WHO RULES IN TIMES OF THE VIDEO ASSISTANT REFEREE?

DECISION MAKING IN DUTCH FOOTBALL

Harry Garretsen (University of Groningen)

Janka Stoker (University of Groningen)

Rob Alessie (University of Groningen)

Bertus Talsma (Ortec)

Abstract

This paper contributes to the literature on decision making among professionals. We studied the role of personal characteristics of referees and video assistant referees (VARs) in football. We look at age, experience and ranking for explaining the number of VAR moments per match and the confirmation of the initial referee decision. Our main finding is that VAR characteristics matter. Younger, less experienced and lower ranked VARs significantly recommend more reviews and they subsequently see the initial decision significantly more often confirmed. Although the introduction of the VAR intended to be an objective add-on for refereeing, our results show that VAR characteristics matter for the decision making process.

.

Keywords

Football, Decision making, Referee, Video assistant referee, Sports analytics

JEL classification codes

D81, D91, L83, Z20

1. Introduction

Sports data offer an interesting ground for economic research, especially when the rules of the game or external (exogenous) circumstances fundamentally change. In football, a recent prime example of the latter is the COVID-19 pandemic (Endrich & Gesche, 2020; Ferraresi & Gucciardi, 2021). Here we focus on a substantive change in the rules of the game via the introduction of the video assistant referee (VAR) in football. By now, all major football competitions have introduced the VAR. Operating under the philosophy of 'minimal interference, maximum benefit', the VAR system seeks to provide a way to correct for 'clear and obvious errors' and 'serious missed incidents' by the referee on the field[1].

Several studies investigated the effect of the VAR introduction. They focus mainly on the before-after effect (De Oliveira, Steffen, and Trojan, 2023). Our study is different, since we do not focus on comparing pre and post VAR-regimes. Instead, we focus on the characteristics of the VAR and the referee on the field. This an understudied area, and in as far as the characteristics of these officials matter, only characteristics of the referee are included (see e.g. Holder, Ehrmann, & König, 2022). The VAR itself is typically seen as a non-personalized and objective factor, which is remarkable, because the VAR is not a machine or algorithm but also an individual agent with personal characteristics.

Research on decision making among professionals shows that personal characteristics matter. More specifically, variables like status, experience and authority influence decision making among professionals like physicians, judges and teachers (Chan, Gentzkow & Yu, 2022; Figlio & Lucas, 2004; Kleinberg, Lakkaraju, Leskovec, Ludwig & Mullainathan, 2018; Van Parys & Skinner, 2016). Our interest is in a decision-making process where the relevant professionals (referee and VAR) differ in status and authority.

---

[1] See https://www.fifa.com/technical/football-technology/football-technologies-and-innovations-at-the-fifa-world-cup-2022/video-assistant-referee-var

Professional status and formal authority can be aligned, which implies that higher-status professionals have higher *formal* authority over lower-status professionals (e.g., Barley,1986; DiBenigno, 2018). But in many other organizational, and also sports contexts, there is at least *functional* authority to lower-status professionals, to oversee and direct a specific function (or set of functions) performed by higher-status professionals, who in the end makes the final decision (Karunakaran, 2022). In that situation, there is a misalignment or asymmetry between professional status and functional authority, also labeled as 'status–authority asymmetry'. Often, this becomes clear in situations aimed at compliance, e.g. safety auditors in the lab, D&I officers, or 911 dispatchers (Karunakaran, 2022). This line of research suggests that higher-status professionals often disregard the directions given by lower-status professionals with functional authority, as a way to reassert their dominant position in the professional hierarchy (Huising & Silbey, 2013).

The goal of our study is to investigate if the decision-making process between the higher-status official, the referee, and the lower-status official, the VAR, in football is influenced by personal characteristics the referee or VAR. To this end, we analyse data on all VAR moments from the highest Dutch football league for the 2019-2020 and 2020-2021 seasons.

2. Material and methods

A VAR is a match official, with independent access to match footage, who may assist the referee only in the event of a 'clear and obvious error' or 'serious missed incident'. The VAR can only recommend a 'review' to the referee; the referee always makes the final decision. The initial decision given by the referee will not be changed, unless the review shows that the decision was a 'clear and obvious error'.

Our data include all matches from the highest Dutch football league in 2019-2020 (232 matches[2]) and 2020-2021 (306 matches), in total 538 matches – due to missing data our final dataset consists of 531 matches. The mean (median) number of VAR moments in these matches is 0.97 (1.00). The percentage of games without any VAR moments is 39.55. For the characteristics of the VAR and referee, we measure age, experience and ranking. The referee's and VAR's mean age (in years) are 37.45 (SD 5.53) and 33.70 (SD 5.54) respectively. Experience is measured by the number of years that the match official has been active in the highest Dutch football league: the referee's and VAR's mean experience (in years) are 9.17 (SD 4.82) and 5.75 (SD 4.57) respectively. For formal ranking or category, we use the official classification of the Royal Dutch Football Association (KNVB) ranging from 1 (highest) to 5 (lowest). The referee and VAR are drawn from the same pool of match officials (which also means that they can and do switch roles in the course of the season). Individual category, age and experience are highly positively correlated; correlations between age, experience and category of the referee versus VAR (within a game) are not significant (all correlations available on request).

Our main independent variables are referee and VAR characteristics. We have two dependent variables: the decision-making process and outcome. The decision-making process concerns the 'recommending a review' action by the VAR, and consists of the number of total VAR moments in a match. The decision making outcome variable is measured as 'confirmed by referee', which is the case if - after the VAR-review - the initial decision of the referee stands and is thus confirmed (for possible decisions and outcomes, see Appendix). Next to the independent and dependent variables, we include a range of controls: the season (2019-2020 or 2020-2021), home team effect, timing of the VAR review during the game (in quantiles), the tightness of the score-line at the VAR moment (difference of max. two goals is a tight

---

[2] Season 2019-2020 contains fewer matches because it was halted in March 2020 due to the Covid-19 pandemic.

game), a fixed effect for each official (are some officials for reasons unrelated to their age, experience or category more or less inclined to call for a review or to confirm their initial decision in their roles as VAR and referee respectively) and a fixed effect for each team (like major or higher-ranked clubs).

3. Results

Table 1 shows the main estimation results for the number of VAR moments in a match. Column (1-3) show that age, experience and category of the VAR respectively, significantly predict the number of VAR moments in a match, such that the younger, less experienced and low ranked VARs recommend more reviews. The characteristics of the referee matter less; only the age of the referee is marginally significantly associated with number of VAR moments. We find no effect of ranking differences between the VAR and referee, nor any interaction-effects.

# VAR moments	1.	2.	3.
Age referee	-0.016* [0.009]		
Age VAR	-0.020** [0.009]		
Experience referee		-0.013 [-0.011]	
Experience VAR		-0.027*** [0.010]	
Category referee			-0.012 [0.040]
Category VAR			-0.072** [0.036]
Season dummy	Y	Y	Y
Home team dummy	Y	Y	Y
FE official	Y	Y	Y
FE team	Y	Y	Y
Observations	531	531	531
R^2	0.071	0.070	0.062

Table 1. Number of VAR moments

Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p<0.01$. Not shown are sub-sample estimations for a) timing of VAR moment (quantiles), b) tightness score-line (more/less than 2 goal difference) or c) type of VAR event (see Appendix).

In table 2, the results for the outcome of the VAR-moment are shown. The dependent variable is whether after the VAR moment, the decision of the referee is confirmed. Columns 1-3 show

the results for age, experience and category, respectively. We find that the characteristics of the VAR significantly predict the dependent variable. Younger, less experienced and low ranked VARs are associated with more confirmed decisions of the referee. For the referee, only age is again a marginally significant predictor. Again, we find no effect of ranking differences between the VAR and referee, nor any interaction-effects. Also, we do not find any different effect for types of VAR events, as listed in the Appendix.

Decision referee confirmed	1.	2.	3.
Age referee	-0.013* [0.008]		
Age VAR	-0.015** [0.007]		
Experience referee		-0.013 [-0.009]	
Experience VAR		-0.022** [0.009]	
Category referee			-0.012 [0.035]
Category VAR			-0.053* [0.031]
Season dummy	Y	Y	Y
Home team dummy	Y	Y	Y
FE official	Y	Y	Y
FE team	Y	Y	Y
Observations	531	531	531
R^2	0.056	0.058	0.048

Table 2. Decision of the referee confirmed.

Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p<0.01$. Not shown are sub-sample estimations for a) timing of VAR moment (quantiles), b) tightness score-line (more/less than 2 goal difference) or c) type of VAR event (see Appendix).

4. Discussion

We studied the role of referees and VARs characteristics in Dutch football matches. Specifically, we looked at the role of age, experience and ranking for two dependent variables, namely the number of VAR moments per match and the confirmation of the initial decision by the referee after the VAR moment. Results clearly indicate that VAR characteristics matter. Younger, less experienced or lower ranked VARs significantly more often recommend a review of the initial referee decision, but they also subsequently get the initial referee decision more often confirmed.

Our results contribute to research into the effect of the VAR in football. Thus far, these studies treated the VAR as an objective 'machine'. If these studies look at

characteristics of actors in the game at all, they include those of the referee (e.g. Holder et al, 2022). Although it is assumed that the VAR is objective, we show that their decision-making behavior is subjective and dependent on individual characteristics. Second and more generally, our results contribute to the literature on decision making by professionals who differ in status and authority. Chan et al. (2022) showed that inexperienced/less skilled radiologists diagnose more cases with a disease and miss more cases that have the disease. Correspondingly, we find that inexperienced VARs recommend more reviews, and that these reviews lead to more confirmation of the original referee decision, implying they are more 'wrong' in their recommendations. To conclude, although the introduction of the VAR intended to be an objective add-on for refereeing, we must realize that the VAR is a human with personal characteristics that matter significantly for the course of the match.

References

Barley, S. R. (1986). Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Administrative Science Quarterly*, 78-108.

Chan, D. C., Gentzkow, M., & Yu, C. (2022). Selection with variation in diagnostic skill: Evidence from radiologists. *The Quarterly Journal of Economics*, *137*(2), 729-783.

De Oliveira, M. S., Steffen, V., & Trojan, F. (2023). A systematic review of the literature on video assistant referees in soccer: Challenges and opportunities in sports analytics. *Decision Analytics Journal,* 100232.

DiBenigno, J. (2018). Anchored personalization in managing goal conflict between professional groups: The case of U.S. Army mental health care. *Administrative Science Quarterly, 63*, 526–569.

Endrich, M., & Gesche, T. (2020). Home-bias in referee decisions: Evidence from "Ghost Matches" during the Covid19-Pandemic. *Economics Letters*, *197*, 109621.

Ferraresi, M., & Gucciardi, G. (2021). Who chokes on a penalty kick? Social environment and individual performance during Covid-19 times. *Economics Letters*, *203*, 109868.

Figlio, D.N., & Lucas, M.E. (2004). Do High Grading Standards Affect Student Performance? *Journal of Public Economics, 88,* 1815–1834.

Holder, U., Ehrmann, T., & König, A. (2022). Monitoring experts: insights from the introduction of video assistant referee (VAR) in elite football. *Journal of Business Economics, 92*, 285-308.

Huising, R., & Silbey, S.S. (2013). Constructing consequences for noncompliance: The case of academic laboratories. *The Annals of the American Academy of Political and Social Science, 649*, 157–177.

Karunakaran, A. (2022). Status–authority asymmetry between professions: The case of 911 dispatchers and police officers. *Administrative Science Quarterly*, *67*, 423-468.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, *133*, 237-293.

Van Parys, J., & Skinner, J. (2016). Physician Practice Style Variation: Implications for Policy. *JAMA Internal Medicine, 176,* 1549–1550.

Appendix Type of VAR events and possible outcomes

VAR event	Original Referee decision	Outcome (after VAR)
Goal	Confirmed	Goal
Goal	Cancelled	No goal
No goal	Confirmed	No goal
No goal	Cancelled	Goal
Penalty awarded	Confirmed	Penalty
Penalty awarded	Cancelled	No penalty
Penalty not awarded	Confirmed	No penalty
Penalty not awarded	Cancelled	Penalty
Red card given	Confirmed	Red card awarded
Red card given	Cancelled	No red card
Card upgrade	Confirmed	No Red card
Card upgrade	Cancelled	Red card awarded
Mistaken identity	Confirmed	Player originally cautioned/sent off confirmed
Mistaken identity	Cancelled	Ref cautions/sends off a different player