Motivation of basketball players: a random-effects logit model for the probability of winning

Silvia Bacci, Tijan Juraj Cvetković

1. Introduction

Professional sports are getting more competitive as athletes strive to improve their sports performance and sport organizations employ various coaches in order to help athletes in achieving this aim. In an environment where athletes are physically dominant and have high skill mastery, psychological factors can make a difference to prevail over other athletes. For this reason, sport psychology (Perry, 2015) plays an important role in preparing an athlete from a mental perspective, just as a coach prepares from a physical perspective. In sports, motivation is a key factor of success, hence sport organisations, decision makers, sport psychologists, and players themselves must address it constantly in order to keep it and perform at the highest levels.

Psychology offers various theories and theoretical models to explain the motivational process, its benefits, and how to create a motivational climate. In this contribution we considered McClelland's Need achievement theory (McClelland, 1961) and the Nicholls' Achievement goal theory (Nicholls, 1984). These theories have something in common: goal setting, the incentive value of success, and the probability of success. Estimating the probability of success is difficult, subjective and, often, inaccurate. An error in any step of the motivational process may lead to a mistake in the role assignment, performance, and goal setting.

This paper aims at estimating the probability of success and, consequently, at making clear the motivational process such that a team or an athlete can be easily assigned to a certain role, can enhance their performance, and can set a goal as in deciding what segments of a sport must be improved. The estimation of the success probability relies on detecting the variables that affect the probability of winning in a statistically significant way. As these variables differ according to various sports, in this paper we focus on basketball, in particular the U.S. National Basketball Association (NBA). The study is based on the analysis of the traditional box scores of the regular season games played in the seasons 2016-17, 2017-18, 2018-19, and 2020-21. Because of the hierarchical structure of data at issue, with multiple observations for each team, a random intercept logit model was formulated and estimated.

The remaining part of the paper is organized as follows. The theoretical background concerning the motivational process from a psychological point of view is illustrated in Section 2, data are described in Section 3, and the main results related with the random intercept logit model are shown in Section 4. Finally, some remarks conclude the paper.

2. Motivation

Need achievement theory (McClelland, 1961) is a theory that explains what a person goes through when he/she decides to adopt a certain behaviour. McClelland considered the one's implemented behaviour as the result of a combination between personality traits and situational, resultant, and emotional factors, as illustrated in Figure 1. In detail, there are two main personality traits driving the behaviour along alternative paths: "need to achieve" and "need to avoid failure". The need to achieve is characterized by a drive to successfully compete with the standards of excellence, whereas the need to avoid failure distinguishes for a negative motivation oriented to avoid failure and criticism. These factors link with situational factors, including the probability of success and the incentive value of success. A person weights his/her probability to success and what he/she stands to gain from it. This interaction is crucial as its

Silvia Bacci, University of Florence, Italy, silvia.bacci@unifi.it, 0000-0001-8097-3870
Tijan Juraj Cvetković, University of Florence, Italy, tijan.cvetkovic@stud.unifi.it, 0000-0002-6971-4023

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Silvia Bacci, Tijan Juraj Cvetković, *Motivation of basketball players: a random-effects logit model for the probability of winning*, pp. 81-85, © 2021 Author(s), CC BY 4.0 International, DOI 10.36253/978-88-5518-461-8.16, in Bruno Bertaccini, Luigi Fabbris, Alessandra Petrucci (edited by), *ASA 2021 Statistics and Information Systems for Policy Evaluation. Book of short papers of the on-site conference*, © 2021 Author(s), content CC BY 4.0 International, metadata CC0 1.0 Universal, published by Firenze University Press (www.fupress.com), ISSN 2704-5846 (online), ISBN 978-88-5518-461-8 (PDF), DOI 10.36253/978-88-5518-461-8

resultant leads to either approach success or avoid failure. Moreover, emotional factors influence whether we focus on pride or shame. As result of these two main paths, the implemented behaviour consists, respectively, in seeking a challenge and enhanced performance or avoiding challenges, less effort and risk.

Achievement goal theory (Nicholls, 1984) is an orthogonal theory based on the persons approach to a task. Nicholls emphasizes the journey to the goal rather than the results of the goal itself. In the achievement goal theory a person focuses on skill mastery, a self-comparative perspective in which beating a previous personal result is success, or on the ego in which success is determined by comparison with others.

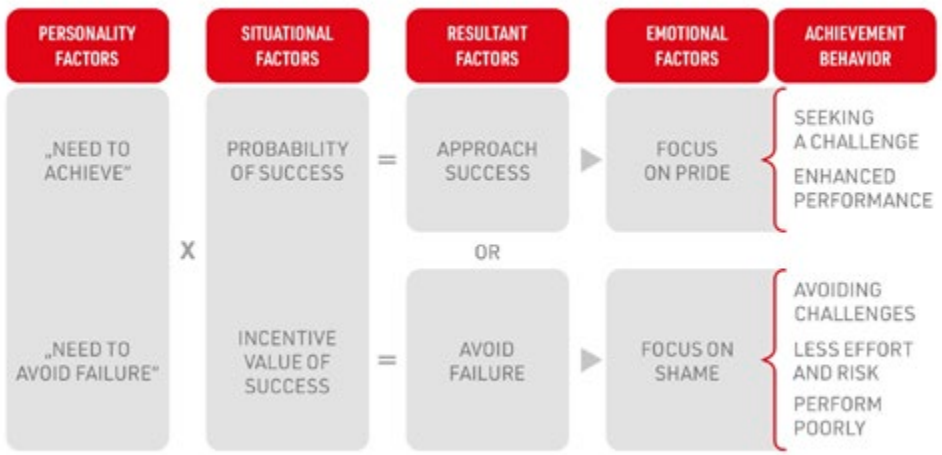

Figure 1 - Illustration of McClelland's (1961) Need Achievement Theory (our elaboration).

3. Data

The data used in the model was collected on websites NBA.com and BasketballReference.com. The dataset was constructed using the traditional box score statistics of the NBA for each game played in the seasons 2016-17, 2017-18, 2018-19 and 2020-21. The traditional box score contains information about: opposing teams, final outcome of the match (in terms of winning and losing), duration of the game (in minutes), total points scored, field goals made, field goals attempted, field goal shooting percentage, 3 point field goals made, 3 point field goals attempted, 3 point field goal percentage, number of free throws made, number of free throws attempted, percentage of free throws, offensive rebounds, defensive rebounds, total rebounds, assists, number of stolen balls, number of lost balls, number of blocks, personal fouls.

Data was arranged with a record per game (i.e., two teams in a single row) and variables were rescaled with extreme values omitted to avoid singularities. Due to changes in the leagues structure, omitted games include games played after the implementation of the Play-in tournament in the 2020-21 season and the 2019-2020 season in order to avoid variance due to circumstances. The resulting dataset is composed of 4,770 games played by 30 teams.

4. Random intercept logit model

To properly address the multilevel data structure, consisting in multiple observations per team, the probability of winning is modelled through a random-intercept logit model, where teams are the upper-level units and games are the lower-level units. The dependent variable of the model is a binary one equal to 1 if the team won the game and 0 otherwise.

As concerns the independent variables, in order to estimate the probability of success for a team we considered the *differences of the per game statistics*, which are an average of the variables based on previous games, *and the opponent per game statistics*, which are averages of other teams performances against the team of interest.

For the sake of clarity, we illustrate how to build the independent variables and to estimate the probability of success of a game played between the Utah Jazz (Team A) and the Sacramento Kings (Team B). Let us consider the following variables:

- FGM: field goals made
- FGA: field goals attempted
- 3PM: 3-point field goals made
- 3PA: 3-point field goals attempted
- FT%: free throw percentage
- DREB: defensive rebounds
- REB: total rebounds
- AST: assists
- STL: steals
- PF: personal fouls

The values of per game statistics representing the offensive performance of Team A and Team B, respectively, are displayed in Table 1, whereas the opponent per game statistics, representing the defensive performance of Team A and Team B, are reported in Table 2.

Table 1 - Per game statistics

Variable	Team A	Team B
FGM	41.3	42.6
FGA	88.1	88.6
3PM	16.7	12.2
3PA	43.1	33.4
FT%	79.3	74.3
DREB	37.6	32
AST	23.6	25.5
STL	6.5	7.5
TOV	14.2	13.4
PF	18.6	19.4

Table 2 - Opponent per game statistics

Variable	Team A	Team B
FGM	40.9	43.6
FGA	91.4	89.4
3PM	10.9	12.4
3PA	31.8	32.6
FT%	76.8	78.7
DREB	32.8	34.6
AST	22.3	25.3
STL	7.7	7.6
TOV	11.5	13.7
PF	19	18.7

Cross averaging Team A per game statistics and Team B opponent per game statistics, and vice-versa, results in a set of variables that take into account both offense and defense of the teams of interest (Table 3). This set of variables considers how the Utah Jazz (Team A) attack will vary against the Sacramento Kings (Team B) defence.

Table 3 - Cross averages of Team A per game and Team B opponent per game statistics, and vice-versa

Variable	Team A	Team B	Difference
FGM	42.4	41.8	0.7
FGA	88.8	90	-1.2
3PM	14.6	11.5	3
3PA	37.8	32.6	5.3
FT%	79	75.6	3.4
DREB	36.1	32.4	3.7
AST	24.4	23.9	0.5
STL	7.1	7.6	-0.5
TOV	13.9	12.4	1.5
PF	18.7	19.2	-0.6

Differences in the last column of Table 3 are then used as independent variables in the random-intercept logit model.

Estimates of the fixed effects of the model are shown in Table 4 (letter "d" before the variable names stays for "difference") and the related correlation matrix in Table 5. The selected model fits data in a very satisfactorily way, being the conditional and marginal pseudo-R^2 equal to 95.5% (Nakagawa and Schielzeth, 2013).

Table 4 – Random-intercept logit model: estimates of fixed effects (significance level 5%)

Variable	**Estimate**	**Std. Error**	**Z-value**	**p-value**
(Intr)	0.000	0.063	0.002	0.999
dFGM	0.745	0.026	27.876	<0.0001
dFGA	-0.126	0.009	-14.474	<0.0001
d3PM	0.624	0.023	26.572	<0.0001
d3PA	-0.053	0.009	-6.114	<0.0001
dFT%	7.538	0.442	17.034	<0.0001
dDREB	0.391	0.018	21.930	<0.0001
dAST	-0.022	0.011	-2.110	0.035
dSTL	0.145	0.020	7.390	<0.0001
dTOV	-0.421	0.023	-17.9	<0.0001
dPF	-0.444	0.016	-26.724	<0.0001

We note that, in addition to variables displayed in Table 2, we investigated other possible determinants that, however, did not result statistically significant. In particular, no significant effect resulted for the difference in field goal percentage, 3-point field goal percentage, free throws made, free throws attempted, offensive rebounds, rebounds, blocks, and for the game season (dummies were added to the model for seasons 2016-2017, 2017-2018, 2018-2019 versus season 2020-2021).

Table 5 – Correlation matrix for significant independent variables.

	dFGM	dFGA	d3PM	d3PA	dFT%	dDREB	dAST	dSTL	dTOV
dFGM									
dFGA	-0.488								
d3PM	0.449	-0.077							
d3PA	0.071	-0.238	-0.587						
dFT%	0.503	-0.173	0.316	-0.086					
dDREB	0.074	-0.006	0.396	-0.038	0.000				
dAST	-0.328	0.102	-0.158	-0.082	-0.060	-0.021			
dSTL	0.122	-0.041	0.142	-0.050	0.139	0.020	0.017		
dTOV	-0.129	0.285	-0.344	0.056	-0.016	-0.717	0.010	0.430	
dPF	-0.696	0.137	-0.512	0.082	-0.439	-0.149	0.110	-0.180	0.050

5. Conclusions

By analysing the traditional box scores of regular season games of the National Basketball Association (NBA) played in the seasons 2016-17, 2017-18, 2018-19 and 2020-21 we found several variables influencing the probability of winning, such as field goals made, field goals attempted, 3-point field goals made, 3-point field goals attempted, free throw percentage, number of defensive rebounds, number of assists, number of steals, number of turnovers and number of personal fouls.

Knowing the effect of these variables on the probability of winning helps a sport organization to improve the motivation of its athletes and to adopt a team-oriented approach to games. By objectively defining the probability of success and knowing what aspects of the game to focus on, the team decision makers can make changes accordingly. For instance, the roles assignment within a team can be improved assembling a team of players that are individually specialized in the significant categories and can consistently obtain values favouring the probability of winning. Moreover, goal setting such as keeping the opposing team under a certain number of made field goals or any other category, or rather prioritizing certain categories to maximise the probability of winning, can be easily identified benefiting both the team as a whole, by improving its chances, and the single athletes, by making him/her more proficient in a single task.

For the future research, we intend to investigate the role of an additional independent variable aimed at considering how injuries of key players affect the probability of winning. The role of team key players will be determined by analyzing their win share statistics. In particular, it will be interesting to assess the effect on the probability of winning of the number of injured or missing key players (none, one, more than one) in a game.

References

"NBA Win Shares - Basketball-Reference.com". Basketball-Reference.com. Retrieved 2016-03-25.

McClelland, D. C. (1961). *The achieving society*. Van Nostrand, Princeton, NJ.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, pp. 133-142.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review,* **91**(3), pp. 328-346.

Perry, J. (2015). *Sport psychology: A complete introduction.* Hodder & Stoughton, UK, pp. 105-123.