

Murat C. Mungan*

Optimal Real-Time Review Standards: Implications for Law Enforcement and Competitive Games

<https://doi.org/10.1515/rle-2024-0117>

Received October 5, 2024; accepted November 24, 2024; published online March 24, 2025

Abstract: Real-time review systems are frequently used in various sports to monitor the decisions of referees and correct their mistakes. Interventions through these systems cause delays in games, which are perceived as being costly. This makes it optimal for these review systems to interfere with the decisions of the referee less frequently than would minimize the costs of decision errors, which I formalize through an analysis of the VAR system in football. This analysis also reveals that optimal review standards ought to be laxer when an important event (e.g., a goal) occurs between the position in which the potential error took place and the VAR intervention. In the near future, it may be possible to introduce similar review systems in the law enforcement context, e.g., by utilizing police officers' body cameras. I compare optimal intervention standards in this context to their analogues in the sports context, and discuss implications.

Keywords: real-time decision review; video review; video assistant referee; VAR; review standard; clear and obvious error

JEL Classification: K0; K4; K41; K42; Z2; Z28; Z29

The author has not received any funding from third parties to write this article.

I thank the editors and an anonymous referee for valuable insights and suggestions on an earlier draft of this article which led to substantial and important revisions.

***Corresponding author: Murat C. Mungan**, Professor, Texas A&M University School of Law, Fort Worth, USA, E-mail: mungan@tamu.edu

 Open Access. © 2025 the author(s), published by De Gruyter.  This work is licensed under the Creative Commons Attribution 4.0 International License.

1 Introduction

Decisions made in legal proceedings and those made by referees in sports share a lot in common. Moreover, while a lot of theorizing about judges' decision making has been provided in the law and economics literature,¹ sports competitions have taken the lead in incorporating new real-time review technologies (e.g., VAR in football, Instant Replay in American football, Hawk-Eye in tennis, DRS in cricket, and TMO in rugby). Soon, it may be possible to introduce similar real-time review tools into law enforcement to serve similar functions (e.g., to review the decisions of police officers through their body cameras). Thus, the technological review tool adoption experiences of various sports competitions can inform the optimal design of how similar decision making tools ought to be incorporated into law enforcement in the future, and analyses of decision making in sports can benefit from incorporating the lessons that emerge from the economics literature on legal procedure. Moreover, this type of academic exchange can also be relevant to studying protocols in seemingly unrelated fields, such as the military context, where similar review technologies exist. Despite this, the law and economics and sports economics literatures have not yet imported many applicable insights from each other.

I illustrate how this type of information exchange may occur by focusing on the similarities between the sports and law enforcement contexts. I examine a game in which a "ground referee" must make a decision and that decision can be reviewed by a "technological referee". In both the sports and law enforcement settings, a technological referee may (on average) improve decision making by overruling erroneous decisions. However, intervention by a technological referee generates costs by slowing down the game (which may be in conflict with the nature of the sport, and it may lead to reduced enforcement efficiency, in the two settings, respectively). Given these benefits and costs, I examine the optimality conditions associated with intervention by a technological referee.

To provide a concrete example I begin my analysis with the case of association football (called 'soccer' in the United States; and henceforth simply 'football' throughout this article). Specifically, I draw on the vast law and economics literature on optimal standards of proof to study a review system that has relatively recently been introduced to football, namely the Video Assistant Referee (VAR) system. In addition to identifying optimal review standards for this system, the analysis more generally highlights how one can think of optimal review processes in sports through the lens of law and economics. Moreover, rationales for the adoption of

¹ Due to the vastness of this literature, it is difficult to exhaustively list the scholarship that has contributed to it; examples include Miceli (1990, 1991), Demougin and Fluet (2005), Hylton and Khanna (2007), Mungan (2011, 2020), Rizzolli and Saraceno (2013), and Mungan and Samuel (2019).

conservative review standards in the VAR process suggests that it may be optimal to use similarly conservative review standards in real-time review technologies of law enforcer decisions that may be adopted in the near future.

Although I focus on the VAR example from football, I suspect the analysis is equally applicable to the real-time review technologies adopted in many other sports. Thus, I first provide a brief review of some relevant characteristics of football and the introduction of VAR.

Football is a fast-paced game involving simultaneous and multiple movements on different parts of the pitch. This makes it difficult for a single referee to catch every incident throughout a match. This has led to many infamous incidents of refereeing errors throughout the history of the game. In an effort to reduce such errors, the VAR system was introduced to football with its full implementation occurring in the 2018 World Cup held in Russia.

One of the key entities in this system is the VAR team, consisting of a VAR official; three assistant video assistant referees (AVARs); and replay operators. The VAR team constantly monitors the match by reviewing video footages of incidents – sometimes from dozens of different cameras² – and analyzing them with advanced technological tools, e.g., the “semi-automated offside technology” (FIFA 2024). The VAR team’s responsibility is to alert the referee in cases where the referee’s decision might be in error, but only if this error relates to four well-defined categories of incidents involving (i) goals and offenses related to goals; (ii) penalties and offenses related to penalty decisions; (iii) direct red cards; and (iv) cases of mistaken identity (e.g., booking the wrong person for a yellow card) (IFAB 2024). Even in these instances, the VAR is directed to communicate with the referee only in cases of “clear and obvious errors” (FIFA 2024).

Soon after its introduction, the system was met with controversy. Many fans complained about VAR reviews slowing down the pace and changing the character of “the Beautiful Game”, and suggested that its interference ought to be limited (see, e.g., Carragher 2023). Others, including coaches, have complained that VAR is being used unfairly to the advantage of larger teams, perhaps to maximize the monetary profit from football events.³ In fact, there have been instances in which the VAR review clearly led to important erroneous decisions, which resulted in the release of VAR audio recordings and subsequent public criticisms of VAR.⁴ Perhaps inspired

2 For instance, Fédération internationale de football association (FIFA), states that “[t]he video assistant referee team has access to 42 broadcast cameras” FIFA (2024). However, some have noted that the numbers of cameras in some competitions are much smaller, see, e.g., Kassouf (2023).

3 See, e.g., Jose Mourinho’s comments regarding VAR being used “to help the money magnet teams or famous teams” as reported in Brown (2024).

4 A great example is the clearly erroneous cancellation of Sivasspor’s goal against Galatasaray upon VAR intervention in the Turkish Super Lig in 2022. After the incident the Turkish referee

by similar events, various academic investigations of VAR's effects on important match statistics have been conducted,⁵ and fans have made comments on whether the VAR protocol can be reformed to reduce its negative impacts on the game (see, e.g., Ashdown 2023).

Here, I provide the first—to the best of my knowledge—economic theory based analysis of *when* VAR ought to alert the referee to a potential mistake. In other words, I identify the optimal review standards for VAR triggered review. I show that the standard currently in place, namely the spotting of a “clear and obvious error” can be rationalized as an attempt to balance undesirable stoppages and pace-slowness effects of VAR review against the objective of minimizing the referee's decision errors. I then ask whether VAR reviews ought to be initiated more liberally or more sparingly in cases where an intervening important event occurs between the potential mistake spotted by the VAR and the next stoppage of the game. A concrete example of this type of situation is where the referee does not call a potential penalty for team X, and subsequently team Y scores a goal on a counter-attack. I show that in these cases, perhaps counter-intuitively, VAR review ought to be implemented more liberally. This is because the costs associated with erroneous decisions are greater in these circumstances, which makes the costs associated with game-stoppage and pace-slowness relatively less important.

After formalizing these points, I note a similar problem that may emerge in the law enforcement context, if police officers' body cameras were used to serve a similar function. As an example, I analyze the case where a police officer's decision to stop a person is reviewed in real-time (whether by other humans or technological tools). In this case, attempting to correct the police officers' mistakes in real time may generate costs by reducing their effectiveness in conducting further enforcement actions. It is then optimal to use a more conservative standard than that which minimizes decision error costs. However, the optimal strength of the standard in absolute terms (as opposed to how it relates to the error-cost minimizing standard) depends on the relative weights of the error costs (i.e., the costs from stopping a person who ought not to be stopped versus letting a person go who ought to be stopped), which may differ greatly from their analogues in the sports contexts. One important difference, for instance, is that a significant part of the costs associated

board admitted the error, and released the VAR audio recordings, in which the VAR referee, Özgür Türkalp, was heard informing the on-field referee that a position *may* involve what he described in Turkish as an “ofsaytımsı” (roughly translatable as “offside-like”) position. This incident fueled public criticism and resulted in Türkalp's removal from future officiating assignments. See, e.g., Habertürk (2023).

⁵ See, e.g., Errekagorri et al. (2020) and Lago-Peñas, Gómez, and Pollard (2021), as well as the sources referenced in de Oliveira, Steffen, and Trojan (2023), which provides a review of existing studies.

with the erroneous stopping of a person can be mitigated ex-post, as opposed to real time, which reduces the ‘irreversible’ cost of a mistaken stop. This may reduce the relative potential gains from implementing real-time review processes in the law enforcement context, and thereby provide a rationale for delaying the adoption of such systems until their costs fall significantly as a result of technological progress.

It is worth noting that the type of video-refereeing technologies analyzed here have been used in other contexts as well. For instance, in the military context, soldiers (analogous to the “ground referee”) may need to obtain permission from an overseer (analogous to the “technological referee”) prior to taking an important action (e.g., a sniper strike). An example is the later disputed account of some service members of the Abbey Gate bombing, in which a man “detonated a bomb outside the Kabul airport in August 2021, killing 170 Afghans and 13 American service members (...). Some service members who were at the airport that day claimed they had spotted the suicide bomber at the site and were ordered not to engage” (Kube and Gains 2024).⁶ The sequence of events and the relationship between the decision makers are slightly different in this context than in the sports and law enforcement contexts.⁷ However, the analysis provided here can be extended to analyze optimal protocols (e.g., when to seek permission versus execute an action without doing so) in military contexts, too.

In the next section, I present a model to study VAR decision making, and use it to derive the results related to optimal VAR review summarized above. In Section 3, I discuss the implications of this model in the law enforcement context. In Section 4, I provide concluding remarks.

2 Model

I consider the decision making process of a Video Assistant Referee team, which I collectively refer to as VAR (pronoun *it*). The VAR continuously monitors a football game and reviews it for discrepancies between what has happened on the pitch and the decisions made by the referee (pronoun *she*). As noted by the International Football Association Board (IFAB), VARs duty to is to intervene and alert the referee

⁶ See, Kube and Gains (2024) for further details, which notes that the account of the service members were denied by a later review which suggested that the service members had the wrong person in their sights.

⁷ This is because the soldier needs to hold off on execution of the action unless he receives approval from the overseer. If the overseer is unable to respond quickly, the opportunity to execute the action may pass. In football, the referee makes a decision which stands unless later over-ruled. Thus, the cost is the slowing of the game as opposed to the costs of missing opportunities to execute beneficial actions.

to what it identifies as “a clear and obvious error”, but only in the four specific cases noted in the introduction. I collectively refer to these specific cases as incidents, and model the decision making process of the VAR in each such incident.

2.1 The Set-Up

At each incident, the VAR observes the referee’s decision and forms a belief about the likelihood of it being correct. Given the binary nature of the issues about which the referee makes a decision in these instances (i.e., goal/no goal, penalty/no penalty, red card/not, mistaken identity/not), I denote the state of the world as $t \in T = \{0, 1\}$ and, without loss of generality, I assume that the referee’s initial decision is correct if $t = 0$. To illustrate this notation more concretely, consider the following simple example. When the referee rewards a goal, the state of the world where the goal should not have been rewarded (e.g., due to a preceding off-side violation) is denoted $t = 1$; and it is denoted $t = 0$, if the goal is a proper one. The VAR does not know the true state of the world, but forms a belief regarding it based on the information it receives, e.g., from video reviews. The VAR’s beliefs are reflected by the likelihood, π , with which it believes the referee’s decision is correct.

Upon forming this belief, the VAR makes a decision about whether to alert the referee to reconsider her call. The VAR’s action is thus denoted $a \in A = \{0, 1\}$ where 1 indicates the decision to alert the referee, and 0 indicates not interfering. Similarly, the referee’s final decision is denoted $d \in D = \{0, 1\}$. If the VAR does not alert the referee, the referee’s existing decision stands, i.e., $d = 0$. On the other hand, if the VAR alerts the referee, the referee reviews the available information, and takes an action based on it: $d = 1$ denotes a decision reversal and $d = 0$ indicates that the referee’s previous decision stands. The referee’s probability of preserving her initial decision is $p(\pi) \in [0, 1]$ with $p' \geq 0$. Here, the dependency of p on π reflects the idea that the available information affects the referee’s decision making process. I assume the VAR correctly assesses this probability, but, as will become apparent below, this simplifying assumption has no impact on the analysis that follows. I also assume that there are a continuum of instances leading to a distribution of π being observed by the VAR, such that the VAR’s problem is to determine an optimal decision rule for all $\pi \in [0, 1]$.

2.2 The VARs Decision Matrix

At every instance, the VAR’s decision may result in a correct or incorrect outcome with different probabilities. First, if the VAR does not interfere, then it preserves the decision on the pitch, which is correct with probability π and incorrect with probability $1 - \pi$. On the other hand, if the VAR alerts the referee, the referee changes

Table 1: VAR actions and outcomes.

		<i>State of the world</i>	
		<i>t = 1 (with prob. $1 - \pi$)</i>	<i>t = 0 (with prob. π)</i>
		<i>Decision/outcome pair</i>	
<i>Don't interfere</i>			
$a = 0$		$d = 0; t = 1$	$d = 0; t = 0$
		<i>Incorrectly preserved</i>	<i>Correctly preserved</i>
<i>Alert referee</i>			
$a = 1$	Ref. preserves decision with prob. $p(\pi)$	$d = 0; t = 1$	$d = 0; t = 0$
		<i>Incorrectly preserved</i>	<i>Correctly preserved</i>
	Ref. changes decision with prob. $1 - p(\pi)$	$d = 1; t = 1$	$d = 1; t = 0$
		<i>Correctly changed</i>	<i>Incorrectly changed</i>

her initial decision with probability $1 - p(\pi)$, which was correct with a probability of π . Thus, with probability $(1 - p(\pi))\pi$, the referee incorrectly changes her standing decision and with probability $p(\pi)(1 - \pi)$ she incorrectly fails to switch her decision; both constituting errors. Finally, with probabilities $p(\pi)\pi$ and $(1 - p(\pi))(1 - \pi)$, the referee correctly changes and correctly refuses to change her decisions, respectively. These possibilities are summarized in Table 1, below.

2.3 Optimal VAR Review

As Table 1 illustrates, the combination of the true state of the world and the VAR's actions can lead to four different outcomes, with correct outcomes when $d = t$ and incorrect outcomes when $d \neq t$. The utility associated with these correct and incorrect decisions, in terms of serving the objectives of the game, can thus be denoted u_{dt} . In addition to this decision related utility, there may be direct costs (or benefits) associated with initiating a VAR review, e.g., from stopping play. Thus, I denote the direct disutility from each action $a \in \{0, 1\}$ as k_a , which I assume for simplicity have an additive relationship with the decision related utilities u_{dt} . Under these assumptions, a VAR seeking to maximize the objectives of the game, perceives the following expected utilities from each action:

$$U(a) = \begin{cases} (1 - \pi)u_{01} + \pi u_{00} - k_0 & \text{if } a = 0 \\ p(\pi)[(1 - \pi)u_{01} + \pi u_{00}] + (1 - p(\pi))[(1 - \pi)u_{11} + \pi u_{10}] - k_1 & \text{if } a = 1 \end{cases} \quad (1)$$

Thus, it is (weakly) optimal for the VAR to alert the referee if, and only if $U(1) \geq U(0)$, which corresponds to the condition that

$$(1 - p(\pi))\{[(1 - \pi)u_{11} + \pi u_{10}] - [(1 - \pi)u_{01} + \pi u_{00}]\} \geq k_1 - k_0 \quad (2)$$

which can be re-written as

$$(1 - p(\pi))[(1 - \pi)C_1 - \pi C_0] \geq K \quad (3)$$

where

$$\begin{aligned} K &\equiv k_1 - k_0; \\ C_1 &\equiv u_{11} - u_{01}; \text{ and} \\ C_0 &\equiv u_{00} - u_{10} \end{aligned} \quad (4)$$

Here, K can be interpreted as the direct cost of initiating a VAR review relative to letting the game play on, and C_t are the relative costs of making a decision mistake when the state of the world is $t \in \{0, 1\}$.

To render the analysis meaningful I assume that it is strictly preferable for the VAR to interfere (i.e., $U(1) > U(0)$) when $\pi = 0$, i.e., when it is certain that the referee's decision is erroneous. In addition, I impose a restrictive assumption regarding C_t to investigate optimal VAR interference standards. In general, the costs of the two types of decision errors that can be committed in a binary decision problem need not equal each other. This may occur, for instance, because the two errors typically lead to qualitatively very different harms (see e.g., Mungan 2011; Rizzolli and Saraceno 2013, in the criminal trial context; the same may be true in the law enforcement context studied in Section 3, below). However, in the VAR context studied here, there are very natural, symmetry related reasons, to assume that these two costs are equal. Whenever the commission of an error leads to an unjust benefit for one team, it naturally unjustly harms the other team: there is a simple mechanical relationship which causes the increase in the probability with which one team wins due to erroneous decisions to equal the probability with which the other team unjustly loses.

Due to these reasons I restrict attention to cases where the cost of the two errors is equal: $C_1 = C_0 = C$ for some $C > 0$. In these cases, it follows that the existing standards for VAR review is consistent with the optimal review process if, and only if, there are direct relative costs associated with initiating VAR review (i.e., $K > 0$). I formalize this result through the following proposition.

Proposition 1. *Suppose $C_0 = C_1 = C$, then:*

(i) *if $K > 0$, there exists $\pi^* < \frac{1}{2}$, such that it is (weakly) optimal for the VAR to initiate review if, and only if, $\pi \leq \pi^*$;*

(ii) if $K = 0$, it is (weakly) optimal for the VAR to initiate review whenever $\pi \leq 1/2$; and

(iii) if $K < 0$, there exists $\pi^{**} > \frac{1}{2}$, such that it is optimal for the VAR to initiate review if $\pi \leq \pi^{**}$.

Proof. When $C_0 = C_1 = C$ the expression in (3) becomes

$$Z(\pi) \equiv (1 - p(\pi))(1 - 2\pi)C - K \geq 0 \quad (5)$$

It follows that $Z'(\pi) < 0$ for all $\pi \leq 1/2$, since $p' \geq 0$. Moreover, $Z(0) > 0$, because it is assumed that the VAR ought to interfere when $\pi = 0$. The expression in (5) and its properties listed above are used in deriving results.

(i) When $K > 0$, it follows that $Z(1/2) < 0$. Since, in addition, $Z(0) > 0$, and $Z'(\pi) < 0$ for all $\pi \leq 1/2$, it follows that among $\pi \in [0, 1/2]$ there's a unique π^* such that $Z(\pi) = 0$. When $K > 0$, it also follows that $Z(\pi) < 0$ for all $\pi > 1/2$, and thus $Z(\pi) \geq 0$ iff $\pi \leq \pi^*$.

(ii) When $K = 0$, it follows that $Z(1/2) = 0$. In addition, $Z'(\pi) < 0$ for all $\pi \leq 1/2$ and $Z(\pi) < 0$ for all $\pi > 1/2$. Thus, $Z(\pi) \geq 0$ iff $\pi \leq 1/2$.

(iii) When $K < 0$, it follows that $Z(1/2) > 0$, which together with the fact that $Z'(\pi) < 0$ for all $\pi \leq 1/2$, implies that $Z(\pi) > 0$ for all $\pi \leq 1/2$. Moreover, since $Z(1/2) > 0$, it follows that $Z(\pi) > 0$ for all $\pi \in [1/2, \pi^{**}]$. Thus, there exists $\pi^{**} > 1/2$ such that $\pi < \pi^{**}$ implies that $Z(\pi) > 0$. ■

The results formalized by Proposition 1 are relatively intuitive. First, if there are no direct consequences associated with VAR decisions, or when the costs or benefits associated with initiating and not initiating a VAR review are equal (i.e., $K = 0$), then the only objective is to minimize expected decision error costs. This is achieved by the VAR alerting the referee whenever they catch a position in which a mistake appears to be more likely than not;⁸ i.e., $\pi > 1/2$. In the related law and economics literature on optimal standards of proof, this standard is often interpreted as corresponding to the preponderance of the evidence standard. The standard called for by IFAB is stronger as it requires “clear and obvious error” as opposed to an error being simply more likely than no error.

On the other hand, when $K > 0$, the VAR ought to apply a stricter standard: It would need to interfere only if the probability with which the referee's call is greater than some threshold which exceeds $1/2$. The reason is that the VAR now has two conflicting goals: one is to alert the referee as little as possible, because

⁸ In this case, the VAR ought to be indifferent when $\pi = 1/2$.

this leads to expected game stoppage costs, the second is to alert the referee to correct mistakes efficiently, which requires alerting the referee whenever $\pi \leq 1/2$. This trade-off leads to an intermediate optimal VAR standard in which the referee is alerted less frequently than would minimize expected decision errors, i.e., $\pi^* < \frac{1}{2}$. It is, of course, difficult (and perhaps counter-productive) to attempt to place an exact probability threshold that corresponds to the “clear and obvious error” standard. But, as long as this standard is stronger than a preponderance of the evidence standard, its use is consistent with the optimal standards that emerge from this model only when one acknowledges the expected direct costs associated with initiating a VAR review.

Finally, I consider the case where $K < 0$ for completeness, although this case appears counter-intuitive. If there were net gains from initiating VAR review, e.g., because it creates a sense of procedural justice among fans, then it would be optimal for VAR to interfere even when it believes that the referee’s decision is more likely to be correct than not. This case is naturally inconsistent with the standard proposed by IFAB, because it calls for a standard that is even weaker than preponderance of the evidence.

2.4 The Impact of an Intervening Important Event

Next, I consider an interesting feature of VAR review in football. If in a given position, the referee makes a decision which does not lead to the stopping of the game, the VAR may review the position that gave rise to the decision while the game is ongoing, and alert the referee regarding potential decision errors once the game stops. In this process, an important intervening event may take place. Again, I provide a concrete example to illustrate this idea. Suppose team X ’s shot on goal is blocked by a defender inside the 18-yard box. The referee spots no illegal moves and allows the game to continue. Subsequently, team Y scores a goal on the counter-attack. During the counter-attack, the VAR reviews the blocked shot in the initial position, and notices the ball bouncing off the defender’s hand. However, it is questionable whether the player is in their natural position, and thus whether the blocked shot should have resulted in a penalty kick in favor of team X . In this example, the intervening important event is the goal scored by team Y .

This type of intervening important event increases the importance of decision errors. In the example provided, if the shot was blocked through an illegal hand-play, and the VAR fails to interfere, then it would be incorrectly rewarding team Y with a goal *and* incorrectly depriving team X of a penalty kick. A similar statement can be made for the other type of decision error. Thus, important intervening events increase the costs of decision errors, relative to cases in which no important

intervening event takes place (e.g., instead of team Y scoring a goal, the game is stopped due to a throw-in around midfield).

In the model described, the occurrence of an important intervening event can be incorporated through an increase in the decision error cost C . An interesting question is whether the optimal standard for VAR intervention becomes stricter or laxer in such cases. To answer this question, I conduct a simple comparative statics analysis to assess the impact of C on the threshold belief that the VAR must possess to alert the referee. In this exercise, I focus on the case where $K > 0$, since this is the only case that is consistent with IFAB's existing guidelines.

Proposition 2. *Suppose $K > 0$. Then, the VAR ought to alert the referee more liberally when an important intervening event occurs after the referee makes the reviewed decision (i.e., $\frac{d\pi^*}{dC} > 0$).*

Proof. To reflect the potential dependency of the VAR's optimal action on C , we can re-write (5) as:

$$Z(\pi, C) = (1 - p(\pi))(1 - 2\pi)C - K$$

From the analysis in Proposition 1-i, it follows that the unique decision threshold, π^* , is obtained when $Z(\pi^*, C) = 0$. Using the implicit function theorem, it follows that

$$\frac{d\pi^*}{dC} = -\frac{Z_C}{Z_\pi} = \frac{(1 - p(\pi))(1 - 2\pi^*)}{p'(\pi^*)(1 - 2\pi^*)C + 2(1 - p(\pi^*))C} = \frac{(+)}{(+)} > 0 \quad (6)$$

where the positivity of the denominator in (6) is guaranteed by the facts that $p' \geq 0$ and $\pi^* < 1/2$. ■

Interestingly, and perhaps against some football fans' intuition, the optimal standard for VAR interference becomes laxer when an intervening important event occurs in football. The rationale is simple: the relative magnitude of decision errors, i.e., what is at stake, becomes larger when an intervening event occurs. Therefore, relative to the objective of not stopping the game, the objective of reducing decision errors becomes more important. This pulls the optimal standard closer to the standard that would be used to minimize decision errors, namely one where the VAR interferes whenever an error is more likely than not.

3 Optimal Review of Law Enforcement Stops

In this section I extend the analysis of the VAR system to analyze a similar problem which may arise in the law enforcement setting through the adoption of

technological review processes. Specifically, I consider the possibility of body cameras worn by police officers being used for purposes of identifying problematic stops. The analytic set-up I consider in this case is very similar to that which arises in the analysis of the VAR system in football. The police officer (pronoun *he*) monitors an area and makes decisions about whether to stop individuals based on suspicion that they may be involved in criminal activity. A technological review system (henceforth ‘TRS’), in turn, monitors the decisions of the police officer. If the police officer makes a stop, the TRS forms a belief regarding the appropriateness of the stop (reflected by π , as in the VAR analysis), and may alert the police officer to suggest that he over-turn his decision and cancel the stop. If the TRS chooses to do so, it also provides a summary of the basis of this recommendation. The police officer reviews the information provided, and determines whether to over-turn his decision (which occurs with probability $1 - p(\pi)$). This review process is distracting and time consuming, and thus leads to the slowing down of law enforcement, which generates a cost of $K > 0$. Thus, the pay-off structure associated with this interaction is similar to that described via (1)–(3) in the VAR context, and the normalized net-gains from the TRS initiating a review can be written as:

$$V(\pi, r, q) = (1 - p(\pi))(1 - \pi(1 + r)) - q \geq 0 \quad (7)$$

where

$$r = \frac{C_0}{C_1}$$

denotes the error cost of not conducting a stop (when it ought to be conducted) relative to the error cost of conducting a stop (when it ought not to be conducted) and

$$q = \frac{K}{C_1}$$

is the normalized cost of slowing enforcement down.

With this notation, the assumption made in the VAR context analyzed in Section 2 corresponds to that of $r = 1$. In the law enforcement context, this assumption may not be warranted, because the two error-costs may carry very different consequences. In particular, an erroneous failure to stop a person may constitute a failed opportunity to prevent criminal harm (reflected by C_0) whereas an erroneous stop may cause inconvenience costs to the suspect (see Mungan 2018 for a more detailed discussion of these costs) in addition to potential wrongful prosecution, and in extreme cases wrongful convictions (which are captured by C_1). Given that expected wrongful prosecution and conviction costs can be reduced through ex-post review of the same information that the TRS reviews to make recommendations, one may expect C_1 to be closer to the inconvenience costs associated with stops. Thus, in what follows I

identify the optimal review standard of the TRS without imposing the constraint that $r = 1$.

For this analysis, it is useful to identify the standard which minimizes decision error costs. These costs are captured by $V(\pi, r, 0)$, which equals the net-gains from the TRS initiating a review without incorporating the negative impacts of the review process on the enforcement function of the police officer. With these definitions in place, the properties of the optimal review process can be summarized as follows.

Proposition 3. (i) Decision error costs are (weakly) minimized by alerting the police officer to a potentially illegal stop when, and only when, $\pi \leq \pi_0^* \equiv \frac{1}{1+r}$. (ii) When $q > 0$, there exists $\pi^* < \pi_0^*$ such that it is (weakly) optimal to alert the police officer to a potentially illegal stop when, and only when, $\pi \leq \pi^*$. (iii) Everything else equal, a greater r is associated with an optimal review rule that is stricter (i.e., $\pi_r^* < 0$).

Proof. (i) The net-gains from TRS when $q = 0$ are given by

$$V(\pi, r, 0) = (1 - p(\pi))(1 - \pi(1 + r)) \geq 0 \text{ iff}$$

$$\pi \leq \frac{1}{1+r}$$

(ii) Next, note that $V_\pi(\pi, r, q) = -p'(1 - \pi(1 + r)) - (1 + r)(1 - p(\pi)) < 0$ for all $\pi \leq \frac{1}{1+r}$. Moreover, $q > 0$ implies that $V(\pi, r, q) \leq 0$ for all $\pi \geq \frac{1}{1+r}$. Thus, there exists a unique $\pi^* \in (0, \frac{1}{1+r})$, such that $V(\pi, r, q) \geq 0 \Leftrightarrow \pi \leq \pi^*$.

(iii) The optimal review rule is characterized by $V(\pi^*, r, q) = 0$. Thus, $\pi_r^* = -\frac{V_r}{V_\pi} = \frac{(1-p(\pi^*))\pi^*}{V_\pi(\pi^*, r, q)} = \frac{(+)}{(-)} < 0$. ■

Unlike in the VAR context, it is not possible to compare the optimal standard of review to the more likely than not standard used in other contexts, without making further assumptions about the relative error costs that can be committed in the enforcement context. However, when the inconvenience costs associated with stops are small relative to the harms from the crimes that the law enforcer is seeking to prevent, it follows that $r > 1$. In this case, an implication of Proposition 3 is that the optimal review standard is stricter.

It is worth noting that the analysis here takes the costs associated with enforcement exogenously given. As illustrated in Mungan (2018), this is a harmless assumption in circumstances where the function of law enforcement is largely preventive. However, in cases where the enforcement policies implemented (reflected by π^* in the current analysis) may have deterrence effects, these costs may be responsive to the review standards adopted by the TRS. Even in these cases, the main results presented here are likely to be preserved as long as the review-standard elasticity of crime is small.

3.1 Extending to the Military Context

Finally, the analysis can be extended to study the interactions in the military context where a soldier must seek approval from an overseer to execute important actions. The Abbey Gate bombing noted in the introduction provides an example of this type of interaction. However, more specific assumptions regarding the interactions between the overseer and the soldier may be appropriate in analyzing these cases. It may be important, for instance, to specify whether the overseer's input to the soldier is a command with which the soldier must comply, or whether it is a recommendation. If it is the former, then it would be more appropriate to model

the probability of action execution as being binary (e.g., $p(\pi) = \begin{cases} 0 & \text{if } \pi < \pi^* \\ 1 & \text{if } \pi \geq \pi^* \end{cases}$).

Another modeling choice which may be important in this context relates to the choice set of the overseer. If certain actions, e.g., sniper strikes, always require prior approval, this may transform the question from 'when' to interfere to the more familiar question in the standard of proof literature of 'how' to interfere. Thus, an interesting issue in this setting relates to the mandatory versus optional input of the overseer as a function of the importance of the action being undertaken. This issue can be analyzed through extensions of the model provided herein.

4 Conclusions

There are many similarities between decisions made in the law enforcement, sports, and even in the military contexts. Yet, there appears to be little information flow between academic studies in these fields. Here, I focused on the real-time review procedures in sports—and the VAR system in football in particular—to discuss optimal intervention standards by bringing insights from the law and economics literature. I then questioned how real-time review processes may look like in the context of law enforcement, if similar processes were implemented in those contexts. In closing, I make a few remarks regarding optimal VAR intervention standards as well as the viability of implementing similar procedures in the law enforcement context, and how these insights could potentially be imported to study decision making in the military context.

VAR was introduced to football with the goal of reducing referee errors. However, its introduction sparked controversy among football fans, especially those who are skeptical about the benefits that changes to the game may bring. One of the primary concerns of football fans regarding VAR is that it leads to frequent stoppages and reviews, which have historically been relatively foreign concepts to the fast-paced game of football. Perhaps anticipating these concerns, the IFAB proposed rules that would have the effect of limiting VAR interventions. Despite these rules,

many fans continue to complain about the frequent interventions of VAR, and note that they cannot even properly cheer after their teams score a goal, out of fear that it may be cancelled after VAR review. Indeed, some fans note that the VAR ought to 'leave subjective decisions to on-pitch referee' (Carragher 2023).

Here, inspired by prior law and economics analyses of standards of proof, I constructed an economic model to study when VAR ought to alert the referee, with the goal of balancing two objectives, namely reducing referee errors and limiting review-based stoppages. My analysis revealed that, in order to alert the referee, the VAR ought to have a belief that the referee's decision is incorrect with a probability that exceeds a threshold that is above 50 %. Absent further restrictive assumptions, it is impossible to pin point the exact threshold probability that the VAR ought to use as a criterion. However, this threshold being above 50 % is consistent with the standard proposed by IFAB, namely the presence of a "clear and obvious error". If this interpretation is correct, then the VAR may use the 50 % figure as a simple and conservative *filter* in making decisions; the referee ought never be alerted for review unless her decision is more likely erroneous than correct. This simple filter may act as a useful rule-of-thumb, and can also be used in the performance evaluations of VAR referees, given that some VAR decisions appear to fail even this very conservative test. I also questioned whether the VAR ought to be more or less hesitant to alert the referee, when an important event, such as a goal, occurs between the position in which the referee may have made a mistake and the stoppage in which the VAR review takes place. My analysis revealed that the VAR ought to alert the referee more liberally in such cases, because these intervening events increase the importance of decision errors relative to game stoppage costs.

I then conducted an analogous analysis of a hypothetical system which monitors and reviews the stopping decisions of a police officer in the processes of law enforcement. This analysis revealed that under plausible conditions the review system ought to interfere with a police officer's decision making only when the officer has committed an error with a large probability, due to the asymmetric costs associated with different types of enforcement mistakes. This suggests that the relative gains from implementing a review system in the law enforcement context may be smaller in the law enforcement context. Therefore, the implementation of such review systems in the law enforcement context may need to be delayed until their implementation costs fall below these relatively small benefits, which may provide a rationale for a conservative approach to adopting real-time review systems in law enforcement.

It is worth emphasizing that dynamics similar to those studied here emerge in other contexts where video reviews are used to guide the decisions of an actor. An example I noted is the military context, in which a soldier's environment is monitored by overseers through videos, and he is required to obtain their approval prior

to executing certain actions (e.g., sniper strikes). The relationship between the two parties in this example is different than the one between the referee and the VAR in the football context in some respects. The soldier must wait for prior approval to take an action whose viability may vanish while seeking approval, while the referee may take actions and later over-turn them. Thus, some errors in the military context generate irreversible harms, while error-costs in sports can largely be reversed through later reviews. Despite these differences, the optimality of military protocols that use video-reviews can be studied by extending the analysis provided herein.

References

- Ashdown, J. 2023. "How to Solve Everything with Eight Suggestions for Improving VAR." *The Guardian*. <https://www.theguardian.com/football/2023/oct/02/football-daily-email-var-tottenham-liverpool> (accessed February 13, 2024).
- Brown, B. 2024. "Jose Mourinho Explains Why VAR Worked Perfectly in AFCON." *NotJustOK*. <https://notjustok.com/sports/jose-mourinho-explains-why-var-worked-perfectly-in-afcon/> (accessed February 13, 2024).
- Carragher, J. 2023. "VAR is Ruining Football — We Must Limit its Influence Now." *The Telegraph*. <https://www.telegraph.co.uk/football/2023/11/10/jamie-carragher-var-limit-influence-ruining-football/> (accessed February 13, 2024).
- Demougin, D., and C. Fluet. 2005. "Deterrence Versus Judicial Error: A Comparative View of Standards of Proof." *Journal of Institutional and Theoretical Economics* 161: 193–206.
- Errekagorri, I., J. Castellano, I. Echeazarra, and C. Lago-Peñas. 2020. "The Effects of the Video Assistant Referee System (VAR) on the Playing Time, Technical-Tactical and Physical Performance in Elite Soccer." *International Journal of Performance Analysis in Sport* 20: 808–17.
- FIFA. 2024. "Video Assistant Referee (VAR)." *Inside FIFA*. <https://www.fifa.com/technical/football-technology/football-technologies-and-innovations-at-the-fifa-world-cup-2022/video-assistant-referee-var> (accessed February 13, 2024).
- Haberturk. 2023. "Özgüç Türkalp'in 'Ofsaytımsı' ifadesi gündem oldu." <https://www.haberturk.com/ozguc-turkalp-in-ofsaytimsi-ifadesi-gundem-oldu-3553015-spor> (accessed October 18, 2024).
- Hylton, K., and V. Khanna. 2007. "A Public Choice Theory of Criminal Procedure." *Supreme Court Economic Review* 15: 61–118.
- IFAB. 2024. "Video Assistant Referee (VAR) Protocol, IFAB." *Video Assistant Referee (VAR)*. <https://www.theifab.com/laws/latest/video-assistant-referee-var-protocol/#principles>.
- Kassouf, J. 2023. "VAR in the NWSL Lacks Investment to Fix Refereeing Problems." *ESPN*. https://www.espn.com/soccer/story/_/id/38566958/nwsl-lacks-cameras-investment-var-fix-refereeing (accessed February 13, 2024).
- Kube, C., and M. Gains. 2024. "Kabul Airport Bomber Was an ISIS Operative Freed from Prison by the Taliban." *NBC News*. <https://www.nbcnews.com/news/kabul-airport-bomber-was-isis-operative-freed-prison-taliban-rcna147810> (accessed October 18, 2024).
- Lago-Peñas, C., M. Gómez, and R. Pollard. 2021. "The Effect of the Video Assistant Referee on Referee's Decisions in the Spanish LaLiga." *International Journal of Sports Science & Coaching* 16: 824–9.

- Miceli, T. 1990. "Optimal Prosecution of Defendants Whose Guilt is Uncertain." *Journal of Law, Economics, and Organization* 6: 189–201.
- Miceli, T. 1991. "Optimal Criminal Procedure: Fairness and Deterrence." *International Review of Law and Economics* 11: 3–10.
- Mungan, M. 2011. "A Utilitarian Justification for Heightened Standards of Proof in Criminal Trials." *Journal of Institutional and Theoretical Economics* 167: 352–70.
- Mungan, M. 2018. "Optimal Preventive Law Enforcement and Stopping Standards." *American Law and Economics Review* 20: 289–317.
- Mungan, M. 2020. "Justifications, Excuses, and Affirmative Defenses." *Journal of Law, Economics, and Organization* 36: 343–77.
- Mungan, M., and A. Samuel. 2019. "Mimicking, Errors, and the Optimal Standard of Proof." *Economics Letters* 174: 18–21.
- Oliveira, M., V. Steffen, and F. Trojan. 2023. "A Systematic Review of the Literature on Video Assistant Referees in Soccer: Challenges and Opportunities in Sports Analytics." *Decision Analytics Journal* 7: 100232.
- Rizzolli, M., and M. Saraceno. 2013. "Better that Ten Guilty Persons Escape: Punishment Costs Explain the Standard of Evidence." *Public Choice* 155: 395–411.