranking were relegated to the secondary league.

Fifth, the top five soccer teams in the final ranking cumulated points more efficiently than their direct competitors during the league season under consideration. This is why these soccer teams were entitled to participate in the next season's tournaments organized by the *Union of European Football Associations* (UEFA). For example, PAOK (#2) (Asteras Tripolis #5) cumulated more points than Panaitolikos (#9) despite that these soccer teams performed rather similarly in defense (offense).

Sixth, some soccer teams followed similar (offensive and defensive) strategies during the league season under consideration. For example, Atromitos (#3), Asteras Tripolis (#5), and Panaitolikos (#9) placed less (more) emphasis on offense (defense) than other soccer teams. This is evident from the fact that their offensive (defensive) efficiencies are smaller (greater) than the estimated average efficiency in offense (defense). On the contrary, Levadiakos (#8), Panthrakikos (#10), Kalloni (#12), Platanias (#14), Skoda Xanthi (#16), Apollon Athens (#17), and Aris (#18) placed more (less) emphasis on offense (defense) than other soccer teams. This is evident from the fact that their offensive (defensive) efficiencies are greater (smaller) than the estimated average efficiency in offense (defense). From the above results, it becomes clear that the more effective (offensive and defensive) strategies, which yielded the more points, were followed by the former group of soccer teams. In addition, it follows that, to enhance their athletic efficiency, soccer teams of the former (latter) group should have performed more efficiently in offense (defense).

Seventh, Levadiakos (#8) and Panaitolikos (#9) performed nearly identically in the points' accumulation sub-process during the league season under consideration. However, these soccer teams presented some performance differences in offense and defense. Thus, their managers had to deal with different efficiency enhancing game aspects for improving their on-field performance in the next season. For this purpose, it could have been helpful for them to consider the playing style, structure, and players' characteristics of their team's benchmarks provided (for each sub-process) by (1). For instance, as Table 4 reveals, Panaitolikos (#9) may have improved its on-field performance in the next season by following the example of: (*i*) OSFP (#1) and Panthrakikos (#10) in offense; (*ii*) OSFP (#1) in defense; (*iii*) Atromitos (#3) and OFI (#6) in the points' accumulation sub-process. It should be noted here that OSFP (#1) is the most popular benchmark in offense as well as defense and Atromitos (#3) in the points' accumulation sub-process.

Table 4: Benchmarks, 2013-14 Greek Premier Soccer League

Teams	Offense	Defense	Points' accumulation sub-process
1. OSFP	1	1	1
2. PAOK	1, 8	1	1, 3
3. ATROMITOS	1, 8, 12	1	1, 3
4. PAO	1, 10	1	1, 3
5. ASTERAS TRIPOLIS	1, 10	1	1, 3
6. OFI	1, 10	1	3, 6
7. ERGOTELIS	1, 8, 10, 12	1	3, 6
8. LEVADIAKOS	8	1	3, 6
9. PANAITOLIKOS	1, 10	1	3, 6
10. PANTHRAKIKOS	10	1	3, 6
11. PAS GIANNINA	1, 8, 10, 12	1	3, 6
12. KALLONI	12	1	3, 6
13. PANIONIOS	13	1	3, 6
14. PLATANIAS	1, 10, 12	1	3, 6
15. VERIA	1, 10	1	3, 6
16. SKODA XANTHI	1, 8	1	3, 6
17. APOLLON ATHENS	17	1	3, 6
18. ARIS	18	1	18
Times served as a benchmark for another team			
Teams	Offense	Defense	Points' accumulation sub-process
1. OSFP	11	17	4
3. ATROMITOS	0	0	15
6. OFI	0	0	11
8. LEVADIAKOS	5	0	0
10. PANTHRAKIKOS	8	0	0
12. KALLONI	4	0	0

Eighth, according to the results provided by our proposed framework, the *proportional* output reduction strategy of Lins *et al.* (2003), and the *equal* output reduction strategy of Collier *et al.* (2011), only OSFP (#1) performed efficiently in all three sub-processes. In addition, only the specific soccer team was athletically efficient among the high-ranked soccer teams of the league season under consideration. These facts might explain the rather low rank correlations (presented in Table 5) among offensive, defensive, and athletic efficiency and between athletic efficiency and accumulated points.

From Table 5, it can also be seen that: (*i*) the rank correlations between defensive and athletic efficiency are generally greater than the rank correlations between offensive and athletic efficiency; (*ii*) the rank correlations between defensive efficiency and accumulated points are positive and quite strong ranging from 66 to 77 percent; (*iii*) the rank correlations between offensive efficiency and accumulated points are negative and quite weak ranging from -29 to -27 percent. From the above results, it follows that it is defensive efficiency, which mostly led soccer teams under evaluation to athletic efficiency and to a better final league ranking.

Table 5: Rank Correlation Coefficients, 2013-14 Greek Premier Soccer League

	O	D	A
Y	-0.28	0.77	0.46
O		-0.46	0.33
D			0.44
	$\hat{O}$	$\hat{D}$	$\hat{A}$
Y	-0.27	0.77	0.46
$\hat{O}$		-0.45	0.34
$\hat{D}$			0.44
	$\check{O}$	$\check{D}$	$\check{A}$
Y	-0.29	0.66	0.37
$\check{O}$		-0.56	0.29
$\check{D}$			0.19

Notes: 1) Y refers to the *Points* that teams cumulated,

2) O refers to the offensive efficiencies, D to the defensive efficiencies, and A to the athletic efficiencies of teams produced by our proposed framework,

3) $\hat{O}$ refers to the offensive efficiencies, $\hat{D}$ to the defensive efficiencies, and $\hat{A}$ to the athletic efficiencies of teams provided by the *proportional* output reduction strategy of Lins *et al*. (2003),

4) $\check{O}$ refers to the offensive efficiencies, $\check{D}$ to the defensive efficiencies, and $\check{A}$ to the athletic efficiencies of teams provided by the *equal* output reduction strategy of Collier *et al*. (2011).

5 Discussion

As mentioned in the previous section, the points' accumulation sub-processes of Atromitos (#3) and OFI (#6) serve as benchmarks for the corresponding sub-process of Panaitolikos (#9). However, according to the results produced by our proposed framework, Atromitos (#3) and OFI (#6) were athletically inefficient in the league season under consideration. The reason behind this paradox is that athletic efficiency is defined on the production possibility sub-set of the points' accumulation sub-process. Therefore, if this sub-process was *independently* evaluated by a single-stage DEA model, then the above two soccer teams would be athletically efficient. [4] This is why they are included (despite being athletically inefficient) among the benchmarks provided (for the points' accumulation sub-process) by our proposed framework. [5] From the above, it follows that the two-stage network DEA model used in this paper discriminates soccer teams under evaluation better than a two-stage network DEA model (such as, for example, the one employed in Kern *et al*., 2012), according to which each sub-process would be *independently* evaluated.

6 Concluding Remarks

In this paper, we presented for the first time the on-field production process of soccer teams as a *mixed* (serial and parallel) structure two-stage network of three sub-processes (namely, offense, defense, and points' accumulation). Furthermore, based on a two-stage network DEA model, we estimated the offensive, defensive, and athletic efficiency of soccer teams during a league season. According to our proposed framework, these three different efficiency scores were provided (for each soccer team under evaluation) by a *single* linear programming problem. For this purpose,

---

[4] For brevity, the results obtained from the *independent* evaluation of the points' accumulation sub-process under consideration are not presented here but are available upon request.

[5] A similar situation is observed in the results provided by the Sexton and Lewis (2003) two-stage network DEA model that was employed for the performance evaluation of baseball teams. For another similar situation, see Lewis and Sexton (2004a, p. 1384) that also evaluated the performance of baseball teams.

aggregate-over-games statistics from the 2013-14 Greek premier soccer league were used.

According to our findings: (*i*) soccer teams under evaluation generally scored goals more efficiently than prevented their opponents from scoring; (*ii*) it is defensive efficiency that mostly led soccer teams under evaluation to athletic efficiency; (*iii*) to improve their final league ranking, soccer teams under evaluation should have generally performed more efficiently in defense; (*iv*) only the champion of the league season under consideration performed efficiently in all three sub-processes and was athletically efficient among the high-ranked soccer teams; (*v*) the champion (second runner-up) of (in) the league season under consideration is the most popular benchmark in both offense and defense (the points' accumulation sub-process); (*vi*) the top five soccer teams in the final ranking cumulated points more efficiently than their direct competitors during the league season under consideration; (*vii*) some soccer teams followed similar (offensive and defensive) strategies during the league season under consideration.

As a final remark, it should be noted that the empirical application of this paper could be extended in the future to consider more seasons. In this way, a better picture for the soccer league under consideration could be produced. Moreover, the theoretical framework of this paper could be extended in the future to also consider aspects of the off-field performance of soccer teams, such as their efficiency in revenue generation or their impact on fans, spectators, etc. Note, finally, that our proposed framework could also be used in the future for the performance evaluation in other sports, such as American football, baseball, basketball, (ice) hockey, handball, volleyball, water polo, tennis, etc.

Conflicts of Interest: The author declares that there are no conflicts of interest.

References

Bi, G., W. Song, and M. Song (2015), "An output-constrained DEA model as applied to efficiency analysis of football clubs", *Journal of Intelligent & Fuzzy Systems*, 28(2), 561-570.

Boscá, J. E., V. Liern, A. Martínez, and R. Sala (2009), "Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football", *Omega*, 37(1), 63-78.

Bouzidis, Th. and G. Karagiannis (2019), "Output interdependency and efficiency in soccer", submitted to the *Journal of Sports Economics*.

Carmichael, F., D. Thomas, and R. Ward (2001), "Production and efficiency in association football", *Journal of Sports Economics*, 2(3), 228-243.

Chen, Y., W. D. Cook, and J. Zhu (2010), "Deriving the DEA frontier for two-stage processes", *European Journal of Operational Research*, 202(1), 138-142.

Chen, Y., L. Liang, and J. Zhu (2009), "Equivalence in two-stage DEA approaches", *European Journal of Operational Research*, 193(2), 600-604.

Chen, Y. and J. Zhu (2004), "Measuring information technology's indirect impact on firm performance", *Information Technology and Management*, 5(1-2), 9-22.

Collier, T., A. L. Johnson, and J. Ruggiero (2011), "Measuring technical efficiency in sports", *Journal of Sports Economics*, 12(6), 579-598.

Dawson, P., S. Dobson, and B. Gerrard (2000), "Estimating coaching efficiency in professional team sports: Evidence from English association football", *Scottish Journal of Political Economy*, 47(4), 399-421.

Espitia-Escuer, M. and L. I. García-Cebrián (2004), "Measuring the efficiency of Spanish first-division soccer teams", *Journal of Sports Economics*, 5(4), 329-346.

Espitia-Escuer, M. and L. I. García-Cebrián (2006), "Performance in sports teams: Results and potential in the professional soccer league in Spain", *Management Decision*, 44(8), 1020-1030.

García-Sánchez, I. M. (2007), "Efficiency and effectiveness of Spanish football teams: A three-stage-DEA approach", *Central European Journal of Operations Research*, 15(1), 21-45.

Halkos, G. E., N. G. Tzeremes, and S. A. Kourtzidis (2014), “A unified classification of two-stage DEA models”, *Surveys in Operations Research and Management Science*, 19(1), 1-16.

Kern, A., M. Schwarzmann, and A. Wiedenegger (2012), “Measuring the efficiency of English premier league football: A two-stage data envelopment analysis approach”, *Sport, Business and Management: An International Journal*, 2(3), 177-195.

Lewis, H. F., K. A. Lock, and T. R. Sexton (2009), “Organizational capability, efficiency, and effectiveness in major league baseball: 1901-2002”, *European Journal of Operational Research*, 197(2), 731-740.

Lewis, H. F. and T. R. Sexton (2004a), “Network DEA: Efficiency analysis of organizations with complex internal structure”, *Computers & Operations Research*, 31(9), 1365-1410.

Lewis, H. F. and T. R. Sexton (2004b), “Data envelopment analysis with reverse inputs and outputs”, *Journal of Productivity Analysis*, 21(2), 113-132.

Lins, M. P. E., E. G. Gomes, J. C. C. B. Soares de Mello, and A. J. R. Soares de Mello (2003), “Olympic ranking based on a zero sum gains DEA model”, *European Journal of Operational Research*, 148(2), 312-322.

Rossi, G., D. Goossens, G. L. Di Tanna, and F. Addesa (2018), “Football team performance efficiency and effectiveness in a corruptive context: The Calciopoli case”, *European Sport Management Quarterly*, DOI: https://doi.org/10.1080/16184742.2018.1553056.

Sala-Garrido, R., V. Liern Carrión, A. Martinez Esteve, and J. E. Boscá (2009), “Analysis and evolution of efficiency in the Spanish soccer league (2000/01-2007/08)”, *Journal of Quantitative Analysis in Sports*, 5(1), Article 3, DOI: https://doi.org/10.2202/1559-0410.1143.

Sexton, T. R. and H. F. Lewis (2003), “Two-stage DEA: An application to major league baseball”, *Journal of Productivity Analysis*, 19(2-3), 227-249.

Zambom-Ferraresi, F., B. Iráizoz, and F. Lera-López (2019), “Are football managers as efficient as coaches? Performance analysis with *ex ante* and *ex post* inputs in the premier league”, *Applied Economics*, 51(3), 303-314.