Beyond Perception: Impact of Precision of Monitoring on Performance in Competitive Settings

Divyanshu Jain *
(Economics area, Indian Institute of Management Ahmedabad)

September 2024

Abstract

This paper studies how technology-enabled improvements in precision of monitoring impact performance in competitive settings. More precise monitoring can reduce uncertainty by providing accurate and reliable outcome measures, and facilitate human capital development using better data in training and feedback. Using a difference-in-differences framework and match-level data, I evaluate whether the adoption of Electronic Line Calling (ELC) technology in tennis tournaments after July 2020 impacted player performance. The main finding is that ELC adoption leads to an increase in aces and the point-winning probability of the server, and a reduction in serve faults. These results suggest that tech-enabled precision of monitoring improves performance in competitive settings. Furthermore, additional evidence of increased serve speeds and shorter rally lengths due to ELC adoption points to risk-taking behavior as a potential mechanism underlying the performance improvement. Lastly, the effects of more precise monitoring are found to be more pronounced in early-stage matches, lower-intensity rivalries, and among female competitors.

Keywords: *Monitoring; Technology adoption; Difference-in-differences; Sports economics*

JEL Classification: *C93; D91; L83; O33; Z22; Z28.*

*Email: phd21divyanshuj@iima.ac.in

1 Introduction

Monitoring individual actions is key for managerial oversight in a large number of settings, especially when rewards or punishments are based on the monitoring outcomes. Recent advancements in technology have enabled much greater precision of monitoring than before. For example, in the retail and logistics industry, Fujitsu uses a monitoring device that closely studies the posture of employees to determine whether employees are shirking or not (Marr, 2015). Similarly, Amazon uses wristbands on warehouse workers to monitor their precise location, hand movements, and speed of packaging (Yeginsu, 2018). Other examples include biometric attendance in government workplaces (Dhaliwal and Hanna, 2017), body cameras in policing (Zamoff et al., 2022), and real-time precise measurement of pollution levels (Axbard and Deng, 2024). Increasing use of technology implies that precise monitoring may be ubiquitous in the future (Holt et al., 2017). This paper studies the impact of technology-enabled improvements in the precision of monitoring on individual performance.

Studying this research question is challenging for several reasons. First, the introduction of monitoring technology may be endogenous or influenced by natural changes over time, making it difficult to establish causal effects. Second, precise data may be difficult to obtain for researchers in these settings. I overcome these challenges by exploiting a natural quasi-experimental setting in professional sports, and study the impact of tech-enabled precision of monitoring on performance in a competitive setting.

Several papers suggest the potential importance of precise monitoring. First, precise monitoring may strengthen the monitoring incentives due to increased scrutiny of actions (Parsons et al., 2011). Second, precise monitoring reduces uncertainty (or margin of error) of the monitored outcomes, which increases the accuracy, reliability, and perceived fairness of the performance measures (Mills, 2017; Ahmed et al., 2022; Abernethy et al., 2023). Finally, precision of monitoring may impact individuals through the human capital development process, as high-quality data generated by the monitoring process can be incorporated in future training, evaluation, and feedback mechanisms (Taylor and Tyler, 2012; Rockoff et al., 2012).

Precision of monitoring may be related to performance through the following mechanisms. The increased reliability and perceived fairness of performance measures may make monitored individuals more confident that they will not be erroneously penalized (Giacosa et al., 2023). Moreover, the reduced uncertainty around monitored outcomes may enable the individuals to put more cognitive effort, i.e., focus, on their main task, thereby improving performance. On the other hand, individuals may become overcautious and focus too much on avoiding mistakes at all times due to increased scrutiny of their actions, possibly hampering performance (Siegel et al., 2022).

The context of this study is the Electronic Line Calling (ELC) technology in tennis. Using multiple cameras situated around the playing arena, ELC precisely traces a ball's trajectory and bounce. Till July 2020, ELC was used only to recheck line calls made by human umpires. However, post-July 2020, some tournaments completely adopted ELC to make line calls, while others continued employing human umpires to make the decisions. I leverage this difference in the adoption of ELC by various tournaments to study its impact on performance, using a difference-in-differences (DiD) framework.

The first difference is between tournaments that completely adopted ELC (referred as "with-ELC" group going forward) versus tournaments that continued with the status quo of human line umpires (referred as "without- ELC" group going forward). The second difference is between the post-July 2020 and the pre-July 2020 time period. ELC adoption was homogeneous, one-shot, and in an absorbing state. This implies that all "with-ELC" tournaments happening after July 2020 used the same ELC technology and never stopped using it. Finally, for my analysis to be valid, the key outcome measures must evolve in the same way for both the “with-ELC” and the “without-ELC” tournaments before July 2020 in the pre-period. In the Empirical Analysis section, I show that this holds, thus satisfying the key identification assumption of parallel trends.

Using a professional sports context, such as tennis, to study this research question offers many advantages. First, the rules and individuals’ actions are well-defined and easily observable (Kahn, 2000). Data from these settings contain large amounts of detailed information that might be unobserved in other contexts, enabling robust empirical estimations and heterogeneity analysis based on gender, ability, experience, etc. (Bar-Eli et al., 2020). Further, the subjects of analysis are high-ability individuals in high-stakes competitive environments, which helps in applying these insights in consequential real-life situations (Balafoutas et al., 2019). Finally, the choice of this particular context to study the impact of tech-enabled precision of monitoring on individual performance is supported by the anecdotal evidence on ELC by top tennis players (see Appendix A.1), as their quotes highlight the importance of mechanisms like enhanced focus on the main task and increased accuracy, perceived fairness, and reliability of performance measures.

The central finding of this paper is that tech-enabled precision of monitoring increases the aces and the point-winning probability of the server, and reduces the serve faults. These results indicate that tech-enabled precision of monitoring improves performance in competitive settings, potentially driven by the reduced uncertainty, and increased reliability and perceived fairness of the monitoring system. Further, individuals may put more cognitive effort in the main task rather than worry about the uncertain monitoring outcomes. This is reflected in the increase in the average speed of serve and a reduction in the average length of rallies due to ELC, indicating increased risk-taking behavior. Finally, I also find that the effect of ELC is stronger in cases of early stages of competition, low-intensity rivalry, and among women participants.

This study contributes to the monitoring literature by providing robust evidence on the effect of precision of monitoring on performance in a competitive setting. By advancing the understanding of the role of precise monitoring, it particularly contributes to the literature on the use of technology in monitoring (Axbard and Deng, 2024; Debnath et al., 2023; Muralidharan et al., 2021; Berry et al., 2021; Dhaliwal and Hanna, 2017; Duflo et al., 2012). The findings suggest that precision of monitoring is important for the monitored individuals in improving performance in competitive settings.

The study also adds to the sports monitoring research where increasing use of technology has potentially influenced many stakeholders, including players, coaches, referees, and even the spectators (Almog et al., 2024; Abernethy et al., 2023; Mills, 2017; Parsons et al., 2011). Finally, the study is part of a growing body of research that uses tennis data to study important concepts, including momentum effects (Gauriot and Page, 2019), loss aversion (Anbarci et al., 2018), and applicability of mixed strategy predictions in real-life (Gauriot et al., 2023). This study utilizes rich and unique tennis data to empirically estimate the impact of increased precision of monitoring (by analyzing ELC), which may be useful not only to sports stakeholders but also to managers and policymakers overseeing experienced and high-ability individuals operating in a high-stakes environment.

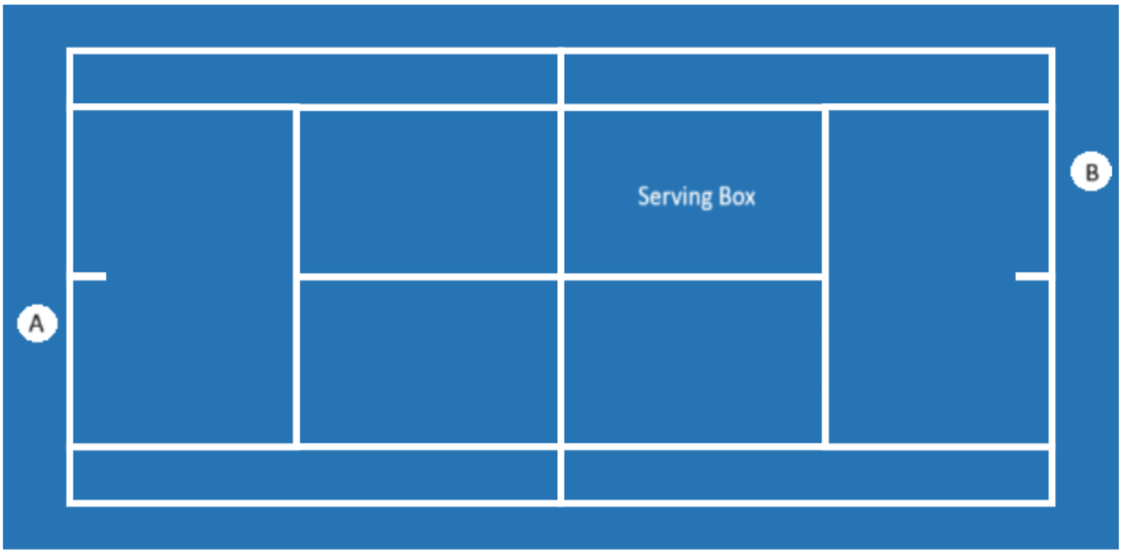
2 Context and Research design

2.1 Background

Electronic Line Calling (ELC) technology in tennis, primarily provided by the company Hawk-Eye Innovations, precisely traces a ball's trajectory and bounce using multiple cameras situated around the playing arena (or tennis court).[1] Figure 1 depicts a top-down view of a typical tennis court. In a one-on-one (singles) match, players stand on opposite ends and hit the ball toward each other till the time one of them is not able to return the ball in a legitimate manner. At this moment, the

[1] More information about Hawk-Eye Innovations is available at their website: https://www.hawkeyeinnovations.com/

Figure 1: Tennis court



Notes: The figure depicts a top-down view of a tennis court of a match between players A and B. The meaning and purpose of the labels are explained later in the main text. The source of the image is the Australian Open 2024 website. The labels are made by the author.

other player wins a point. To win the match, a series of such points is accumulated to win games and then sets.

ELC captures whether a ball was played inside the relevant court lines or outside. In tennis, a player serving a ball outside the court line is considered invalid, and the player incurs a penalty. The precision of ELC is extremely high, with a mean prediction error of just 3.6 mm. For comparison, court lines are 50 mm wide, and a tennis ball is approximately 67 mm in diameter (Gauriot and Page, 2019; Almog et al., 2024).

Till 2006, only human line umpires had the sole responsibility of deciding whether a ball bounced inside or outside the line (in other words, making the line calls). However, from the 2006-07 season, the apex tennis authorities, including the Association of Tennis Professionals (ATP), started using ELC to allow players to challenge the line calls by human umpires.[2] If challenged, the decision of ELC was

[2]The ELC review system was majorly used in hard court and grass court tournaments and not in clay court tournaments. On clay courts, the ball usually leaves a visible mark on the surface. The umpires generally rely on those marks in case of any dispute over a line call.

accepted.[3]

2.2 Research design

Till July 2020, ELC was used only to recheck the line calls made by human umpires if players challenged them. After July 2020, some tournaments, such as the Australian Open, adopted ELC to make line calls (replacing the human line umpires), while others, such as Wimbledon, continued employing human umpires to make the decisions and used ELC only for challenges. This complete adoption of ELC to make line calls by tournaments in July 2020 is the main "treatment" in this study which captures the tech-enabled precision of monitoring. The adoption of ELC increases the precision of monitoring. How this increased precision of monitoring impacts the performance of the server is the subject of analysis going forward.

To answer the research question, I use a standard DiD methodology. The "with-ELC" tournaments are the treatment group, and the "without-ELC" tournaments form the control group. The pre-period is the time duration before July 2020 (which is the date of ELC adoption), and the post-period is the time duration after July 2020. This setting satisfies the standard assumptions of the DiD framework. For the "with-ELC" group, the ELC technology was the same all across the tournaments, thus maintaining its homogeneity. Moreover, all "with-ELC" tournaments happening after July 2020 used ELC and never stopped using it. Further, no other major rule change occurred in tennis at the same time as ELC was adopted.

A review of various news portals (such as The New York Times, ESPN, and The Telegraph) and ATP press releases suggests that the adoption of ELC was driven by reasons unrelated to playing behavior. First, the adoption of ELC was expedited in some tournaments by COVID-related restrictions and social distancing norms (Clarey, 2020; Block, 2022). More importantly, the main source of difference in the adoption of ELC was the trade-off between reducing errors and keeping the jobs of

[3] On an average, 12% of the line calls by human line umpires were found to be incorrect (Almog et al., 2024).

human line umpires (Clarey, 2020; Briggs, 2023). The presence and involvement of line umpires on the court is considered to be an important part of the viewing experience and tradition of tennis tournaments.[4] While some tournaments decided that reducing errors on line calls is more important than the jobs of human line umpires, others were concerned about job loss (and its consequences) due to replacing human umpires with ELC (Clarey, 2020; Briggs, 2023). To substantiate these claims with data, in the Empirical Analysis section, I show that playing behavior (in the pre-period) did not predict ELC adoption, but certain tournament characteristics (such as level and surface type) did. Accordingly, I control for all time-invariant tournament characteristics throughout the analysis.

2.3 Key outcome measures

Most of the key outcome measures to study performance in tennis matches come from serve-level information. A serve marks the beginning of each point with players alternating serving each game. To visualize, refer to Figure 1, where player A is the server at a particular point serving to player B standing at opposite ends of the court. The line in the middle represents the net in the court. The server (player A) is required to serve the ball within the serving box. If the first serve does not land correctly, it is considered a fault, and the server is given another chance to serve (known as a second serve). If even the second serve does not land correctly, then the point is awarded to the opponent (known as a double fault). I primarily focus on serve-level information for player behavior since it is the only action in the match that is in complete control of the player serving the ball (also argued by Anbarci et al. (2018)). The shots played after the serve are dependent on it.

On average, players win around 60-70% of the points when they serve (Cohen-

[4] As articulated in the ATP press release in April 2023 (https://www.atptour.com/en/news/electronic-line-calling-release-april-2023) where they also announce that ELC will be used across all its tours from 2025 onwards. This suggests that the without-ELC tournaments are not structurally different from the with-ELC tournaments as they all will have ELC in the future, and only the timing of adoption will differ.

Zada et al. (2017); also see Table 1). Considering the level of control on the serve and the potential advantage to the server, Cohen-Zada et al. (2017) suggest that a server losing a point is a good proxy for low performance. Extending this argument, I consider serving an ace (winning the point on serve because the receiver fails to touch it) as an indicator of good performance and incurring a serve fault as an indicator of low performance. Thus, aces, serve faults, and points won by the server are used to operationalize performance in this context.

Researchers can also gain insight into the performance of tennis players by observing the risk-taking behavior using the speed and the direction of a serve (Klaassen and Magnus, 2009). A fast-paced serve is considered to be riskier as the player has less control compared to a slower serve. Further, serving close to the permitted court lines also implies higher risk as the chances of the ball landing outside the line are higher as compared to serving around the middle of the serving box. If players are serving more risky serves, then a point may finish quickly (i.e., shorter rally length) because the opponent is less likely to return properly or the player is more likely to make an error (Dona et al., 2024). So, observing shorter rallies would also imply that risk-taking was high on that serve. Thus, speed of serve, direction of serve (whether landing close to the line or not), and rally length are used as a proxy for risk-taking behavior in this context.

If players are able to successfully serve faster and farther, and play shorter rallies, then it would be difficult for the opponent to win the point (Ely et al., 2017). Therefore, observing an improvement in performance indicators with a simultaneous increase in risk-taking behavior might suggest that risk-taking behavior is one of the potential drivers of the improvement in performance. Therefore, in additional analysis, I study the impact of ELC on risk-taking behavior to explore this potential mechanism.

3 Data and Empirical strategy

Data in this study corresponds to the match-level information from Grand Slams, ATP tours, and Women's Tennis Association (WTA) tours.[5] I have collected data from August 2011 to July 2023 and defined each year-ending July (i.e., August-July) as the unit of time. The unit of observation is match-level variables on the key outcome measures (described in the previous section). As each match may have a different number of serves played in it, the outcome variables are standardized on a per-serve per-match basis. This data is supplemented with player information on gender, rankings, and age from ATP and WTA websites. This will help in exploring whether the effect of precision of monitoring varies based on the individual's gender, ability (proxied by ranking), and age.

Table 1 contains the sample description of the dataset, which includes all singles matches from Grand Slams, ATP, and WTA tours conducted between August 2011 and July 2023. It comprises 58093 matches across various levels of tournaments (Grand Slams, ATP1000, ATP500, and so on) and from various stages of the tournament (final, semi-final, and so on). The sample is evenly distributed across male and female players. Of the 197 tournaments, 40 form the "with-ELC" group and 157 form the "without-ELC" group. The lower panel of the table indicates the typical values of the key outcome measures per match till July 2020. For example, on average, 6 percent of the serves are aces, 31 percent of the serves are faulty, and around 60 percent of the points are typically won by the server.

Information on other variables, including speed of serve, serve direction [6], and rally length, is not consistently available for all matches across the pre- and post-

[5] This is sourced from Jeff Sackmann's Github repository (https://github.com/JeffSackmann). This data has been used by related literature to study player behavior in tennis (Dona et al., 2024). Outside of the official data with tennis authorities, broadcasters, and Hawk-Eye, no such publicly available data on tennis exists. According to the website, data in this repository is relied upon even by the coaching staff of some top players. To check the veracity of this source, I randomly selected 2 matches from each Grand Slam tournament in the sample and watched their replays from the tournament websites. Furthermore, I watched highlights (on YouTube) of approximately 50 matches from ATP and WTA tournaments.

[6] Serve direction takes value 1 if the serve is near the court lines, 0 otherwise.

Table 1: Sample description

Sample Description	Total
Players	
Men	853
Women	854
Tournaments	
Grand Slams	4
ATP 1000 or WTA 1000	12
ATP 500 or WTA 500	33
ATP 250 or WTA 250	130
Others	18
All	197
Matches	
Final	1398
Semi-final	2794
Quarter-final	5480
Round of 16	10940
Others	37481
All	58093

Outcome variables (per serve per match)	**Mean (SD)**
Aces	0.06
	(0.04)
Fault on serve	0.31
	(0.04)
Point winning probability of server	0.60
	(0.06)

Notes: In the upper panel, the cells contain the number of observations in each respective category. In the lower panel, each cell displays the mean and standard deviation of the outcome variable of interest.

period. Therefore, I do not use these variables in the main analysis. However, I discuss this subset and associated analysis later in the *Additional analysis and robustness checks* section.

Next, I show that the motivation behind ELC adoption was unrelated to playing behavior. To test this, I regress the dummy variable for ELC adoption on the pre-period averages of the key outcome measures along with tournament characteristics, including their level, surface type, and location. As shown in Table 4 in Appendix A.2, the coefficients of all three outcome variables are statistically insignificant at the 95 percent level.

However, ELC adoption is related to certain tournament-level characteristics. First, the higher level tournaments (such as ATP 1000 and Grand Slams) are more likely to adopt ELC as compared to the lower level tournaments (such as ATP250). This is reasonable as ELC is an expensive technology, and it might be difficult to eliminate the role of line umpires from lower-level tournaments as they help in organizing the tournaments (Michaels, 2021). Second, relative to hard court tournaments, ELC is less likely to be adopted by clay court tournaments. This is driven by the nature of clay courts, where the ball marks are visible on the court, and therefore officials may not feel the need to replace human umpires with ELC. Given that these characteristics do not change over time, the DiD framework should take care of these differences. Nevertheless, I include tournament-level fixed effects throughout the analysis to control for any potential time-invariant tournament-level unobserved characteristics. Similarly, I include year-fixed effects to control for any potential tournament-invariant year-level unobserved characteristics.

The impact of ELC on performance is estimated using the following empirical specification representing a DiD analysis.

$$
\begin{aligned}
y_{mct} = \beta_0 + \beta_1(ELC \times Post)_{ct} + \beta_2 StageT_{mct} + \beta_3 DiffRanking_{mct} \\
+ Players_m FE + Competition_c FE + Year_t FE + \epsilon_{mct} \quad (1)
\end{aligned}
$$

In this specification, y_{mct} is the outcome variable in match m in tournament/competition c in the year t. ELC_C variable takes value 1 if the competition c belongs to the with-ELC group, and 0 if it belongs to the without-ELC group. $Post_t$ variable takes value 1 if the match is happening after July 2020, i.e., in the post period and 0 if it is happening before July 2020. β_1 is the key coefficient of interest representing the effect of ELC adoption on the outcome variable. $StageT_{mct}$ variable controls for whether the match m was a final or semi-final, and so on. The $DiffRanking_{mct}$ captures the difference in the ranking of the two players playing match m. This specification also controls for player level, competition level, and year-level fixed effects. The error term is clustered at the competition-year level. This equation is estimated for each of the outcome variables – aces, point winning probability of server, and serve faults.

Before moving on to estimating the effect of ELC on performance, the key identification assumption of parallel trends must be satisfied. To test this, I estimate a modified version of equation (1).

$$y_{mct} = \beta_0 + \sum_{t=2012}^{2023} \beta_1 Year_t + \sum_{t=2012}^{2023} \beta_{2,t} (ELC \times Year)_{ct} + \beta_3 StageT_{mct}$$
$$+ \beta_4 DiffRanking_{mct} + Players_m FE + Competition_c FE + Year_t FE + \epsilon_{mct} \quad (2)$$

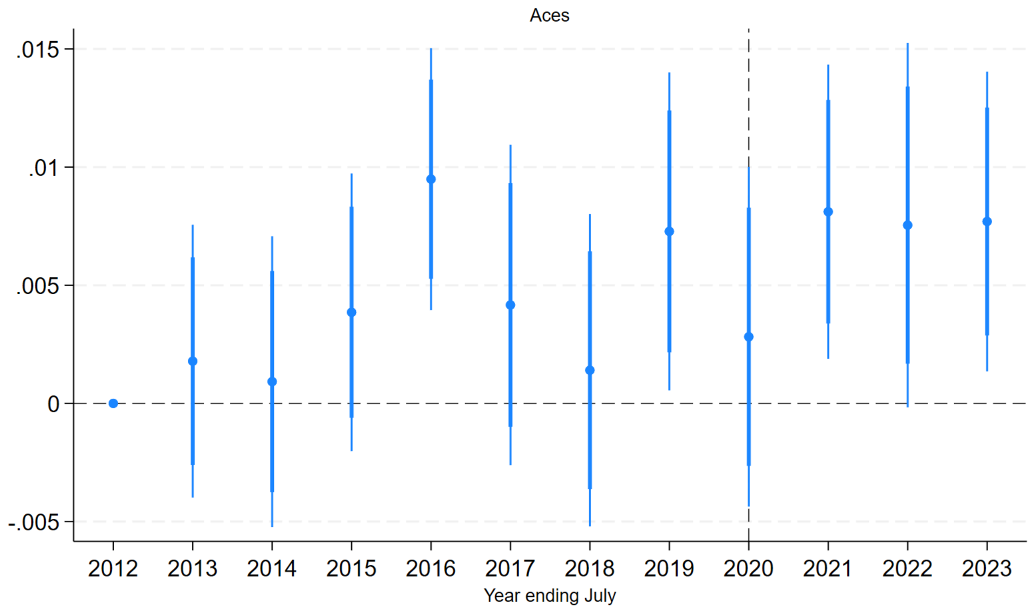
This specification is the same as equation (1) except that the $Post_t$ variable is replaced by a dummy variable for each year separately. Therefore, this gives a separate interaction term coefficient for each year, enabling a year-wise comparison of the respective $\beta_{2,t}$ coefficients. If the coefficients for the pre-period years are not statistically different from the null (with the base being the year ending July 2012), then one may assume that in the absence of the "treatment," the outcome variable for the two groups would have evolved in a similar manner. Given this, any change observed in the post-period can then be attributed to the "treatment", thereby providing the treatment effect.

4 Empirical Analysis

4.1 Main results

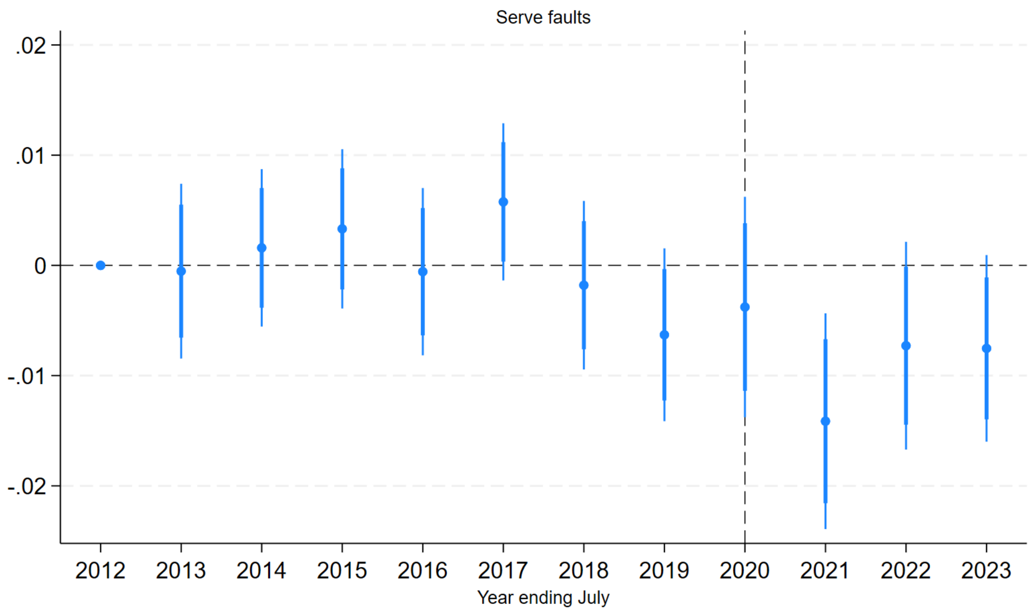
The results of equation (2) for each of the three outcome measures are presented in figures 2-4, respectively, and the results of equation (1) are shown in Table 2.

Figure 2: Impact of ELC on aces



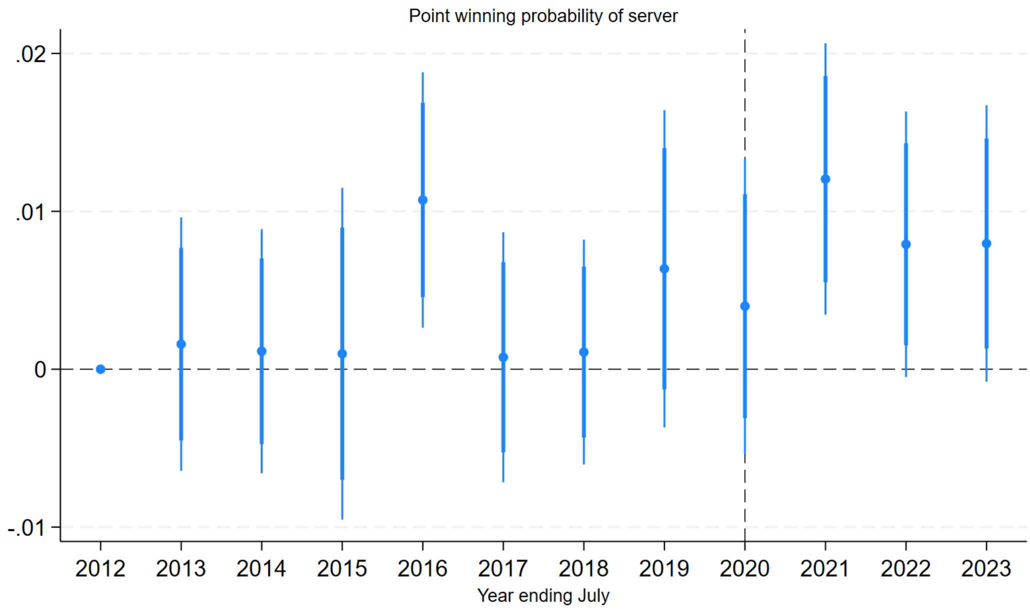
Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise aces (per serve) as the dependent variable. The null is defined with the year ending July 2012 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

Figure 3: Impact of ELC on serve faults



Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise serve faults (per serve) as the dependent variable. The null is defined with the year ending July 2012 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

Figure 4: Impact of ELC on point winning probability of the server



Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise point winning probability of the server as the dependent variable. The null is defined with the year ending July 2012 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

The figures reveal that the parallel trends assumption is largely satisfied for all three variables. The coefficients for pre-period years are statistically significant in only 3 instances out of a total of 27 estimated coefficients, which might simply be a case of random chance. Figures 2, 3, and 4 show that aces increase, serve faults decrease, and point-winning probability of the server increases due to ELC adoption, respectively.

To get precise estimates of the impact of ELC, the results for equation (1) are depicted in Table 2. In particular, the number of aces per serve increases by 0.004 (or a 6.67 percent increase), the point-winning probability of the server also increases by 0.6 percentage points (or a 1 percent increase), and the number of serve faults per serve declines by 0.009 (or a 3 percent decrease). All these effects are significant at the 1 percent level. These findings suggest that tech-enabled precision of monitoring (in the form of ELC) improves performance in competitive settings.

Table 2: Impact of ELC on performance measures

	(1)	(2)	(3)
VARIABLES	**Aces per serve**	**Faults per serve**	**Point winning probability of server**
ELC X Post	0.004***	-0.009***	0.006***
	(0.001)	(0.002)	(0.002)
Stage of Tournament	0.000**	-0.004***	0.001***
	(0.000)	(0.000)	(0.000)
Difference in Ranking	-0.000	0.000*	-0.000***
	(0.000)	(0.000)	(0.000)
Players FE	Yes	Yes	Yes
Tournament FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	0.061***	0.321***	0.599***
	(0.001)	(0.001)	(0.001)
Control group mean	0.06	0.31	0.60
Control group SD	(0.04)	(0.04)	(0.06)
Observations	55,372	55,372	55,372
Adjusted R-squared	0.635	0.310	0.502

Notes: Robust standard errors in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

4.2 Additional analysis and robustness checks

As discussed earlier, to get more insight into the impact of ELC on playing behavior, one can look at variables such as speed of serve, serve direction, and rally lengths. However, data on these variables is only available from August 2014 onwards and that too only for a few matches. For example, data on the speed of serves is available for all US Open and Wimbledon men's matches and only for a subset of other Grand Slam matches (4987 matches in total). Similarly, for serve direction and rally length, data is available for only 6336 and 9156 matches, respectively. The inconsistency in data may be because the source of these data is volunteer-coded information from match recordings and is therefore skewed in favor of popular players and later stages of tournaments (Dona et al., 2024).

While these variables do not form part of the main analysis for the reasons highlighted above, I use them to find additional evidence on whether ELC causes an increase in risk-taking behavior and whether the main results are robust to changes in sample composition. Table 3 shows the results of equation (1) with average speed of serve, serve direction, and average rally length as three additional dependent variables.[7] The coefficients indicate that the average speed of serve increases by 1.46 kmph (a 0.9 percent increase), the average length of rally declines by 0.7 shots (a 18 percent decrease), and the serve direction remains unchanged due to ELC. Taken together, these indicate that ELC adoption leads to an increase in risk-taking behavior.

Furthermore, these results provide a robustness check for the impact of ELC on the main performance indicators. Even with a much smaller and inconsistent sample, the coefficients indicate that ELC adoption leads to an increase in aces and point-winning probability. While the coefficient for serve faults is not statistically significant at the 90 percent level (it becomes significant at the 89 percent level), directionally it is in line with the main result that serve faults are reducing due to ELC. Combined with a simultaneous improvement in performance (as shown above in the main result), these results provide suggestive evidence that tech-enabled precision of monitoring improves performance, potentially driven by an increase in successful risk-taking behavior.

Finally, since the clay court tournaments do not adopt ELC at all, none of these tournaments are part of the "with-ELC group." While these are appropriate as a control group, one may be concerned about the comparability of these tournaments with the "with-ELC group." Therefore, as a robustness test, I rerun equation (1), excluding all clay court tournaments. The results of this robustness check are presented in Table 5 in Appendix A.4. It shows similar results to the main re-

[7] The results for equation (2) with average speed of serve, serve direction, and average rally length as the dependent variables are shown in figures 8, 9, and 10 in Appendix A.3

Table 3: Impact of ELC on the risk-taking and performance measures (smaller sample)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Avg rally length**	**Serve direction**	**Avg speed of serve**	**Aces per serve**	**Faults per serve**	**Point winning probability of server**
ELC X Post	-0.704***	0.019	1.461***	0.004*	-0.009	0.007*
	(0.220)	(0.018)	(0.515)	(0.002)	(0.005)	(0.004)
Stage of Tournament	0.112***	0.003***	0.101*	-0.000	-0.003***	0.000
	(0.011)	(0.001)	(0.050)	(0.000)	(0.000)	(0.000)
Difference in Ranking	0.000**	-0.000*	-0.001	0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Players FE	Yes	Yes	Yes	Yes	Yes	Yes
Tournament FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.499***	0.644***	159.919***	0.066***	0.300***	0.629***
	(0.112)	(0.010)	(0.342)	(0.002)	(0.003)	(0.003)
Control group mean	3.908	0.672	160.37	0.068	0.284	0.630
Control group SD	(1.134)	(0.135)	(13.82)	(0.045)	(0.033)	(0.063)
Observations	8,130	5,410	4,386	9,355	9,355	9,355
Adjusted R-squared	0.469	0.649	0.937	0.659	0.313	0.550

Notes: Robust standard errors in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

sult obtained in Table 2, which suggests that the impact of ELC on the outcome measures remains consistent irrespective of the inclusion/exclusion of clay court tournaments from the sample.

4.3 Heterogeneity analysis

4.3.1 Gender differences

In this section, I explore whether the impact of ELC varies based on the gender of the players. For this purpose, I estimate equation (1) separately for the men and women subsamples. Figure 5 panel (A) shows the results for men and panel (B) shows the results for women. ELC improves performance for both men and women, but the magnitude of improvement is much greater for women as compared to men. For example, the improvement in aces is 10 percent for women as compared to 6 percent for men. Similarly, the improvement in the point-winning probability is 1.4 percent for women as compared to 0.8 percent for men. The magnitude of reduction in serve faults is similar for both men and women. These results indicate that while precision of monitoring is important for everyone, women respond more to the reduced uncertainty and increased reliability of performance measures. A similar pattern is visible for risk-taking behavior, where women increase their risk-taking behavior much more as compared to men. This finding is notable given the existing research highlighting that women are much more risk averse as compared to men (Charness and Gneezy, 2012; Böheim et al., 2022). Finally, this result also supports the potential mechanism of increased risk-taking leading to improved performance.

Figure 5: Gender differences in the impact of ELC

Variables	Aces per serve	Faults per serve	Point winning probability of server	Avg speed of serve	Serve direction	Avg rally length
			Panel A: Men			
ELC X Post	0.005***	-0.009***	0.005***	0.938	0.010	-0.639***
	(0.002)	(0.002)	(0.002)	(0.584)	(0.019)	(0.209)
Control group mean	0.082	0.307	0.634	172.05	0.712	3.843
Control group SD	(0.046)	(0.038)	(0.055)	7.66	0.126	1.119
Observations	29,467	29,467	29,467	2,236	3,073	4,489
Adjusted R-squared	0.586	0.293	0.358	0.799	0.636	0.510
			Panel B: Women			
ELC X Post	0.004***	-0.009***	0.008***	1.974***	0.034*	-0.819***
	(0.002)	(0.003)	(0.002)	(0.536)	(0.018)	(0.246)
Control group mean	0.041	0.310	0.564	148.43	0.616	3.990
Control group SD	(0.030)	(0.043)	(0.053)	6.65	0.127	1.147
Observations	25,905	25,905	25,905	2,150	2,325	3,626
Adjusted R-squared	0.449	0.334	0.231	0.732	0.565	0.421

Notes: This table reports the output of equation (1) separately for the men and women subsamples. The fixed effects and other control variables are not shown here for brevity.

4.3.2 Stage of tournament

In advanced stages of the competition, the level of performance and risk-taking may be higher as compared to that in the initial stages of the competition. This might be because of higher stakes and motivation in the top stages of competition. For example, in quarter-finals and beyond, the performance indicators and risk-taking measures (except rally length) are higher than those in round-of-16 and lower matches. Therefore, in Figure 6, I explore whether the impact of ELC varies based on the stage of tournament. The results indicate that while ELC leads to better performance and increased risk-taking in matches prior to quarter-finals, performance shows smaller improvements and risk-taking behavior shows no significant change in quarter-finals and beyond. This suggests that the level of focus and motivation in the advanced stages of competition may already be high, leaving limited room for precise monitoring to drive further performance gains.

Figure 6: Impact of ELC: Heterogeneity by stage of tournament

Variables	Aces per serve	Faults per serve	Point winning probability of server	Avg speed of serve	Serve direction	Avg rally length
Panel A: Quarter finals and beyond						
ELC X Post	0.005**	-0.009***	0.003	1.335	-0.008	0.065
	(0.002)	(0.003)	(0.003)	(1.204)	(0.020)	(0.101)
Control group mean	0.067	0.300	0.608	163.20	0.736	4.309
Control group SD	0.047	0.039	0.065	13.41	0.118	0.984
Observations	8,985	8,985	8,985	274	1,027	1,186
Adjusted R-squared	0.659	0.305	0.520	0.958	0.575	0.709
Panel B: Prior to Quarter finals						
ELC X Post	0.004**	-0.009***	0.006***	1.579***	0.017	-0.849***
	(0.001)	(0.002)	(0.002)	(0.549)	(0.020)	(0.253)
Control group mean	0.062	0.310	0.600	160.12	0.649	3.821
Control group SD	0.044	0.040	0.064	13.83	0.134	1.145
Observations	46,107	46,107	46,107	4,016	4,095	6,672
Adjusted R-squared	0.628	0.301	0.499	0.935	0.645	0.440

Notes: This table reports the output of equation (1) separately for the quarter finals and beyond subsample and the prior to quarter finals subsample. The fixed effects and other control variables are not shown here for brevity.

4.3.3 Rivalry

Similar to the above argument, performance and risk-taking may be higher in cases of high-intensity rivalry between the competitors as compared to that in low rivalry competitions. This might be due to a higher level of motivation when playing against a close competitor. In the context of this study, I define high-rivalry competitions as those matches where the difference in ranking of the two players is less than 15. In such cases, the performance indicators and risk-taking measures (except rally length) are higher than those in low-rivalry competitions. Therefore, in Figure 7, I explore whether the impact of ELC varies based on the intensity of rivalry between the competitors. The results indicate that while ELC leads to better performance and more risk in both cases, the magnitude of impact is slightly greater in low-rivalry matches. This might be due to the fact that the level of focus and motivation in a high-rivalry competition may already be strong enough to not leave much scope for improvement for precise monitoring to have a major impact.

Figure 7: Impact of ELC: Heterogeneity by intensity of rivalry

Variables	Aces per serve	Faults per serve	Point winning probability of server	Avg speed of serve	Serve direction	Avg rally length
			Panel A: High-intensity rivalry			
ELC X Post	0.004**	-0.012***	0.005*	1.141	0.014	-0.293**
	(0.002)	(0.002)	(0.003)	(0.941)	(0.020)	(0.139)
Control group mean	0.064	0.306	0.604	162.88	0.711	4.012
Control group SD	0.044	0.040	0.063	13.74	0.122	1.154
Observations	10,893	10,893	10,893	551	1,014	1,487
Adjusted R-squared	0.620	0.323	0.499	0.949	0.606	0.638
			Panel B: Low-intensity rivalry			
ELC X Post	0.004***	-0.009***	0.006***	1.420**	0.021	-0.718***
	(0.001)	(0.002)	(0.002)	(0.561)	(0.019)	(0.225)
Control group mean	0.062	0.309	0.601	159.90	0.660	3.882
Control group SD	0.044	0.041	0.064	13.79	0.137	1.127
Observations	44,392	44,392	44,392	3,843	4,393	6,815
Adjusted R-squared	0.638	0.309	0.503	0.929	0.643	0.456

Notes: This table reports the output of equation (1) separately for the high and low intensity subsamples. The fixed effects and other control variables are not shown here for brevity.

5 Conclusion

This study contributes to the monitoring literature by providing robust evidence on the effect of tech-enabled precision of monitoring on performance in a competitive setting. The findings, along with anecdotal evidence, suggest that the reduced uncertainty and increased reliability of the monitoring system are important for the monitored individuals, and they perform better with an increase in the precision of monitoring. Notably, the benefits of increased precision of monitoring are more substantial in the early stages of competition, in matches characterized by lower intensity, and for female participants.

The study adds to the sports monitoring research where increasing use of technology has potentially influenced many stakeholders, including players, coaches, referees, and even the spectators (Almog et al., 2024; Abernethy et al., 2023; Mills, 2017; Parsons et al., 2011). While I focus on the impact of technology adoption on players in this paper, future research can analyze the impact on the behavior of other stakeholders as well.

This study utilizes a unique setting in tennis to empirically estimate the impact of increased precision of monitoring (by analyzing ELC), which may be useful not only to sports stakeholders but also to managers and policymakers overseeing experienced and high-ability individuals operating in high-stakes environments. However, as indicated by Finan et al. (2017), the incentive structure and the overall monitoring scheme of the context may be important to consider while assessing the role of tech-enabled precision of monitoring.

The analysis is limited in terms of the level of detail in the outcome variables. I use publicly available tennis data on standard match outcomes. Access to comprehensive and exhaustive data on all key outcome measures for all tournaments, especially speed of serve (and shots) and the exact distance of the ball from the court lines (captured by Hawk-Eye), can improve the rigour of the findings. Currently, the unit of analysis is match-level variables. In further analysis, I am aiming to conduct a player-level analysis and explore how the impact of tech-enabled precision of monitoring changes depending on the individual's ability and age.

A Appendix

A.1 Tennis player quotes on ELC

- **Naomi Osaka (former world no. 1):** "I don't mind it [ELC] at all. It saves me the trouble of attempting to challenge or thinking about 'Did they call it correctly or not?' It actually gets me really focused."
- **Dominic Thiem (rank 3 in 2020):** "No offense at all, but there are just no mistakes happening, and that's really good in my opinion. If the electronic call is out, the ball is out, so there's no room for mistakes."
- **Stefanos Tsitsipas (rank 3 in 2021):** "We must keep growing and adding new things to the sport that will help make it better and more fair."
- **Noah Rubin (former Wimbledon juniors champ):** [ELC] "takes away that anxiety of, 'I really hope the line judge or chair umpire doesn't mess this one up.'"

A.2 Relation of ELC adoption with pre-period characteristics

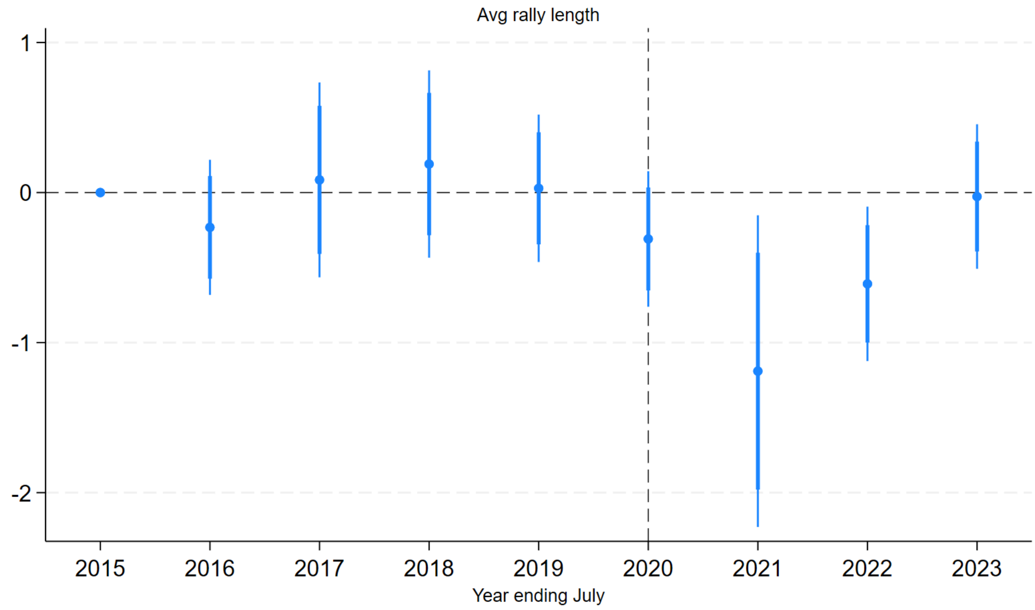
Table 4: Relation of ELC adoption with pre-period characteristics

Variables	**Dependent variable: ELC**
Aces per serve	-5.223
	(3.639)
Faults per serve	3.495
	(2.219)
Point winning probability of server	3.849
	(2.340)
ATP 1000	0.360***
	(0.099)
ATP 500	0.388***
	(0.086)
ATP Challenger	-0.139
	(0.330)
Grand Slams	0.408***
	(0.118)
Multi-country	0.678**
	(0.306)
WTA 1000	0.229
	(0.237)
WTA 125	0.039
	(0.120)
WTA 250	-0.004
	(0.079)
WTA 500	-0.230*
	(0.120)
WTA Finals	-0.096
	(0.375)
Carpet	-0.326
	(0.319)
Clay	-0.303***
	(0.090)
Grass	-0.065
	(0.135)
Country FE	Yes
Constant	-2.840*
	(1.583)
Observations	196
Adjusted R-squared	0.422

*Notes: The base variable for tournament level is ATP 250. The base variable for surface type is hard court. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$*

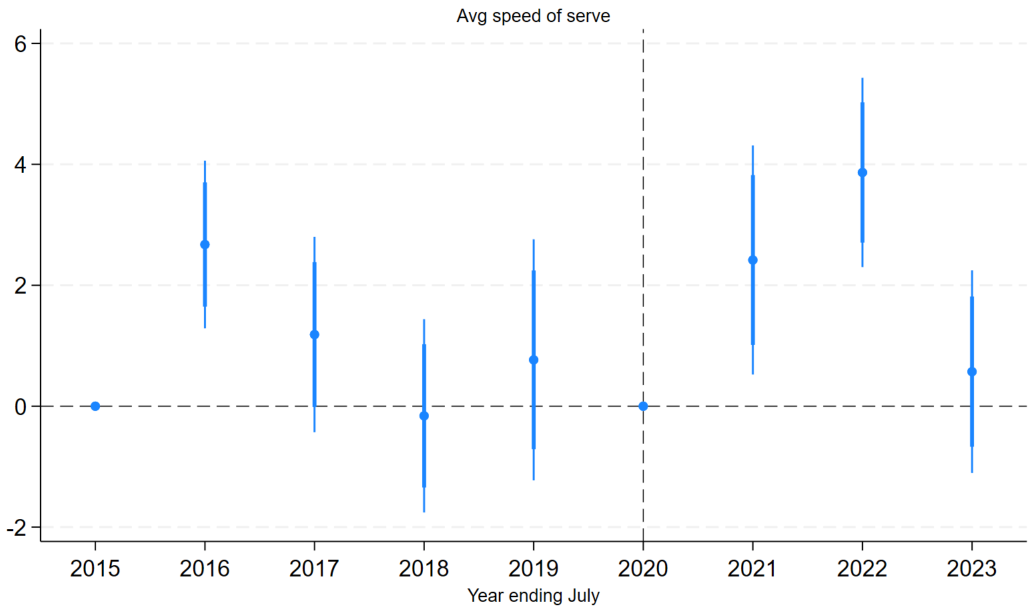
A.3 DiD results and parallel trends - small sample

Figure 8: Impact of ELC on average rally length



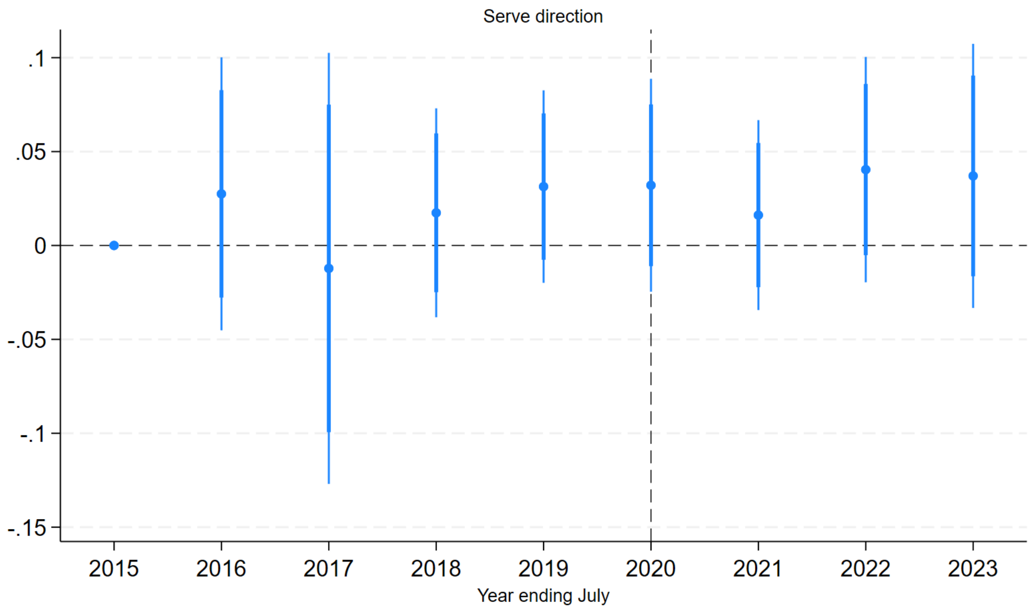
Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise average rally length as the dependent variable. The null is defined with the year ending July 2015 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

Figure 9: Impact of ELC on average speed of serve



Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise average speed of serve as the dependent variable. The null is defined with the year ending July 2015 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

Figure 10: Impact of ELC on serve direction



Notes: Plots the $\beta_{2,t}$ coefficient from equation (2) for each year separately with matchwise serve direction (per serve) as the dependent variable. The null is defined with the year ending July 2015 as the base year. The confidence intervals are drawn at the 95% and 99% levels.

A.4 DiD results - excluding clay court tournaments

Table 5: Robustness check of the main result (excluding clay court tournaments)

	(1)	(2)	(3)
VARIABLES	**Aces per serve**	**Faults per serve**	**Point winning probability of server**
ELC X Post	0.003**	-0.007***	0.003*
	(0.002)	(0.002)	(0.002)
Stage of Tournament	0.000**	-0.004***	0.001***
	(0.000)	(0.000)	(0.000)
Difference in Ranking	0.000	0.000	-0.000***
	(0.000)	(0.000)	(0.000)
Players FE	Yes	Yes	Yes
Tournament FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	0.067***	0.323***	0.603***
	(0.001)	(0.001)	(0.001)
Control group mean	0.069	0.311	0.605
Control group SD	(0.046)	(0.040)	(0.065)
Observations	39,107	39,107	39,107
Adjusted R-squared	0.643	0.305	0.517

Notes: Robust standard errors in parentheses.
**** p<0.01, ** p<0.05, * p<0.1*

References

Abernethy, M., K. Chin, S. Nair, and N. Soderstrom (2023). Can technology-enabled advanced monitoring systems influence individual performance and team dynamics? *European Accounting Review 32*(3), 577–605.

Ahmed, S., Y. M. Kundi, and N. Dar (2022). Effects of electronic performance monitoring on employee work engagement: A multilevel investigation. In *Academy of Management Proceedings*, Volume 2022, pp. 12746. Academy of Management Briarcliff Manor, NY 10510.

Almog, D., R. Gauriot, L. Page, and D. Martin (2024). Ai oversight and human mistakes: Evidence from centre court. *arXiv preprint arXiv:2401.16754*.

Anbarci, N., K. P. Arin, T. Kuhlenkasper, and C. Zenker (2018). Revisiting loss aversion: Evidence from professional tennis. *Journal of Economic Behavior & Organization 153*, 1–18.

Axbard, S. and Z. Deng (2024). Informed enforcement: Lessons from pollution monitoring in china. *American Economic Journal: Applied Economics 16*(1), 213–252.

Balafoutas, L., S. M. Chowdhury, and H. Plessner (2019). Applications of sports data to study decision making.

Bar-Eli, M., A. Krumer, and E. Morgulev (2020). Ask not what economics can do for sports-ask what sports can do for economics.

Berry, J., S. Mehta, P. Mukherjee, H. Ruebeck, and G. K. Shastry (2021). Crowd-out in school-based health interventions: Evidence from india's midday meals program. *Journal of Public Economics 204*, 104552.

Block, M. (2022). At the u.s. open, line judges are out. automated calls are in. https://www.npr.org/2022/09/10/1121801484/us-open-tennis-human-line-judges-replaced. npr.

Böheim, R., M. Lackner, and W. Wagner (2022). Raising the bar: Causal evidence on gender differences in risk-taking from a natural experiment. *Journal of Sports Economics 23*(4), 460–478.

Briggs, S. (2023). Wimbledon facing battle to keep line judges. https://www.telegraph.co.uk/tennis/2023/04/28/wimbledon-facing-battle-to-keep-line-judges/. The Telegraph.

Charness, G. and U. Gneezy (2012). Strong evidence for gender differences in risk taking. *Journal of economic behavior & organization 83*(1), 50–58.

Clarey, C. (2020). Automated line calls will replace human judges at u.s. open. https://www.nytimes.com/2020/08/03/sports/tennis/us-open-hawkeye-line-judges.html. The New York Times.

Cohen-Zada, D., A. Krumer, M. Rosenboim, and O. M. Shapir (2017). Choking under pressure and gender: Evidence from professional tennis. *Journal of Economic Psychology 61*, 176–190.

Debnath, S., M. Nilayamgode, and S. Sekhri (2023). Information bypass: Using low-cost technological innovations to curb leakages in welfare programs. *Journal of Development Economics 164*, 103137.

Dhaliwal, I. and R. Hanna (2017). The devil is in the details: The successes and limitations of bureaucratic reform in india. *Journal of Development Economics 124*, 1–21.

Dona, N. E., P. S. Gill, and T. B. Swartz (2024). What does rally length tell us about player characteristics in tennis? *Journal of the Royal Statistical Society Series A: Statistics in Society*, qnae027.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American economic review 102*(4), 1241–1278.

Ely, J., R. Gauriot, and L. Page (2017). Do agents maximise? risk taking on first and second serves in tennis. *Journal of Economic Psychology 63*, 135–142.

Finan, F., B. A. Olken, and R. Pande (2017). The personnel economics of the developing state. *Handbook of economic field experiments 2*, 467–514.

Gauriot, R. and L. Page (2019). Does success breed success? a quasi-experiment on strategic momentum in dynamic contests. *The Economic Journal 129*(624), 3107–3136.

Gauriot, R., L. Page, and J. Wooders (2023). Expertise, gender, and equilibrium play. *Quantitative Economics 14*(3), 981–1020.

Giacosa, E., G. M. Alam, F. Culasso, and E. Crocco (2023). Stress-inducing or performance-enhancing? safety measure or cause of mistrust? the paradox of digital surveillance in the workplace. *Journal of Innovation & Knowledge 8*(2), 100357.

Holt, M., B. Lang, and S. G. Sutton (2017). Potential employees' ethical perceptions of active monitoring: The dark side of data analytics. *Journal of Information Systems 31*(2), 107–124.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of economic perspectives 14*(3), 75–94.

Klaassen, F. J. and J. R. Magnus (2009). The efficiency of top agents: An analysis through service strategy in tennis. *Journal of Econometrics 148*(1), 72–85.

Marr, B. (2015). The quantified workplace: Big data or big brother. `https://www.forbes.com/sites/bernardmarr/2015/05/11/the-nanny-state-meets-the-quantified-workplace/?sh=56d932a169fa`. Forbes.

Michaels, J. (2021). "there are just no mistakes happening": Hawk-eye live gains more support at australian open. `https://www.espn.in/tennis/story/_/id/30877297/hawk-eye-live-gains-more-support-australian-open`. ESPN.

Mills, B. M. (2017). Technological innovations in monitoring and evaluation: Evidence of performance impacts among major league baseball umpires. *Labour Economics 46*, 189–199.

Muralidharan, K., P. Niehaus, S. Sukhtankar, and J. Weaver (2021). Improving last-mile service delivery using phone-based monitoring. *American Economic Journal: Applied Economics 13*(2), 52–82.

Parsons, C. A., J. Sulaeman, M. C. Yates, and D. S. Hamermesh (2011). Strike three: Discrimination, incentives, and evaluation. *American economic review 101*(4), 1410–1435.

Rockoff, J. E., D. O. Staiger, T. J. Kane, and E. S. Taylor (2012). Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review 102*(7), 3184–3213.

Siegel, R., C. J. König, and V. Lazar (2022). The impact of electronic monitoring on employees' job satisfaction, stress, performance, and counterproductive work behavior: A meta-analysis. *Computers in Human Behavior Reports 8*, 100227.

Taylor, E. S. and J. H. Tyler (2012). The effect of evaluation on teacher performance. *American Economic Review 102*(7), 3628–3651.

Yeginsu, C. (2018). If workers slack off, the wristband will know. (and amazon has a patent for it.). `https://www.nytimes.com/2018/02/01/technology/amazon-wristband-tracking-privacy.html`. The New York Times.

Zamoff, M. E., B. N. Greenwood, and G. Burtch (2022). Who watches the watchmen: Evidence of the effect of body-worn cameras on new york city policing. *The journal of law, economics, and organization 38*(1), 161–195.