INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 942

Global Evidence for the Relevance of Irrelevant Events: International Soccer Games and Leader Approval

Kyosuke KIKUTA*
Yoshikuni ONO†

21 August, 2024

Abstract

Despite the intense debate over whether politically irrelevant events (e.g., football games) influence voter attitudes, the existing literature predominantly focuses on cases in the U.S. and a few developed countries, neglecting generalizability across other countries. Our study provides the first global evidence for the relevance of irrelevant events, analyzing the effects of international soccer games on leader approval in more than 100 countries from 2008 to 2023. Leveraging "double" natural experiments, comprising as-if random assignments of interview dates and soccer results conditional on betting odds, we find that lucky victories enhance public approval of national leaders by 7 percentage points, while unlucky losses lack compensating effects. The analysis of causal mechanisms suggests that soccer games serve as a benchmark for leader performance. Notably, these effects exhibit substantial cross-country variation, prompting us to consider contextual conditions across countries.

Keywords: Soccer, Football, Public opinion, Leader approval, Natural experiment

---

* Research Fellow, Africa Study Group, Area Studies Center, Institute of Developing Economies, JETRO (Kyosuke_Kikuta@ide.go.jp)

† Graduate Student. Osaka School of International Public Policy, Osaka University.

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Global Evidence for the Relevance of Irrelevant Events:

International Soccer Games and Leader Approval

(10,437 words)

Abstract

Despite the intense debate over whether politically irrelevant events (e.g., football games) influence voter attitudes, the existing literature predominantly focuses on cases in the U.S. and a few developed countries, neglecting generalizability across other countries. Our study provides the first global evidence for the relevance of irrelevant events, analyzing the effects of international soccer games on leader approval in more than 100 countries from 2008 to 2023. Leveraging "double" natural experiments, comprising as-if random assignments of interview dates and soccer results conditional on betting odds, we find that lucky victories enhance public approval of national leaders by 7 percentage points, while unlucky losses lack compensating effects. The analysis of causal mechanisms suggests that soccer games serve as a benchmark for leader performance. Notably, these effects exhibit substantial cross-country variation, prompting us to consider contextual conditions across countries.

Keywords: Soccer, Football, Public opinion, Leader approval, Natural experiment.

On February 10, 2015, two days after the spectacular victory of Ivory Coast in the final round of the Africa Cup of Nations, thousands of people flooded the streets of Abidjan and celebrated the national heroes including Yaya Toure. However, when Yaya passed the crowds riding on a parade car, a rather unexpected person sat next to him: President Alassane Ouattara. The president even wore Toure's gold medal, symbolizing his contributions. Arguably, the national achievement helped the president's reelection in the same year. As Adekaiyaoja (2024) states, "Ouattara, who has always supported the team, was front and centre during celebrations, eventually parlaying the mood of the country into winning re-election."

The case of Ivory Coast illustrates how seemingly irrelevant events, such as soccer games, shape political outcomes. Politically irrelevant events, such as shark attacks (Achen and Bartels 2017), college football (Busby, Druckman, and Fredendall 2016; Healy, Malhotra, and Mo 2010; M. K. Miller 2013), and UFO sightings (Kitamura 2022) have been proven to influence elections and public opinion in the U.S.; however, the validity of the evidence is still debated (Achen and Bartels 2018; Busby and Druckman 2018; Fowler and Hall 2016; Fowler and Montagnes 2015; Healy, Malhotra, and Mo 2015). Recent studies have extended the analyses, especially those regarding sports events, to different times (Graham et al. 2023b) and different countries such as Finland (Rapeli and Söderlund 2022), Ireland (Müller and Kneafsey 2023), and Spain (Alfano and Ercolano 2023), yielding mixed results (Fowler and Montagnes 2023b, 2023a; Graham et al. 2023a). However, despite the extensive analyses, the literature is still confined to the U.S. and a few developed countries, neglecting the generalizability across countries.

Our study fills this gap by analyzing the effects of international soccer games on leader approval in more than 100 countries over a 15-year period from 2008 to 2023. We leverage one of the most comprehensive public opinion surveys—Gallup World Poll (GWP)—and "double" natural experiments comprising as-if random assignments of interview dates and the

results of soccer games conditional on pre-game betting odds (Card and Dahl 2011). Importantly, although the performance of a national team may depend on government policies, we condition the soccer results on pre-game betting odds and thus analyze "lucky" victories and "unlucky" losses, which are beyond the control of a national government. Since such random outcomes are not indicative of a leader's performance, they should not alter rational voters' attitudes. This feature allows us to test voter rationality (Ashworth, Bueno de Mesquita, and Friedenberg 2018; Healy, Malhotra, and Mo 2010).

The results of our analysis indicate that lucky victories improve leader approval, while unlucky losses lack compensating effects. Victories increased the support for an incumbent leader by 7 percentage points—about a 14% increase from the sample average. Extensive validity tests, including robustness checks with 2,688 specifications of samples, measurements, and methods, indicate that this finding is unlikely to be a false positive. The effects are also large for particularly salient soccer games (e.g., games in popular tournaments, those with high viewership, and matches between historical rivals) and vary substantially across countries. Moreover, our analyses of causal mechanisms indicate that although victories improve people's evaluations of national leader performance, they do not affect local mayoral performance. This indicates that respondents use the results of international soccer games as an irrational yet sensible heuristic for assessing national leadership. We also find suggestive evidence that international soccer games alter the actual results of elections.

To the best of our knowledge, this study is the first *global* test for the relevance of politically irrelevant events. Previous studies have predominantly focused on the U.S. (Achen and Bartels 2017; Graham et al. 2023b; Healy, Malhotra, and Mo 2010; Kitamura 2022), with exceptions in only a few developed countries (Alfano and Ercolano 2023; Müller and Kneafsey 2023; Rapeli and Söderlund 2022). While Kikuta and Uesugi (2023) have analyzed the effects of the European professional soccer games on protests and public opinion in Africa, their

analysis remains confined to fewer than 15 countries on the continent. This narrow focus hinders the generalizability of claims. Moreover, our finding that the effects vary substantially across countries highlights a new direction for the literature; future studies must explain *under what conditions*, not merely *whether*, irrelevant events affect public opinion.

Motivation: Voter Rationality, Political Relevance, and Sports Events

Voter rationality—the assumption that voters update their beliefs following Bayes' rule—is essential to the classical models of voting behavior (see Ashworth, Bueno de Mesquita, and Friedenberg 2017 for review; Besley 2006; Canes-Wrone, Herron, and Shotts 2001; Fearon 1999, 2011; Ferejohn 1986; Lohmann 1998). Although it is debatable whether voter rationality improves social welfare (Ashworth and Bueno de Mesquita 2014), the theoretical predictions of the classical models depend on voter rationality. Thus, testing the validity of this assumption remains important (Ashworth, Bueno de Mesquita, and Friedenberg 2018).

Empirical studies have tested voter rationality by examining whether seemingly irrelevant features influence voters. Previous studies show that voting behavior is influenced by subtle features at polling stations, such as ballot designs (Calvo, Escolar, and Pomares 2009; Ho and Imai 2008; Reynolds and Steenbergen 2006), candidate appearances (Atkinson, Enos, and Hill 2009; Berggren, Jordahl, and Poutvaara 2010; Todorov et al. 2005), candidate names (Auer, Portmann, and Tichelbaecker 2023; Fukumoto and Miwa 2018; Muraoka 2021), and order on the ballot (Belschner 2024; E. Chen et al. 2014). Another group of studies investigates the effects of politically irrelevant events—those beyond the control of politicians and for which no response is expected. Although earlier studies have examined the effects of natural disasters (Fair et al. 2017; Gasper and Reeves 2011; Healy and Malhotra 2009, 2010; Heersink, Peterson, and Jenkins 2017; Lazarev et al. 2014; Malhotra and Kuo 2008; Ramos and Sanz 2020; Reeves 2011; Remmer 2014), politicians are responsible for disaster preparedness and

relief, and thus, natural disasters are expected to influence the beliefs of rational voters (Ashworth, Bueno de Mesquita, and Friedenberg 2018).

Given such difficulties, recent studies have focused on sports events among other politically irrelevant events.[1] Since politicians are not responsible for the results of sports events, nor are they expected to respond, such events should not alter the beliefs of rational voters. Any observable effects of sports events, therefore, constitute crucial evidence against the assumption of voter rationality (Ashworth, Bueno de Mesquita, and Friedenberg 2018; Fowler and Montagnes 2023a; Healy, Malhotra, and Mo 2010). Although politicians may be responsible for the average performance of a national team, they cannot control every aspect of sports events. The outcomes of sports games often depend on elements of luck, such as shots hitting the post in soccer games (Healy and Malhotra 2010; Kikuta and Uesugi 2023). Owing to the importance of this issue, much attention has been paid to the validity of the evidence (Busby and Druckman 2018; Busby, Druckman, and Fredendall 2016; Fowler and Montagnes 2015, 2023a, 2023b; Graham et al. 2023b, 2023a; Healy, Malhotra, and Mo 2010, 2015; M. K. Miller 2013).

However, those studies almost exclusively focus on the U.S., neglecting generalizability across countries. Although Busby and Druckman (2018) call for examining "under what conditions to irrelevant event effects occur" (9), the question remained unanswered until very recently. Rapeli and Söderlund's (2022) study is the first to expand the scope to a multiparty context in Finland, indicating that the Olympic Games did not affect

[1] Other studies leverage UFO sightings (Kitamura 2022), Christmas lottery (Bagues and Esteve-Volart 2016), foreign aid (Cruz and Schneider 2017), local tax increase (Sances 2017), rainfall (Meier, Schmid, and Stutzer 2016), minimum wage (B. Jiménez 2023), sunshine (Bassi 2019), wind (Mo et al. 2023), and natural disasters without any damages (Gallagher 2021).

government approval. Similarly, Müller and Kneafsey (2023) extended the analysis to Gaelic football in Ireland, finding little evidence of its impact. Conversely, Kikuta and Uesugi (2023), who analyzed the effects of European professional soccer on protests and public opinion in Africa, found that losses in soccer games substantially decreased the support for an incumbent leader. Alfano and Ercolano (2023) also found similar results in Spain.

These studies signify a "comparative turn" in the literature, yet they continue to focus on only one or a few countries, neglecting generalizability. As Müller and Kneafsey (2023) state in the conclusion, "examining the robustness of high-profile findings across different case contexts is key to determining the factors that may condition the impact of irrelevant events" (325). Similarly, Rapeli and Söderlund (2022) argue that "[w]hile some scholars question whether looking for political effects of non-political events is a meaningful exercise, there is still room for more contextual variety in the analyses before the idea should be dismissed" (7). We take this issue seriously and analyze the effects of international soccer games on a global scale.

Although our claim—that international soccer games influence leader approval—may appear extraordinary, sports events have been shown to affect a wide array of outcomes, including mood (Ge 2018; Gkorezis et al. 2016, 2016; Otto, Fleming, and Glimcher 2016; Schwarz et al. 1987), hooliganism (Priks 2010), unhealthy eating (Cornil and Chandon 2013), domestic violence (Card and Dahl 2011; Kirby, Francis, and O'Flaherty 2014), crime (Ge, Barbieri, and Schneider 2021; Kalist and Lee 2016; Marie 2016), judiciaries (D. L. Chen 2016; Eren and Mocan 2018), and stock returns (Edmans, García, and Norli 2007; Hirshleifer and Shumway 2003). Other studies also show that sports events influence nationalism, attitudes toward out-groups (Alrababa'h et al. 2021; Arnold 2021; Auer and Ruedin 2023; Gries, Crowson, and Sandel 2010; Kim and Lopez de Leon 2019; Pinto 2024; Rosenzweig and Zhou 2021), and interstate and intrastate violence (Bertoli 2017; Depetris-Chauvin, Durante, and

Campante 2020; Scharpf, Glabel, and Edwards 2023).[2] Given these findings, it is surprising that no consensus has been reached regarding leader approval.

Theory: Rational Update and Psychological Biases

From a rationalist perspective, lucky victories or unlucky losses in international soccer games should not affect leader approval. As mentioned previously, although leaders and governments may influence the overall strength of the national team, they do not control every aspect of soccer games.[3] Thus, even though rational voters can potentially base their evaluations of leaders on the overall performance of their national teams (e.g., FIFA scores),[4] they should not overreact to the random results of individual games.

Although the results of international soccer games can be considered noisy signals of leader performance, rational voters do not change their beliefs when they have access to more reliable indicators such as FIFA scores. These indicators allow rational voters to assess a leader's performance, and the random results of soccer games provide no additional information.[5] Thus, unless accessing those indicators is costly,[6] the random results of soccer games should not influence rational voters' beliefs. Consistent with this perspective, the

[2] Other studies examine the effects of sports participation (Lowe 2021; Mousa 2020).

[3] At least during our study period (2008–2023), instances of bribing judges was extremely rare in the World Cups and regional tournaments.

[4] The FIFA scores and rankings are updated monthly and derived from a game theoretic model (Elo rating system), which calculates a team's probability of winning against an average opponent.

[5] In Appendix A1, we formalize this notion by using a simple model of Bayesian update.

[6] Game results are typically broadcasted along with information about opponent teams, including their FIFA rankings.

consensus in the literature suggests that the random results of sports games should not affect rational people's evaluations of a political leader (Ashworth, Bueno de Mesquita, and Friedenberg 2018; Fowler and Montagnes 2023a; Healy, Malhotra, and Mo 2010). However, in the following subsections, we argue that sports games can affect voter attitudes via psychological channels.

Psychological Biases

While people are assumed to be rational, their rationality can be limited in the real world. For example, they might blindly attribute the results of soccer games to a leader (blind retrospection; Achen and Bartels 2017). Even though people can more accurately evaluate leader performance by looking at the national team's average performance (e.g., FIFA scores),[7] individual soccer games are often more impressive. To illustrate, consider a stylized example: a citizen in a middle-ranked country (e.g., Benin) evaluates the performance of their national leader. This citizen's belief is likely to be more positively updated when their team miraculously defeats France than when their team's FIFA ranking rises by 10 places. This is because the former event is more psychologically impressive, even though the latter is a more accurate measure of the team's average strength. As a result, the citizen may misperceive that the result of the soccer game would be a more indicative measure of leader performance. From this perspective, people mistakenly use soccer results as a benchmark for leader performance (*misperception mechanism*).

Moreover, international soccer games can affect public opinion by eliciting emotional reactions. People may simply act depending on their mood, even without using international soccer games as a benchmark for leader performance. Their national team's victories in international soccer games may enhance their moods, leading them to positive evaluations of

[7] See footnote 4 for details of the FIFA scores.

a leader (or anyone else), while losses in international soccer games may induce sadness or anger, resulting in negative evaluations of a leader (or anyone else; *mood mechanism*). Similarly, international soccer games can rally people around the national flag, and nationalistic sentiment may boost popular support for a national leader. In contrast, losses in international soccer games can hurt nationalism and disarray unified support for a leader (*rally mechanism*). In all cases, it is expected that people evaluate their leaders negatively after losses and positively after victories.

Asymmetric Effects

These effects, however, can be asymmetric, as people react differently to soccer victories and losses (Kikuta and Uesugi 2023). First, people can more strongly affiliate themselves with their national teams upon victories than upon losses. The "basking-in-reflected-glory" in the social psychology literature suggests that people associate themselves with the success of others, whereas they distance themselves from others' failures (cutting-off-reflected-failure; Boen et al. 2002; Cialdini et al. 1976). From this perspective, victories in international soccer games can attract more casual fans and hence affect a greater number of people, while losses may make people indifferent about international soccer and thus have weaker or even nonexistent effects. It is therefore expected that soccer victories boost leader approval, while losses lack compensating effects (*positive symmetric effect*).[8]

[8] The positive symmetric effect can be even larger due to leaders' selective credit claiming. Leaders often claim credits for victories, while they have little incentive to highlight their contribution upon losses (Ramos and Sanz 2020). Moreover, soccer victories can wash out negative images of a leader, while losses may not produce equivalent effects (Glaßel, Scharpf, and Edwards 2024; Scharpf, Glabel, and Edwards 2023).

Second, people can also asymmetrically attribute victories and losses in international soccer games, albeit in a different manner. The success-failure bias in psychological research suggests that people blame others for failures and attribute successes to themselves (D. T. Miller and Ross 1975). More broadly, people are more sensitive to negative stimuli than to positive ones (general negativity bias; Baumeister et al. 2001). If this holds, people would blame unlucky losses on a leader, who is one of the most prominent figures in a country, and attribute victories to themselves (e.g., people may perceive that their strong fandom is rewarded with a victory). The egoistic attribution implies that while victories in international soccer games do not change leader approval, losses can decrease it (*negative asymmetric effect*).

In summary, as presented in Table 1, while rational updating constitutes a null expectation, psychological biases (i.e., misperception, mood, and rally mechanisms) suggest alternative possibilities about the effects of victories and losses, which may be either symmetric or asymmetric. Given the theoretical uncertainty regarding whether the asymmetricity in affiliation (i.e., positive symmetric effect) or that in attribution (i.e., negative asymmetric effect) predominates, we subject these possibilities to empirical testing.

Table 1. Predicted Effects of Victories and Losses on Leader Approval

Mechanisms	Predicted effects	Lucky victories	Unlucky losses
Rational update	Null effect	0	0
Psychological biases (misperception, mood, rally)	Symmetric effect	+	−
	Positive asymmetric effect	+	0
	Negative asymmetric effect	0	−

The table shows the predicted effects of lucky victories and unlucky losses in international soccer games on public support for national leaders. +, −, and 0 indicate an increase, decrease, and no change in leader approval, respectively.

Case: The World Cup and Regional Tournaments

Testing these hypotheses involves significant challenges. It is crucial, both theoretically and empirically, to isolate the effects of random results from those of the national team's average performance. The team's average performance is endogenous to various social, political, and

economic conditions (L. F. Jiménez 2016; Miguel, Saiegh, and Satyanath 2011). Meanwhile, we are interested in the effects of random game results. While voters may rationally refer to the team's average performance as a benchmark for evaluation, they should not be influenced by the random results of individual games.

We address these challenges by examining international soccer games, including the men's and women's World Cup, the European Football Championship (Euro), Africa Cup of Nations, Copa America, Gold Cup, and Asian Cup.[9] These games are among the most popular sporting events globally. According to Nielsen Sports (2018), on average, over 40% of adults in 18 surveyed countries expressed an interest or strong interest in soccer. FIFA (2018, 2022) has also reported that over 3.5 billion people watched the 2018 World Cup, and the number even increased to more than 5.9 billion in 2022—accounting for over 74% of the global population. While the men's World Cup enjoys global popularity, regional tournaments such as the Euro, Africa Cup of Nations, and Copa America are equally popular within their respective regions. It has been reported that approximately 2 billion people watched the 2021 Africa Cup—a figure surpassing the total population of Africa (Confederation of African Football 2022). These games are live-covered in nearly every participating country, and the results of the games are broadcasted extensively across various media platforms including newspapers, radio, television, and social media. The overwhelming popularity of these

[9] Since the Women's World Cup, Gold Cup, and Asian Cup attract less attention, we also conduct subsample analysis for each tournament, details of which are discussed in a subsequent section about effect heterogeneity. Our sample includes the playoffs, group stage, and qualification rounds, except for Copa America and Gold Cup. For these tournaments, due to their limited number of participating countries, we use only the playoffs and group stage, as qualification stages are not always conducted.

international soccer games offers a unique opportunity to conduct a *global* test on the relevance of seemingly irrelevant events.

Importantly, unlike the Olympic Games, betting on international soccer games is permitted. Therefore, we can use pre-game betting odds as a measure of expected outcomes to identify the effects of *random* results (Card and Dahl 2011; Healy and Malhotra 2010). As emphasized by Healy et al. (2015), this feature is important for causal identification, distinguishing our research from some of the recent studies (Alfano and Ercolano 2023; Müller and Kneafsey 2023; Rapeli and Söderlund 2022). Thus, international soccer games provide an excellent platform to enhance both internal and external validity.

Design: Double Natural Experiments

We identify the causal effects of international soccer games on leader approval by leveraging two sources of natural experiment. First, we exploit the as-if random assignment of interview dates in the Gallup World Poll (GWP)—the most comprehensive global survey. Gallup has conducted extensive surveys across more than 160 countries since 2006.[10] Importantly, the dates of these survey interviews were pre-determined and thus unaffected by the scheduling of international soccer games. This means that respondents are randomly assigned to be interviewed either before or after a game. This as-if random assignment enables us to identify the causal effects of soccer games (Muñoz, Falcó-Gimeno, and Hernández 2020; Seo and Horiuchi 2023).

Note, however, that a simple before-and-after comparison does not allow us to identify the effects of soccer *results*. For example, comparing respondents interviewed before and after a winning game does not distinguish between the overall effects of soccer games and the specific effects of victories. Neither can we use the difference-in-differences (DiD) approach

[10] The interviews are conducted either face-to-face or via telephone.

by comparing the changes in public opinion after winning, drawing, and losing games. Stronger teams are more likely to win, and their supporters are more engaged and thus likely to be more responsive to game outcomes. More importantly, none of these designs identifies the quantity of our primary interest—the effect of *random* results.

We address these issues by conditioning on pre-game betting odds (Card and Dahl 2011; Healy, Malhotra, and Mo 2010). Since inaccurate betting odds can result in substantial financial losses, bookmakers have strong incentives to make accurate predictions. Studies have shown that betting odds can predict outcomes more accurately than even sophisticated statistical models (Leitner, Zeileis, and Hornik 2010; Wunderlich and Memmert 2018). Therefore, when two games share similar betting odds, we can infer that any difference in their outcomes is unexpected and thus as-if random.[11]

Our design combines two natural experiments, which effectively results in a DiD, where the treatment is randomly assigned over time (i.e., interview dates) and across units (i.e., game results). This means that our main approach is the natural experiments, with DiD serving as supplementary. When both interview dates and game results are randomly assigned, conditioned on betting odds, the common trend assumption—a core assumption of DiD—is also satisfied. Thus, while we acknowledge the recent development in DiD methods (see Xu 2023 for a review), we do not employ them but instead literally calculate the "difference-in-differences." Our design does not involve staggered adoption, two-way fixed effects, or clear violations of the common trend assumption (see the results of a later event study).

[11] We use betting odds, as they are available for every game. We later show that matched pairs have similar FIFA scores and rankings.

Sample and Unit

The unit of our analysis is respondent i interviewed t days before or after an international soccer game j of their national team. We first link respondents in the GWP to soccer games of their national teams by using the closest dates.[12] Following the precedence of previous studies, we exclude soccer games if all respondents were interviewed either entirely before or after a game (Goldsmith, Horiuchi, and Matush 2021; Kikuta and Uesugi 2023; Seo and Horiuchi 2023). Since the effect of soccer games should be instantaneous, we analyze only those respondents who were interviewed within three days before or one day after soccer games.[13] To avoid the compound effects of multiple games, we also exclude games that occurred within three days before or one day after another game.[14] This results in 621 games and 90,565

[12] If there were no games within a week before or after survey interviews, the respondents are excluded. Since the U.K. has four different teams (England, Wales, Scottland, and Northern Ireland), we assign the British respondents' national teams based on their residential locations. Removing the U.K. from the sample does not change the results (see a subsequent robustness check, leave-one-country-out tests).

[13] We subsequently conduct robustness checks with different pre-treatment time periods (1 to 7 days). We also perform an event study to examine the effects during subsequent periods.

[14] A related issue is that respondents may be affected by previous games. For instance, if two different games are held three days before and two days after an interview, only the second game is linked to the respondent. Therefore, the respondent is considered a control unit, despite the potential influence from that game. In a subsequent robustness check, we exclude respondents if there was another game within 7, 14, 21, 28, or 35 days before the game being analyzed.

respondents across 126 countries for the 2008–2023 period.[15] The summary statistics of our data are available in Appendix A1.

Treatment Variables

The first treatment variable $after_i$ takes a value of 1 if respondent i was interviewed after an international soccer game of their national team. For the respondents who were interviewed on the days of soccer games, we assign them to a treated group if the game started before 8 AM at a local time, a control group if the soccer game started after 8 PM at a local time, and otherwise missing treatment status (i.e., we are not sure whether they are interviewed before or after a game).[16] The second treatment variables win_j and $loss_j$ take 1 if a respondent's national team won or lost game j, respectively.

Outcome Variable

The outcome variable $approval_i$ is respondent i's answer to the question: "Do you approve or disapprove of the job performance of the leadership of this country?" This item is "the most commonly used question to measure the approval rating of a political leader" (Seo and Horiuchi 2023, 14). The outcome variable takes the value of 1 for approval, 0 for disapproval, and missing for "don't know" and non-responses.[17]

Control Variable

The control variables $p_{win,j}$ and $p_{loss,j}$ are the expected probabilities of a victory and loss in game j, respectively. We scrape the pre-game betting odds from Odds Portal,[18] and transform

[15] Although the GWP data are also available for the 2006-2007 period, the interview dates are missing from the dataset.

[16] We subsequently conduct a robustness check by excluding all those respondents.

[17] We subsequently conduct a robustness check by using a dummy variable for each response.

[18] https://www.oddsportal.com (scraped on March 3, 2024).

the odds to expected probabilities of wins, draws, and losses (i.e., the inverse of decimal odds). We then standardize the probabilities to remove bookmakers' profit margins (e.g., $p_{win,j} = \frac{\tilde{p}_{win,j}}{\tilde{p}_{win,j}+\tilde{p}_{draw,j}+\tilde{p}_{loss,j}}$, where the tilde indicates probabilities before transformation).

Specification

With the expected probabilities of treatment assignment, we implement nearest-neighboring matching without replacement.[19] While regression control for $p_{win,j}$ or $p_{loss,j}$ relies on functional-form assumptions, the matching allows us to compare only similar units, reduce noises, and check balances without parametric assumptions (Ho et al. 2007).[20] To avoid matching the games of very different $p_{win,j}$ or $p_{loss,j}$, we use an unstandardized caliper of 0.1 (i.e., the treated and control games differ in $p_{win,j}$ or $p_{loss,j}$ by at most 0.1).[21] We implement matching separately for winning and losing games. The matched sample of winning versus other games contains 310 games and 36,037 respondents in 104 countries (2008–2023). Similarly, the sample of losing versus other games contains 285 games and 32,006 respondents in 102 countries (2008–2003).

Figure 1 shows the geographic distribution of respondents after matching (the two samples are combined). Although we do not intend to claim that the samples are globally representative, they cover nearly all regions. The exceptions are North Africa and the Middle East, where the outcome variable is missing in the GWP. The samples also tend to over-represent developed countries and under-represent autocracies (e.g., China) and countries under armed conflicts (e.g., Somalia), thus reflecting the challenges associated with conducting

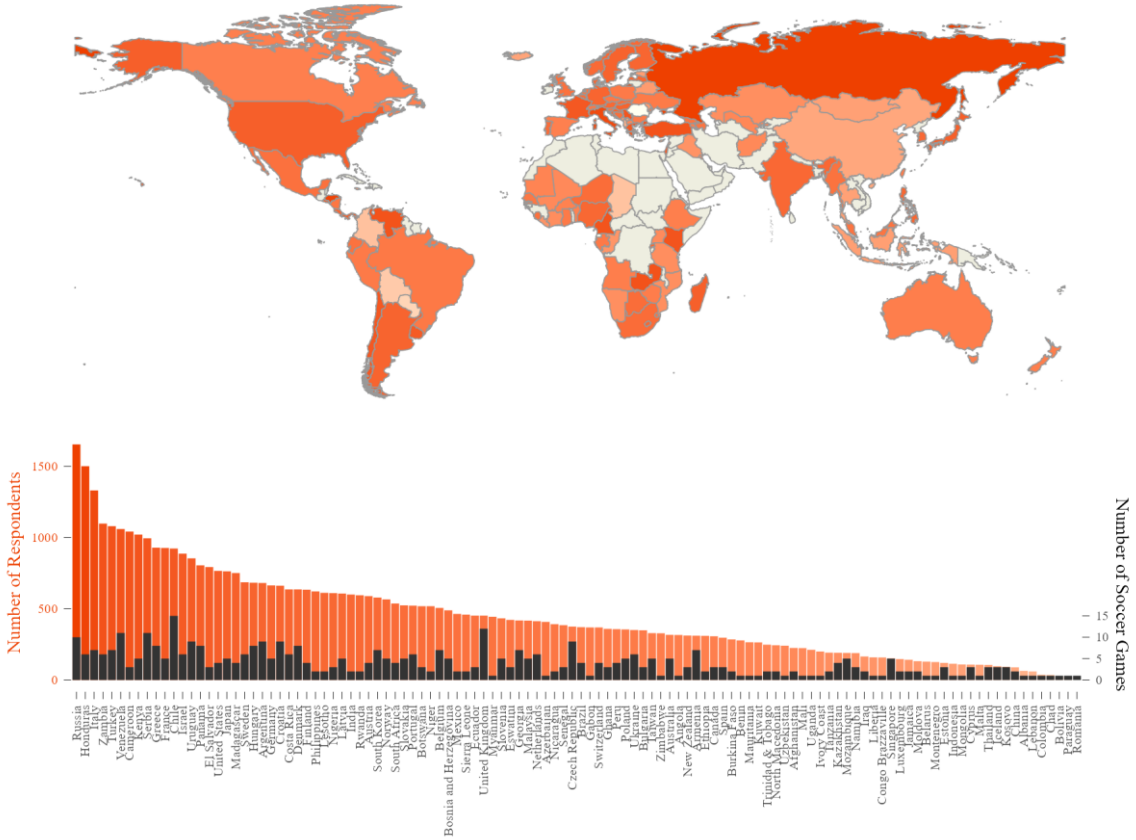
[19] Since we have only one control variable, we use the simplest matching method. T.

[20] We subsequently conduct a robustness check with regression control and "doubly robust" regressions with $p_{win,j}$ and $p_{loss,j}$.

[21] We subsequently conduct robustness checks with various caliper sizes.

survey interviews in these areas. Importantly, most of the missing data come from the missing values in the GWP, rather than from matching outcomes or a lack of overlap between soccer games and survey interview dates. Thus, despite these limitations, we believe our samples have the best feasible coverage.[22]

Figure 1. Geographic Distribution of Respondents (After Matching)



Note: The upper panel is the map of countries included in the matched samples. The color tones correspond to the number of available respondents shown in the bottom panel (vertical axis on the left). The bottom panel also shows the number of available soccer games (vertical axis on the right).

With the matched samples, we use linear regressions;[23]

$$approval_i = \alpha_{win} + \beta_{win,1} after_i + \beta_{win,2} win_j + \delta_{win}\, after_i\, win_j + \varepsilon_{win,i}; \quad (1)$$

$$approval_i = \alpha_{loss} + \beta_{loss,1} after_i + \beta_{loss,2} loss_j + \delta_{loss}\, after_i\, loss_j + \varepsilon_{loss,i}. \quad (2)$$

[22] The World Value Survey has much fewer respondents and lower coverage than the GWP. Among the Barometer Surveys, interview dates are available only in Afrobarometer.

[23] We use fixest package in R (Berge et al. 2020).

Because the righthand sides of (1) and (2) contain only dichotomous variables and are saturated, the models do not depend on functional forms, and $\delta_{\cdot}$ is literally a "difference-in-differences" (e.g., the average change in $approval_i$ before and after winning games minus the average change before and after the other games).[24] The estimates $\hat{\delta}_{\cdot}$ correspond to the average treatment effect on the treated (ATT) within the matched sample. The constituent terms $\alpha_{\cdot}$, $\beta_{\cdot,1}$, and $\beta_{\cdot,2}$ do not have causal interpretations. Since the treatment variables are randomly assigned, we do not include demographic covariates or fixed effects.[25] The DiD accounts for any static confounders. Although the soccer games tend to be held on weekends, the DiD cancels out the calendar effects. Similarly, even though politicians and other actors may take anticipatory actions before the soccer games (e.g., visiting a host country), this does not cause problems as far as the results or the soccer games are random conditional on $p_{win,j}$ and $p_{loss,j}$. Finally, as the treatment is assigned at the level of soccer games, we cluster the standard errors thereby (Abadie et al. 2023).[26]

Result: Irrelevant Events Do Influence Leader Approval

Table 2 shows the results of the main analysis. As shown in Model 1 of Table 2, the national team's victories increased leader approval by 7 percentage points. This is equivalent to a 14% increase from the sample average (0.51). By contrast, the effect of the national team's losses on leader approval was close to zero (Model 2 of Table 2). This suggested that people gave

[24] For this reason, we do not use logit or probit models.

[25] We subsequently conduct robustness checks with demographic covariates, unit fixed effects, calendar fixed effects, and control for differential time trends (Goodman-Bacon 2021).

[26] We subsequently conduct a robustness check with standard errors two-way clustered by games and national teams. Without clustering, nearly everything becomes statistically significant.

credit to leaders for victories, but did not blame them for losses. These results are consistent with the notion of the positive asymmetric effect, though we cannot reject the symmetric effect given the relatively large standard errors for losses.

Table 2. The Effect of Victories and Losses on Leader Approval

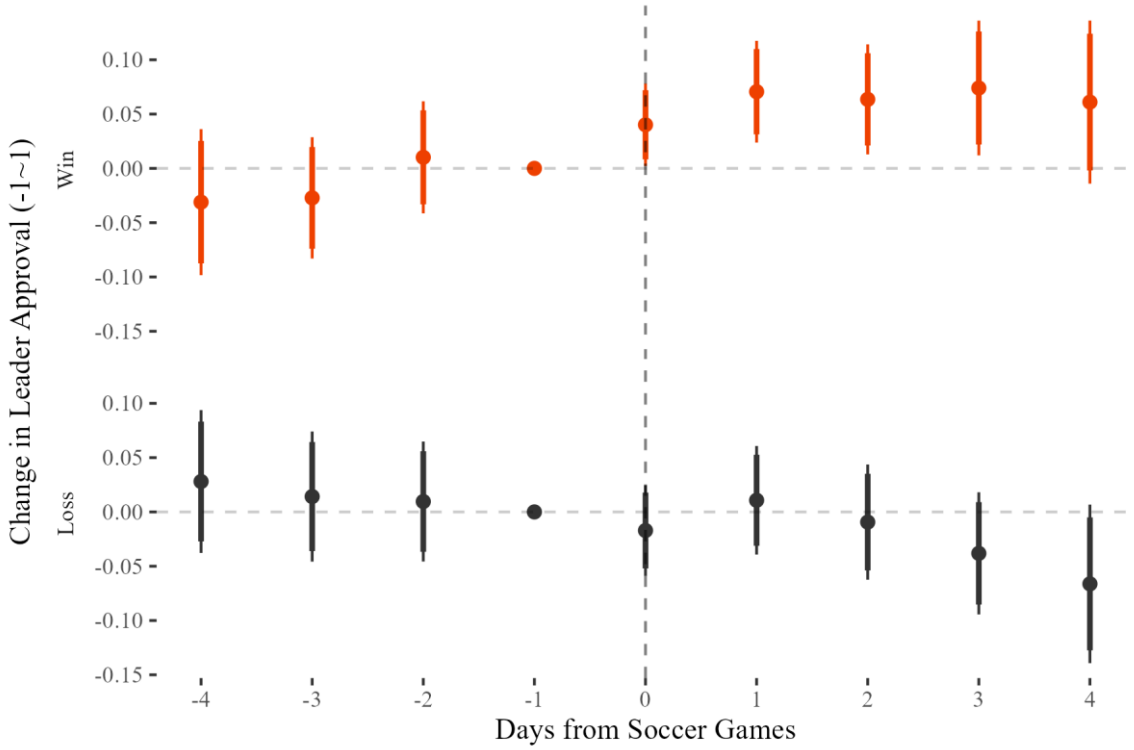
	Leader Approval	
	1	2
Win × After	0.07**	
	(0.03)	
Win	-0.02	
	(0.03)	
Loss × After		0.00
		(0.03)
Loss		0.08*
		(0.03)
After	-0.05*	0.00
	(0.02)	(0.02)
(Intercept)	0.51**	0.47**
	(0.02)	(0.02)
N	36,037	32,006

The table shows the estimates of the coefficients. The standard errors clustered by games are in parentheses. The first (second) column includes 104 (102) countries and 310 (285) games for 2008–2023. $**$ $p < 0.01$; $* p < 0.05$; † $p < 0.1$.

Figure 2 shows the results of the event study, where the average leader approval at $t = -1$ is compared to those in other time periods. As seen in the top panel of Figure 2, victories instantaneously increased leader approval. By contrast, losses had no immediate impact, and there was a slight decrease in leader approval at $t = 4$. While the effects of victories are robust to various changes, as we will demonstrate later, the effects of losses are null across most specifications, cautioning against overinterpreting the results of the event study. Additionally, the event study indicates no clear time trend in the pre-treatment periods ($t \in [-4, -1]$), providing a credence to the common trend assumption. This is not surprising, as our main

design is natural experiments, and hence, the treatment is randomly assigned over time and across units.

Figure 2. The Results of Event Study



The figure shows the results of event study, where the average leader approval at $t = -1$ is compared to that in other time periods. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Validity Checks

To verify whether respondents were randomly assigned to interviews before or after soccer games, we checked the covariate characteristics in Table 3. As shown in this table, the data indicate no consistent evidence of imbalances: standardized mean differences are close to zero, variance ratios are also close to 1, and the p-values of the difference-in-means tests are generally above 0.1, with only a few exceptions.

Table 3. Balance Checks

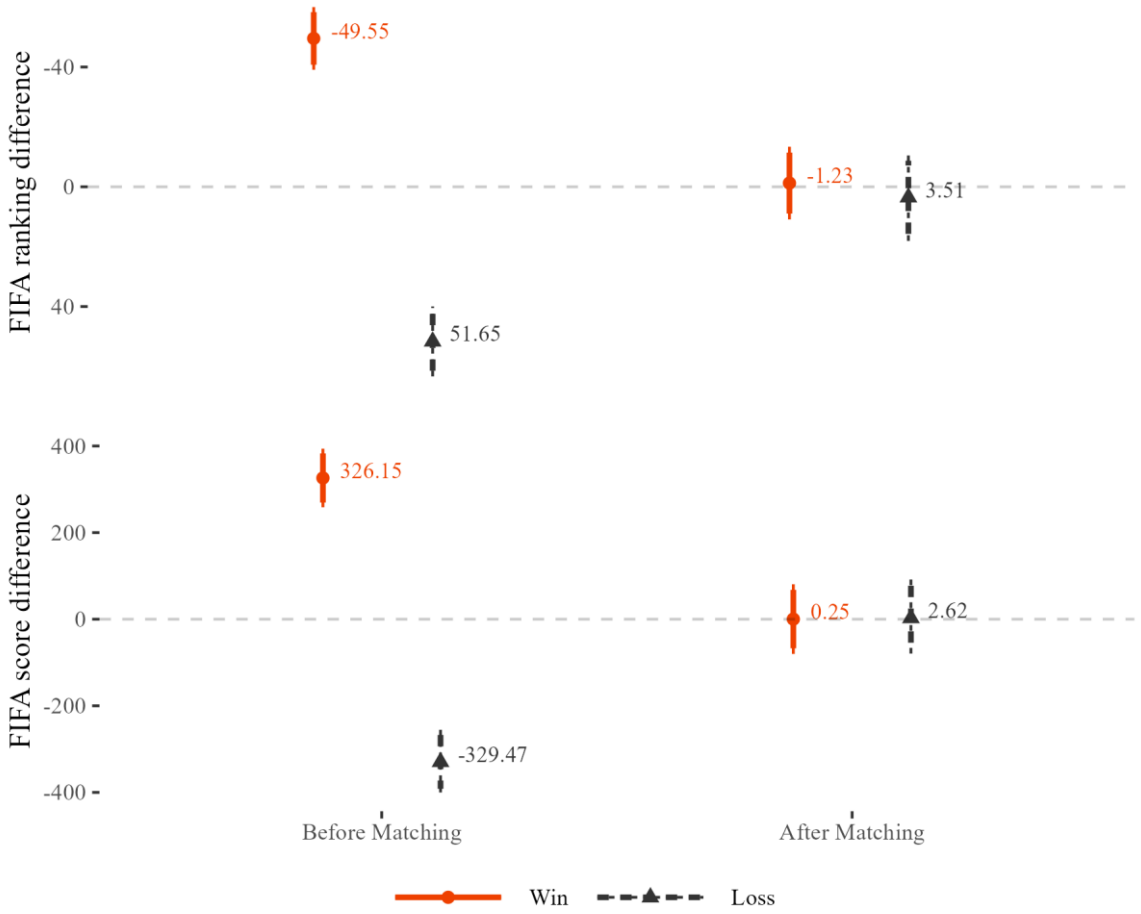
	Mean (Treated)	Mean (Control)	Std. difference	Var. ratio	p-value
Female	0.54	0.54	0.00	1.00	0.71
Age	42.72	43.19	−0.02	0.96	0.35
Educ. (2nd)	0.53	0.53	0.00	1.00	0.82
Educ. (higher)	0.18	0.19	0.00	0.99	0.83
Married	0.45	0.47	−0.03	0.99	0.09
Income	2.66	2.72	−0.04	1.01	0.10
Christian	0.70	0.68	0.04	0.96	0.19
Muslim	0.12	0.13	−0.03	0.92	0.37

The table shows the results of the balance check. The "control" and "treated" respondents are those interviewed before and after soccer games, respectively. As a rule of thumb, they are considered balanced if the standardized mean differences (Std. difference) are between -0.2 and 0.2, the variance ratios (Var. ratio) between 0.5 and 2.0, and the p-values of the difference-in-means tests are above 0.1. The standard errors of the difference-in-means tests are clustered by soccer game.

In Figure 3, we also checked whether the results of soccer games were randomly assigned conditional on betting odds. To this end, we compare the national teams' FIFA rankings and scores with and without matching.[27] Without matching (lefthand panel of Figure 3), winning teams shown with circles (●) have better rankings and scores than losing teams shown with triangles (▲). However, as seen in the righthand panel of Figure 3, matching nearly eliminates those differences. This finding confirms that the game results are randomly assigned conditional on betting odds. In Figure A 3-1 of Appendix A3, we also confirm that the betting odds accurately predict the results of international soccer games.

[27] We used the latest ranks and scores before each soccer game. The FIFA rankings and scores were not used in the matching process.

Figure 3. Randomization Checks



The figure shows the national teams' FIFA rankings and scores relative to those of an opponent team before (right panel) and after (left panel) the matching. The vertical axis of the top panel is reversed as a smaller value in the FIFA ranking means a stronger team. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Robustness Checks

It is still possible that our findings may be false positives. The above-stated analysis rests on arbitrary choices in the sample selection, measurement, and specification. This raises a concern that we might "have taken a forking path that happens to produce an unreliable estimate, or the result could just be a false positive that arose through bad luck" (Fowler and Montagnes 2023a, 1). We address this concern by conducting extensive robustness checks.

First, as summarized in Table 4 and detailed in Appendix A3, we replicate the analysis by using different pre-treatment time periods (1 to 7 days), excluding games held within 7 to 35 days after another game, excepting over-sampled countries (Honduras, Italy, and Russia),

omitting each country individually, employing different measurements of the outcome and treatment variables, adopting various empirical approaches (regression control, matching with different caliper sizes, and doubly robust estimators), adding diverse control variables and fixed effects, and using different calculation of the standard errors. As presented in Table 4, the results remain robust across such changes.[28]

Table 4. Robustness Checks

	Win	Loss	Appendix
1. Different pre-treatment time periods (1 ~ 7 days)	$+^{**1}$	null	Figure A 4-1
2. Omission of games within 7, 14, 21, 28, or 35 days after the last game	$+^{**}$	null	Figure A 4-2
3. Omission of Honduras, Italy, and Russia	$+^{*}$	null	Table A 4-1 Table A 4-2
4. Omission of each country (leave-one-country-out test)	$+^{**2}$	null	Figure A 4-3
5. Measurement of the outcome (DK/Refuse as a baseline)	$+^{*}$	null	Table A 4-1 Table A 4-2
6. Measurement of the treatment (exclusion of all respondents at $t = 0$)	$+^{*}$	null	Table A 4-1 Table A 4-2
7. Regression control for betting odds	$+^{\dagger}$	null	Table A 4-1 Table A 4-2
8. Caliper sizes (0.02 ~ 0.21)	$+^{**3}$	null	Figure A 4-4
9. Doubly robust regression	$+^{**}$	null	Table A 4-1 Table A 4-2
10. Control for demographic covariates	$+^{**}$	null	Table A 4-1 Table A 4-2
11. Country-game fixed effect	$+^{*}$	null	Table A 4-1 Table A 4-2
12. Year, month, weekday, and day fixed effects	$+^{**}$	null	Table A 4-1 Table A 4-2
13. Control for differential time trends (Goodman-Bacon 2021)	$+^{*}$	null	Table A 4-1 Table A 4-2
14. Standard errors two-way clustered by game and national team	$+^{**}$	null	Table A 4-1 Table A 4-2

The table summarizes the results of the robustness checks. $** \; p < 0.01$; $* \; p < 0.05$; $\dagger \; p < 0.1$.
Note 1: Significant at the 5% level when the pre-treatment time period is 2 days before soccer games.
Note 2: Significant at the 5% level when Rwanda or Russia is dropped.
Note 3: Significant at the 5% level when the caliper size is 0.02.

[28] The estimates become smaller and less statistically significant when we control for $p_{win,j}$ and $p_{loss,j}$ instead of using matching (robustness check 7 in Table 4). This is not surprising, as the analyses without matching include incomparable cases, increase noises, and decrease the power of the analysis (Ho et al. 2007).

To further examine the robustness of our findings, we also consider how the *combinations* of different specifications change the results. We identify 2,688 combinations of specifications that characterize the forking paths in our analysis, including the measurement of the outcome and treatment variables, statistical methods, time window of the pre-treatment period, different subsamples, additional control variables, and fixed effects (see Appendix A5 for details). To make inferences *across* the specifications, we follow Simonsohn et al. (2020) and create 500 placebos by randomly assigning the results of soccer games based on their pre-game betting odds.[29] We then estimate the effects of the treatment and placebos for each specification. This effectively results in $2{,}688 \times (1 + 500) = 1{,}346{,}688$ regressions.[30] The results are summarized by the median of the treatment effects, proportion of the estimates with the same sign (i.e., "dominant sign"), median of the p-values, proportion of p-values less than 0.05, and proportion of the estimates with the same sign and statistical significance (Simonsohn, Simmons, and Nelson 2020). These statistics are compared with their distributions under the null hypothesis (i.e., the results with the 500 placebos).

Table 5 presents the results of the analysis. Across 2,688 specifications, the median treatment effect of victories is 0.03, with the sign being the same (i.e., positive) in 96% of the specifications. The treatment effects of victories are statistically significant at the 5% level and have the same sign in 45% of the specifications. Although these numbers may or may not appear robust, it is very unlikely that we would observe such results if victories had no effect.

[29] $placebo_{win,j} \sim \text{Bernoulli}(p_{win,j})$ and $placebo_{loss,j} \sim \text{Bernoulli}(p_{loss,j})$.

[30] We limit the number of placebos to 500 because running more than a million regressions with the sample of over 30,0000 respondents takes an extensive amount of time even with parallelization.

As shown in the brackets of Table 5, we find larger or more statistically significant estimates in less than 2% of the placebos. By contrast, the effects of losses are not robust.

Table 5. Specification Checks

	Win	Loss
Median(est.)	0.03	-0.01
	[0.02]	[0.78]
Pr(dominant sign)	0.96	0.72
	[0.06]	[0.60]
Median(p-value)	0.07	0.59
	[0.02]	[0.69]
Pr(p < .05)	0.45	0.00
	[0.01]	[0.74]
Pr(dominant sign & p < .05)	0.45	0.00
	[0.01]	[0.70]

The table shows the results of the forking path analysis, in which we estimate the treatment effects of victories and losses under 2,688 specifications. The p-values are in brackets.

Additionally, the effects of victories become weaker under only two scenarios: when we do not use matching or when we include team-game fixed effects. Among 2,304 specifications that use matching but do not include team-game fixed effects, the median effect of victories is 0.08, the median p-value is less than 0.01, and 95% of the estimates have the same sign and are statistically significant at the 5% level.[31] These results are not surprising, as the analyses without matching include incomparable games, increase statistical noises, and thus lower efficiency (Ho et al. 2007). Similarly, because the treatment (i.e., results of soccer games) is assigned at the level of team and game, team-game fixed effects absorb much of the variation in the treatment variable, thereby decreasing the power (Imai and Kim 2019). Thus, with these caveats, we find no evidence that forking paths influence our results.

[31] The p-values of those statistics are less than 0.01.

Mechanisms

Given the validity and robustness of the findings, a natural question arises: why do international soccer games change leader approval? Theoretically, victories can increase leader approval by altering people's subjective perceptions of leader performance (*misperception mechanism*), irrationally referring to random results of soccer games when evaluating their leaders' performance. Alternatively, people might feel good upon victories and thus evaluate their leaders more positively (*mood mechanism*). It is also possible that the national team's victories rally people around the flag, thereby boosting support for the national leader (*rally mechanism*).

To assess the misperception mechanism, we use an item measuring people's evaluation of a leader's job performance. Since this variable highly correlates with the main outcome variable ($r = 0.61$), we also use two additional items. The first item assesses whether a respondent approves the leader performance of an opponent country in a soccer game. If people use the results of soccer games as a benchmark for leader performance, they should positively (negatively) evaluate the leader of a winning (losing) country, regardless of whether it is their own or an opponent country. By contrast, if people emotionally react to soccer games, they should have negative (positive) attitudes toward an opponent country when their national team loses (wins). Unfortunately, however, this item is only available for a very few respondents.[32] The second item is whether a respondent approves the leader of their local city or town. While national leaders might be perceived as responsible for international soccer games, local mayors are not responsible for any aspect of *international* games. Thus, if our results are driven by

[32] The GWP has items about attitudes toward leaders of major and regional powers. When respondents' teams fight against teams from these countries and the survey interviews occur within three days before or a day after the games, these responses are included in the analysis. As expected, the sample size is significantly limited.

misperceptions about leader performance, international soccer games should not affect the evaluations of local mayors. By contrast, if people simply follow their moods, they should positively and negatively evaluate local mayors upon victories and losses, respectively. This question item is available for approximately one-third of the respondents.

To evaluate the mood mechanism, we employ three items: whether a respondent felt enjoyment, anger, and sadness the day before a survey interview. These items are available for most respondents and are particularly relevant for evaluating the mood mechanism. Since these items are specific to the day before an interview, we exclude all respondents interviewed on the days of soccer games (see p.14).[33]

Finally, we assess the rally mechanisms by examining their effects on people's attitudes toward immigrants. Admittedly, these are indirect measures and the GWP does not include other questions specifically related to nationalism. Nevertheless, as Rosenzweig and Zhou (2021) have demonstrated, international soccer games can strengthen nationalism and increase hostility toward immigrants (see also Bertoli 2017; Depetris-Chauvin, Durante, and Campante 2020). Therefore, we use respondents' attitudes toward having foreign immigrants in their countries, as neighbors, and as spouses of their relatives as proxies for these sentiments. The details of these survey items are available in Table A 6-1 of Appendix A6.

Table 6 shows the effects of victories.[34] First, victories positively influenced people's evaluation of their leader's job performance (Model 1). By contrast, following victories, people negatively evaluated leaders of opponent countries in soccer games (Model 2). While caution must be exercised given the small sample size, this result implied that people did not merely react based on their moods. Instead, they consistently benchmarked leader performance;

[33] The results are similar even without excluding these respondents.

[34] The effects of losses are null for all outcome variables. See Table A 6-2 in Appendix A6.

whether it was their own or an opponent country, people positively evaluated the leaders of winning countries. In addition, no evidence suggests that people credit local mayors for victories (Model 3). This also confirmed that people did not simply follow their moods; rather, they selectively attributed victories to politicians who were perceived to be responsible. These results are consistent with the misperception mechanism.

Table 6. Mechanism Checks (Victories)

	Misperception			Mood			Immigrants		
	Leader	Opponent Country	Mayor	Enjoy	Anger	Sad	In my Country	Neighbor	Relatives Married
	1	2	3	4	5	6	7	8	9
Win × After	0.06*	-0.28*	-0.04	-0.03	0.00	-0.01	0.06	0.13*	0.10†
	(0.03)	(0.12)	(0.04)	(0.02)	(0.01)	(0.02)	(0.06)	(0.06)	(0.06)
Win	0.00	0.09	0.02	0.02	-0.01	0.01	-0.07	-0.09	-0.09
	(0.03)	(0.14)	(0.04)	(0.02)	(0.01)	(0.01)	(0.07)	(0.07)	(0.07)
After	-0.06**	0.27*	0.01	0.02	0.00	0.01	-0.07	-0.11*	-0.09†
	(0.02)	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
(Intercept)	0.58**	0.39**	0.56**	0.71**	0.18**	0.21**	0.70**	0.71**	0.66**
	(0.02)	(0.12)	(0.02)	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)
N	31,549	913	10,752	33,918	34,066	34,077	5,455	5,426	5,341

The table shows the effects of victories on the outcome variables noted in the first row. The details of the survey items are available in Table A 6-1 of Appendix A6. The standard errors clustered by games are in parentheses. $** \; p < 0.01; * \; p < 0.05; \dagger \; p < 0.1$.

By contrast, we find no evidence supporting the mood mechanism. The effects of victories on respondents' moods are statistically indistinguishable from zero (Model 4-6 of Table 6). These results align with those of some previous studies (Busby and Druckman 2018; Busby, Druckman, and Fredendall 2016; Kikuta and Uesugi 2023). Moreover, it is also possible that the framing of the survey may have weakened the effects. Questions about moods were asked in the section of respondents' ordinary lives. Thus, respondents may have answered questions regarding feelings considering their living standards rather than special events such as international soccer games. The mood indicators only weakly correlate with leader approval

($r \in [-0.1, 0.1]$). Therefore, while we do not entirely dismiss the potential roles of moods, we conclude that the evidence supporting the mood mechanism is scarce in this study.

Finally, we find evidence against the rally mechanism. Victories in international soccer games improved respondents' attitudes toward immigrants (Model 7-9 of Table 6). Thus, it is unlikely that victories heighten nationalism and boost support for national leaders. Instead, victories in international soccer games facilitated more conciliatory attitudes toward immigrants. These results are consistent with findings from previous research suggesting that sports can contribute to inter-group cooperation (Alrababa'h et al. 2021; Depetris-Chauvin, Durante, and Campante 2020; Mousa 2020) and that many national teams often include players of immigrant origin (e.g., Mbappé). Overall, our evidence is supportive of the misperception mechanism.

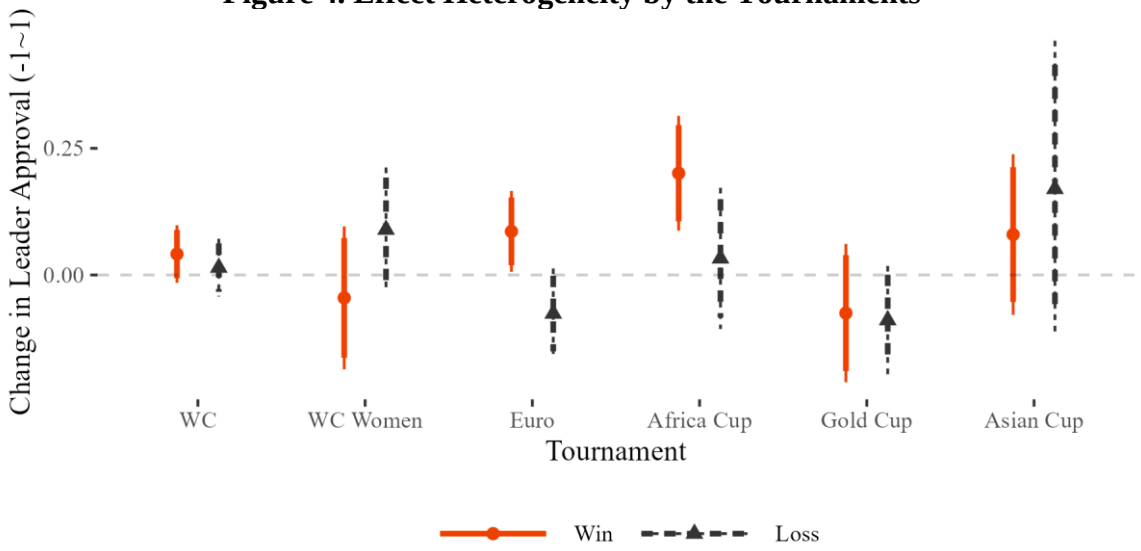
Heterogeneity I: Salience

Thus far, we have discussed the average effects. However, these effects can be heterogeneous. To explore this, we assess whether the effects are more pronounced when the treatment is especially salient. First, as shown in Figure 4, we examine heterogeneity across different tournaments. Copa America has been excluded from the analysis as the size of the subsample is small and hence the effect cannot be fully estimated.[35] According to the results in this figure, victories have larger effects in the Euro and Africa Cup of Nations, smaller and less precise effects in the men's World Cup and Asian Cup, and null effects in the women's World Cup and Gold Cup. This variation mostly correlates with the popularity of the tournaments; the men's World Cup, Euro, and Africa Cup of Nations are among the most popular tournaments. The relatively weaker effect observed in the men's World Cup could be because it includes

[35] See footnote 9 for more details on Copa America.

countries where soccer is less popular among the public, unlike in Europe and Africa where soccer enjoys widespread popularity.

Figure 4. Effect Heterogeneity by the Tournaments



The figure shows the effects of victories and losses in international soccer games on leader approval by the tournaments of the soccer games. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

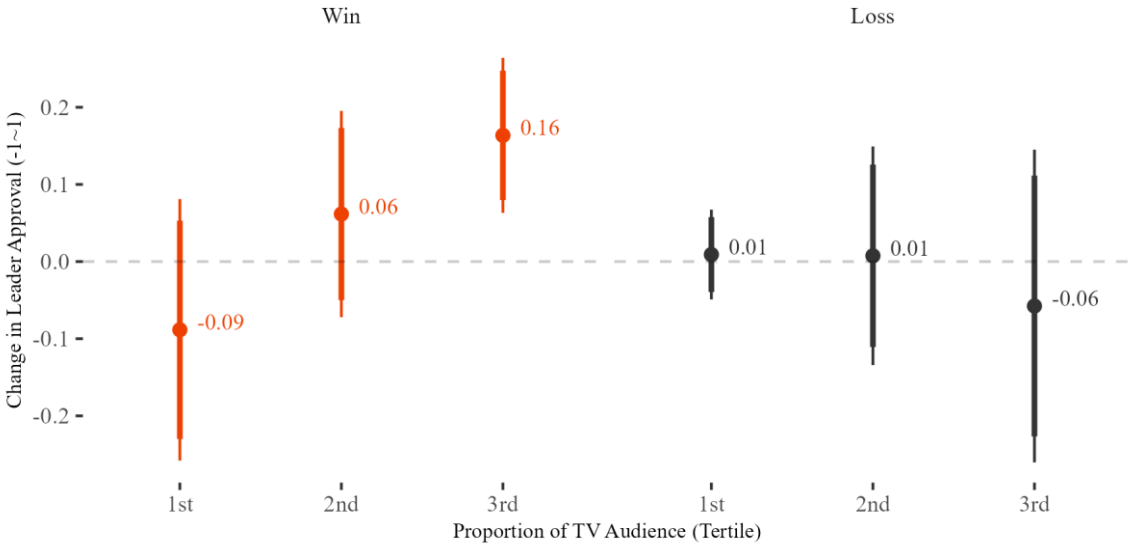
We further explore this possibility by examining whether the effects of the men's World Cup vary depending on the viewership of the games. Although viewership data are rarely available, FIFA has published reports on TV audiences following the 2010 and 2014 World Cups.[36] These reports contain the total audience reach in each country, defined as the number of unique individuals who watched the event on TV at home for more than 20 consecutive minutes.[37] We calculate the proportion of TV audience by dividing the total audience reach by

[36] After 2014, FIFA published only the summaries of the viewership reports without providing country-specific data. Despite attempts to obtain this information by contacting FIFA, we did not receive any responses.

[37] The TV programs include live broadcasts and other related content (e.g., news programs about the games). The coverage includes games from both the qualification and final stages.

the population. Figure 5 shows the effects categorized by tertiles of TV audience proportions.[38] While the effects of losses were null, victories had larger impacts when a higher fraction of the population watched the games. These results support the observation in Figure 4 that the men's World Cup has a weaker effect because it involves countries where soccer is less popular.

Figure 5. Effect Heterogeneity by Viewership (2010 and 2014 World Cups)



The figure shows the effects of victories and losses in international soccer games on leader approval by the tertiles in the proportion of TV audience. The samples include only the 2010 and 2014 World Cup games. The tertiles are 0.35 and 0.72 in the sample of victories and other games (top left), and 0.31 and 0.72 in the sample of losses and other games (top right). The thick and thin intervals are 90% and 95% confidence intervals, respectively.

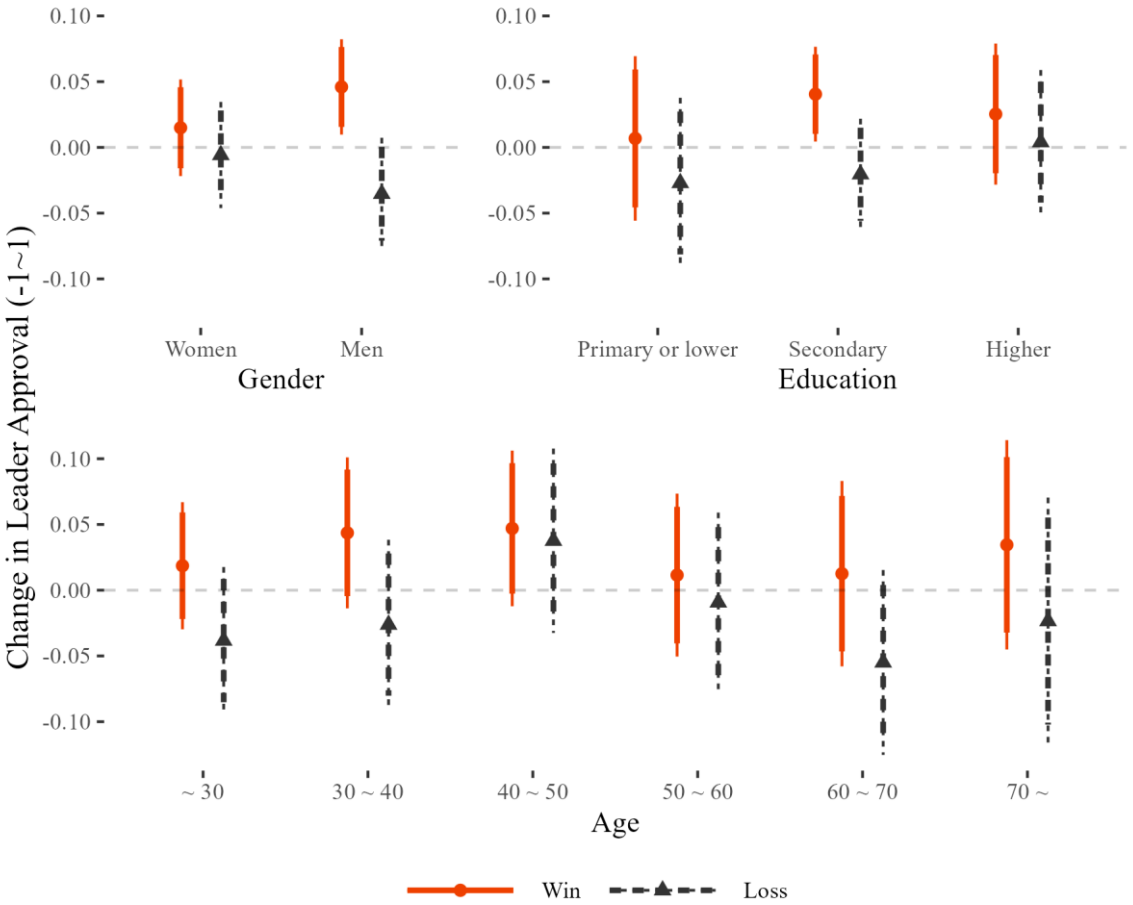
We also assess the plausibility of our findings by examining heterogeneity at the individual level. Figure 6 shows the heterogeneous effects by the demographic covariates of respondents; gender, education, and age.[39] As shown in the figure, the effect is larger among

[38] When we analyze effect heterogeneity by a continuous covariate (e.g., the proportion of TV audience), we follow the approach of Hainmueller et al. (2019) by binning the covariate into tertiles.

[39] The other demographic covariates (e.g., marital status, income, and religion) are less relevant and are therefore not reported. However, the results involving these variables are available upon request.

men and is somewhat larger among younger and middle-aged respondents (under 50 years old). These results are consistent with the observation that soccer games tend to be more popular among such demographic groups. In contrast, the effects are largely similar across different educational levels, suggesting that education may not necessarily facilitate the rational evaluation of a leader.

Figure 6. Effect Heterogeneity by Demographic Covariates



The figure shows the effects of victories and losses in international soccer games on leader approval by demographic covariates. The other demographic covariates (marriage, income, are religion) are less relevant and thus not reported. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

In Appendix A7, we also showed that the effects of victories were larger when the games involved historical rivals and thus captivating (Figure A 7-1); to a lesser extent, when the results were unexpected and thus surprising (Figure A 7-2); when the results were decisive and thus impressive (Figure A 7-3); and when a significant amount of time had passed since

the last games and thus the games were fresh (Figure A 7-4). The effects were consistent across playoff matches, group stages, and qualification games (Figure A 7-5).[40]

Heterogeneity II: Countries

Thus far, we have analyzed the *average* effects of international soccer games using samples from over 100 countries. However, this does not imply that the findings are equally applicable to each country; the effects can be heterogeneous across countries. To assess this possibility, we employ a random slope model;[41]

$$approval_i = \alpha_{win,k} + \beta_{win,1,k} after_i + \beta_{win,2,k} win_j + \left(\delta_{win} + \delta_{win,k}\right) after_i \, win_j + \varepsilon_{win,i} \quad (3)$$

$$approval_i = \alpha_{loss,k} + \beta_{loss,1,k} after_i + \beta_{loss,2,k} loss_j + \left(\delta_{loss} + \delta_{loss,k}\right) after_i \, loss_j + \varepsilon_{loss,i} \quad (4)$$

While $\delta_{\cdot}$ is an effect averaged over countries, $\delta_{\cdot,k}$ represents an effect specific to country k. The country-specific parameters are assumed to follow independent normal distributions. Note that several countries appear only in either Equation (3) or (4) as survey interviews did not coincide with instances of both winning and losing games.

Figure 7 shows the estimates of $\delta_{win,k}$ (horizontal axis) and $\delta_{loss,k}$ (vertical axis). Each flag image in the figure represents a country. When a country is included only in either

[40] There were more qualification games in the men's and women's World Cups and Asian Cup compared to the Euro, Africa Cup of Nations, Gold Cup, and Copa America. This disparity likely explains the small differences among the playoff, group-stage, and qualification games. Subsetting by both stages and tournaments leads to excessively small samples sizes.

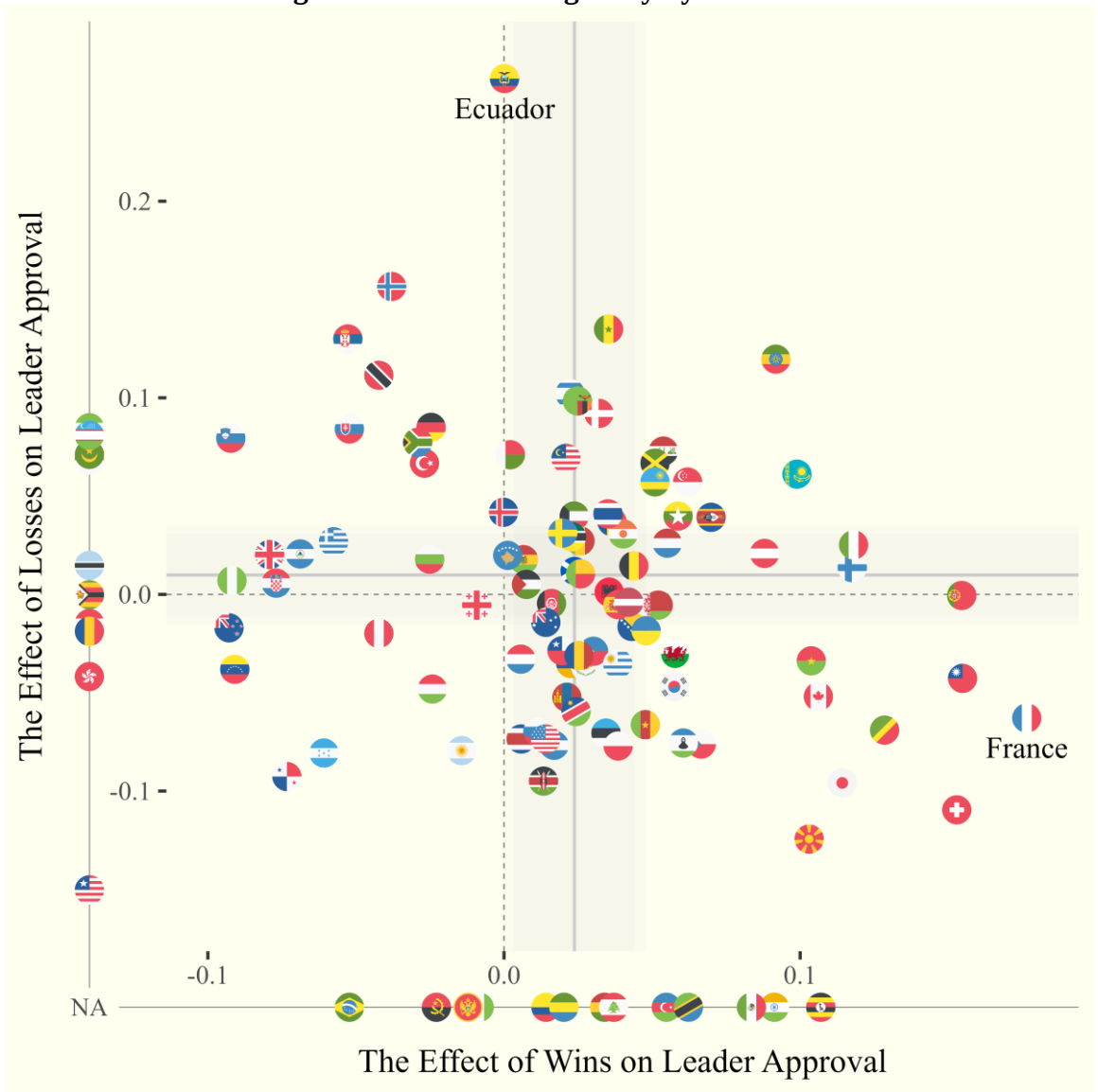
[41] We employ the random slope model rather than the fixed slope model to avoid overfitting. On average, the sample contains only three games per national team. For this analysis, we use the lme4 package in R (Bates 2010). Unfortunately, standard errors are not clustered by game, as lme4 and other packages (e.g., clubSandwitch) do not support clustering by a non-nested group. The coefficients are estimated with restricted maximum likelihood.

Equation (3) or (4), the country is displayed in the "NA" row or column. In addition, when the effects of victories or losses are statistically significant at the 1% level,[42] they are displayed with their country names. The gray lines and intervals are the estimates and confidence intervals for the average effects δ_{win} and δ_{loss}.

As seen in this figure, the effects vary substantially across countries. The standard deviations of the country-specific effects are 0.10 and 0.11 for victories and losses, respectively. These values are large compared to the average effects of victories and losses. Although it is beyond the scope of this paper to identify the causes of this heterogeneity across countries, these findings warn against overgeneralizing the results from single-country studies; the findings in the U.S. (Healy, Malhotra, and Mo 2010), Finland (Rapeli and Söderlund 2022), or Ireland (Müller and Kneafsey 2023) may not necessarily apply to other contexts.

[42] We employ the 1% level as the figure becomes too complicated with the 5% level.

Figure 7. Effect Heterogeneity by Countries



The figure shows the effects of victories (horizontal axis) and losses (vertical axis) in international soccer games on leader approval by countries. Each flag image corresponds to a country in the sample. When survey interviews did not coincide with both winning and losing games and thus a country is available in either Equation (3) or (4), the country is displayed in the NA row or column. When either the effect of victories or losses are statistically significant at the 1% level, their names are displayed (i.e., Ecuador and France; the figure becomes complicated with the 5% significance level). The light gray lines and intervals are the estimates and confidence intervals of the average effects.

Figure 7 also indicates both a representative case (France), where victories increased leader approval, and an outlier case (Ecuador), where losses *increased* leader approval. In Table A 8-1 of Appendix A8, we enumerate all soccer games involving these countries within our sample. The positive effect of losses in Ecuador might be coincidental; Ecuador lost two games held on Friday and Thursday but won a game on Tuesday. Thus, the treated respondents

were interviewed around weekends, while most of the control respondents were interviewed on weekdays. As people tend to evaluate leaders more positively on weekends, this timing could skew the estimates positively. Although this is a small-sample error and is unlikely to explain the main findings,[43] it cautions against overinterpreting the individual estimates in Figure 7. By contrast, the positive effect of victories in France cannot be explained by the weekday effects. France won two games held on Friday and Tuesday and lost two games held on Fridays. Even excluding the Tuesday game, the effect remains positive. Given the overwhelming popularity of soccer in France, it is not surprising that the victories significantly boosted support for the president.[44]

Macro-level Analysis

Finally, we extend our analysis to explore the substantive implications of our micro-level findings by examining the effects of international soccer games on electoral outcomes. To this end, we have compiled the results of national presidential and legislative elections held within three weeks before or after international soccer games.[45] We have obtained the first polling dates for these elections from the National Elections across Democracy and Autocracy dataset (NELDA; 1945-2020; Hyde and Marinov 2012)[46], excluding any elections where opposition parties were not allowed, voters lacked choices, or there were significant concerns about the

[43] Controlling for years, months, weekdays, and days of soccer games does not change our main results. See Table 4 in the main text, and Table A 4-1 and Table A 4-2 in Appendix A3 for more details.

[44] Interestingly, French presidents—Sarkozy, Hollande, and Macron—often played soccer in front of the media to make appeals to constituents.

[45] We analyze the effects at different times by conducting an event study.

[46] If an election involves multiple rounds, we split them to different observations.

freedom and fairness of the elections.[47] For presidential elections, the outcome variable is the vote share of an incumbent or, in cases where the incumbent did not run, the vote share of a candidate from the incumbent's party.[48] For legislative elections, the outcome variable is the vote share of the incumbent's party.[49] Our sample contains 91 elections across 38 countries (from 2004 to 2020).[50] Given the small sample size, we pool presidential and legislative elections[51] and control for $p_{j,win}$ and $p_{j.lose}$ rather than using matching. The observations are weighted by the number of valid votes to ensure that the results from small countries (e.g., Cape Verde) do not sway the overall results.[52] The standard errors are two-way clustered by country and soccer game.

As shown in Table 7, the results are consistent with our main findings; victories in international soccer games increased the vote shares of leaders or their parties in national

[47] If multiple soccer games were linked, we use a soccer game held closest to the election date for analysis. We exclude elections if they were held on the same day of soccer games and the games started between 8AM and 8PM local time.

[48] The observations are excluded if the incumbent did not belong to a party or if no candidate belonged to the incumbent's party.

[49] Same as footnote 48.

[50] The sample contains 74 games: 54 from the men's World Cup, 6 from the women's World Cup, 25 from the Euro, and 6 from the Africa Cup of Nations. The games from other tournaments were excluded as they were not held around the dates of national elections. See Appendix A9 for a list of countries included in the sample.

[51] The sample includes 34 presidential elections and 57 legislative elections.

[52] The weighting accounts for different variances across elections. See p.281-282 of Wooldridge (2009).

elections.[53] The effect size is substantively large; if a national election is held within three weeks after a lucky victory in an international soccer game, the vote shares of leaders or their parties increase by 20 percentage points. This estimate remains similar even with the inclusion of country-fixed effects. However, given the small sample size, we cannot rule out the possibility that the estimates are exaggerated or false positives.[54] Thus, we present these findings as suggestive evidence, leaving the task of further extending our analysis to voting behaviors to future research.

[53] The results of the event study are reported in Figure A 9-1 of Appendix A9.

[54] Although the estimates cease to be statistically significant with the fixed effects of election days, days of week, months, and years, the point estimate remains similar for victories. This is not surprising as the fixed effect model is equivalent to adding 52 dummies, and the sample contains only 91 observations.

Table 7. The Effects on Incumbent Vote Shares

	No FE		Country FE	
	1	2	3	4
Win × After	0.20**		0.19*	
	(0.06)		(0.09)	
Win	-0.13†		-0.17*	
	(0.07)		(0.07)	
Loss × After		-0.08		-0.07
		(0.06)		(0.08)
Loss		0.07		0.08
		(0.06)		(0.08)
After	-0.05	0.08†	-0.09*	0.03
	(0.04)	(0.04)	(0.03)	(0.08)
Pr(Win)	-0.24	-0.36	-0.09	0.08
	(0.30)	(0.30)	(0.26)	(0.32)
Pr(Loss)	-0.37	-0.47	-0.34	-0.13
	(0.37)	(0.38)	(0.32)	(0.39)
(Intercept)	0.67**	0.68*		
	(0.23)	(0.26)		
N	91	91	91	91

The table shows the effects of victories and losses on the vote shares of leaders or their parties. The sample contains 38 countries (2004-2020). The standard errors two-way clustered by countries and games are in parentheses. $** p < 0.01$; $* p < 0.05$; $\dagger p < 0.1$.

Discussion

This study provided the first global evidence for the relevance of irrelevant events—the random results of international soccer games influence how people evaluate their national leaders. Moreover, the effects are found to be asymmetric; while unlucky losses did not affect leader approval, lucky victories improved public attitudes toward leaders. Our extensive robustness checks indicate that such findings are unlikely to be false positives or results of a narrow forking path. The most plausible explanation for these findings is that people use the results of soccer games as a benchmark to evaluate a leader's job performance.

These findings initially appear consistent with those of Healy et al. (2010) and other studies (Graham et al. 2023b; Kikuta and Uesugi 2023) and contradictory to those of Fowler

and Montagnes (2015) and subsequent studies (Fowler and Montagnes 2023a, 2023b; Müller and Kneafsey 2023; Rapeli and Söderlund 2022). However, our findings should not be interpreted in such a dichotomized manner. We observed substantial heterogeneity across different countries. Given this heterogeneity, the effects critically depend on sample composition. Thus, it is not surprising that the effects vary over time (Graham et al. 2023a, 2023b) and across space (Müller and Kneafsey 2023; Rapeli and Söderlund 2022).

A crucial task for future studies is to examine under what conditions, not just whether, politically irrelevant events affect public opinion. Although the debate has largely focused on technical details, often assuming that the effect would be constant, our analysis demonstrates that the effects vary substantially across contexts. Given the heterogeneity across different contexts, a more meaningful approach would involve examining different settings (e.g., survey experiments), treatments (e.g., Olympic games), outcomes (e.g., voting and approval), units (e.g., individuals, constituencies, and countries), and time (e.g., history) and consider generalizability (Findley, Kikuta, and Denly 2021).

Although the cross-country heterogeneity can be explained by numerous factors, a few of them are particularly noteworthy. First, as prior research suggests, the effect of politically irrelevant events depends on leaders' credit claiming (Barron and McLaughlin 2024; Cruz and Schneider 2017; Grimmer, Messing, and Westwood 2012; Ramos and Sanz 2020). Leaders can signify their contributions by attending soccer games, participating in victory parades, and delivering public speeches (e.g., Ivory Coast President Ouattara at the beginning of this paper). The dynamics for credit claiming, in turn, vary across political regimes (e.g., autocracy and democracy), ideological orientations (e.g., right- and left-wing leaders), and social cleavages (e.g., leaders leveraging sports to foster national unity; Depetris-Chauvin, Durante, and Campante 2020).

Second, the impact of politically irrelevant events also depends on how they are framed (Carew, Noor, and Burns 2019; Chong and Druckman 2007; Druckman 2005). Even when leaders effectively claim their credit, the media can frame it differently; coverage may range from compliments on leaders' contributions to critiques against the politicization of sports. Framing, in turn, depends on press freedom, audience reach, and media partisanship. Previous studies have shown that autocratic governments strategically use international sports events and media coverage to obscure violent repression (Glaßel, Scharpf, and Edwards 2024; Scharpf, Glabel, and Edwards 2023). Future studies should examine how credit claiming, framing, and other dynamics explain the heterogeneous effects of politically irrelevant events.

Finally, it is equally important to explain the causal mechanisms in a more nuanced manner. Our findings imply that people did not simply act on their emotions. Instead, they followed an irrational yet sensible heuristic; they used their national team's lucky victories as a benchmark for evaluating the leader's performance. Thus, instead of categorizing voters as rational or irrational, future studies must explore various reasons for irrationality. For instance, people may simply follow their moods, or they may use irrelevant events as heuristics. There exists a spectrum of rationality, ranging from full rationality, limited rationality, and finally to emotional irrationality. This field must focus beyond polarized debates, acknowledge the spectrum of irrationality, and analyze the causal mechanisms further.

Reference List

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138(1): 1–35.

Achen, Christopher H., and Larry M. Bartels. 2017. Democracy for Realists *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton, NJ: Princeton University Press.

Achen, Christopher H., and Larry M. Bartels. 2018. “Statistics as If Politics Mattered: A Reply to Fowler and Hall.” *The Journal of Politics* 80(4): 1438–53.

Adekaiyaoja, Afolabi. 2024. “Analysis: Is Ivory Coast’s AFCON the Latest African ‘Sportswashing’ Case?” *Al Jazeera*. https://www.aljazeera.com/features/2024/1/17/analysis-is-ivory-coasts-afcon-the-latest-african-sportswashing-case (April 1, 2024).

Alfano, Vincenzo, and Salvatore Ercolano. 2023. “Unveiling the Real Madrid Effect: The Impact of Football-Related Acrimony on Elections.” *Economia Politica*.

Alrababa’h, Ala’, William Marble, Salma Mousa, and Alexandra A. Siegel. 2021. “Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes.” *American Political Science Review* 115(4): 1111–28.

Arnold, Richard. 2021. “Nationalism and Sport: A Review of the Field.” *Nationalities Papers* 49(1): 2–11.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2014. “Is Voter Competence Good for Voters? Information, Rationality, and Democratic Performance.” *American Political Science Review* 108(3): 565–87.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg. 2017. “Accountability and Information in Elections.” *American Economic Journal: Microeconomics* 9(2): 95–138.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg. 2018. “Learning about Voter Rationality.” *American Journal of Political Science* 62(1): 37–54.

Atkinson, Matthew D., Ryan D. Enos, and Seth J. Hill. 2009. “Candidate Faces and Election Outcomes: Is the Face–Vote Correlation Caused by Candidate Selection?” *Quarterly Journal of Political Science* 4(3): 229–49.

Auer, Daniel, Lea Portmann, and Thomas Tichelbaecker. 2023. “Electoral Discrimination, Party Rationale, and the Underrepresentation of Immigrant-Origin Politicians.” *American Journal of Political Science*.

Auer, Daniel, and Didier Ruedin. 2023. “How One Gesture Curbed Ethnic Discrimination.” *European Journal of Political Research* 62(3): 945–66.

Bagues, Manuel, and Berta Esteve-Volart. 2016. “Politicians’ Luck of the Draw: Evidence from the Spanish Christmas Lottery.” *Journal of Political Economy* 124(5): 1269–94.

Barron, Nathan T., and Peter T. McLaughlin. 2024. “Experimental Evidence of the Benefits and Risks of Credit Claiming and Pork Busting.” *Legislative Studies Quarterly* 49(1): 103–29.

Bassi, Anna. 2019. “Weather, Risk, and Voting: An Experimental Analysis of the Effect of Weather on Vote Choice.” *Journal of Experimental Political Science* 6(1): 17–32.

Bates, Douglas M. 2010. “Lme4: Mixed-Effects Modeling with R.” https://www.researchgate.net/profile/Dimitris-Kavroudakis/post/What_is_the_appropriate_package_to_use_for_performing_NML_using_R/attachment/59d62f0dc49f478072e9f5e7/AS%3A272534858076166%401441988781179/download/lrgprt.pdf (April 18, 2024).

Baumeister, Roy, Ellen Bratslavsky, Catrin Finkenauer, and K. Vohs. 2001. “Bad Is Stronger than Good.” *Review of General Psychology* 5.

Belschner, Jana. 2024. “Too Young to Win? Exploring the Sources of Age-Related Electoral Disadvantage.” *Electoral Studies* 88: 102748.

Berge, Laurent, Sebastian Krantz, Grant McDermott, and Maintainer Laurent Berge. 2020. “R Package Fixest.” http://cran.salud.gob.sv/web/packages/fixest/fixest.pdf (July 22, 2022).

Berggren, Niclas, Henrik Jordahl, and Panu Poutvaara. 2010. “The Looks of a Winner: Beauty and Electoral Success.” *Journal of Public Economics* 94(1–2): 8–15.

Bertoli, Andrew D. 2017. “Nationalism and Conflict: Lessons from International Sports.” *International Studies Quarterly* 61(4): 835–49.

Besley, Timothy. 2006. *Principled Agents? The Political Economy of Good Government*. Oxford University Press, USA.

Boen, Filip, Norbert Vanbeselaere, Mario Pandelaere, Siegfried Dewitte, Bart Duriez, Boris Snauwaert, Jos Feys, Vicky Dierckx, and Eddy Avermaet. 2002. “Politics and Basking-in-Reflected-Glory: A Field Study in Flanders.” *Basic and Applied Social Psychology* 24: 205–14.

Busby, Ethan C., and James N. Druckman. 2018. “Football and Public Opinion: A Partial Replication and Extension.” *Journal of Experimental Political Science* 5(1): 4–10.

Busby, Ethan C., James N. Druckman, and Alexandria Fredendall. 2016. “The Political Relevance of Irrelevant Events.” *The Journal of Politics* 79(1): 346–50.

Calvo, Ernesto, Marcelo Escolar, and Julia Pomares. 2009. “Ballot Design and Split Ticket Voting in Multiparty Systems: Experimental Evidence on Information Effects and Vote Choice.” *Electoral Studies* 28(2): 218–31.

Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. “Leadership and Pandering: A Theory of Executive Policymaking.” *American Journal of Political Science* 45(3): 532–50.

Card, David, and Gordon B. Dahl. 2011. “Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior.” *The Quarterly Journal of Economics* 126(1): 103–43.

Carew, Mark T., Masi Noor, and Jan Burns. 2019. “The Impact of Exposure to Media Coverage of the 2012 Paralympic Games on Mixed Physical Ability Interactions.” *Journal of Community & Applied Social Psychology* 29(2): 104–20.

Chen, Daniel L. 2016. IAST Working Papers *This Morning’s Breakfast, Last Night’s Game: Detecting Extraneous Factors in Judging*. Institute for Advanced Study in Toulouse (IAST). https://ideas.repec.org/p/tse/iastwp/31020.html (December 17, 2020).

Chen, Eric, Gábor Simonovits, Jon A. Krosnick, and Josh Pasek. 2014. “The Impact of Candidate Name Order on Election Outcomes in North Dakota.” *Electoral Studies* 35: 115–22.

Chong, Dennis, and James N. Druckman. 2007. “Framing Theory.” *Annual Review of Political Science* 10: 103–26.

Cialdini, Robert, Richard Borden, Avril Thorne, Marcus Walker, Stephen Freeman, and Lloyd Sloan. 1976. “Basking in Reflected Glory: Three (Football) Field Studies.” *Journal of Personality and Social Psychology* 34: 366–75.

Confederation of African Football. 2022. *Activity Report 2021-2022*. https://www.cafonline.com/media/sqsd0gxk/urv1nsswxbcjx8gjqrwe.pdf (April 2, 2024).

Cornil, Yann, and Pierre Chandon. 2013. “From Fan to Fat? Vicarious Losing Increases Unhealthy Eating, but Self-Affirmation Is an Effective Remedy.” *Psychological Science* 24(10): 1936–46.

Cruz, Cesi, and Christina J. Schneider. 2017. “Foreign Aid and Undeserved Credit Claiming.” *American Journal of Political Science* 61(2): 396–408.

Depetris-Chauvin, Emilio, Ruben Durante, and Filipe Campante. 2020. “Building Nations through Shared Experiences: Evidence from African Football.” *American Economic Review* 110(5): 1572–1602.

Druckman, James N. 2005. “Media Matter: How Newspapers and Television News Cover Campaigns and Influence Voters.” *Political Communication* 22(4): 463–81.

Edmans, Alex, Diego García, and Øyvind Norli. 2007. “Sports Sentiment and Stock Returns.” *The Journal of Finance* 62(4): 1967–98.

Eren, Ozkan, and Naci Mocan. 2018. “Emotional Judges and Unlucky Juveniles.” *American Economic Journal: Applied Economics* 10(3): 171–205.

Fair, C. Christine, Patrick M. Kuhn, Neil Malhotra, and Jacob N. Shapiro. 2017. “Natural Disasters and Political Engagement: Evidence from the 2010–11 Pakistani Floods.” *Quarterly Journal of Political Science* 12(1): 99–141.

Fearon, James D. 1999. “Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance.” In *Democracy, Accountability, and Representation*, Cambridge; New York: Cambridge University Press, 55–97.

Fearon, James D. 2011. “Self-Enforcing Democracy.” *The Quarterly Journal of Economics* 126(4): 1661–1708.

Ferejohn, John. 1986. “Incumbent Performance and Electoral Control.” *Public choice*: 5–25.

FIFA. 2018. “More than Half the World Watched Record-Breaking 2018 World Cup.” https://inside.fifa.com/tournaments/mens/worldcup/2018russia/media-releases/more-than-half-the-world-watched-record-breaking-2018-world-cup (April 2, 2024).

FIFA. 2022. “FIFA WORLD CUP QATAR 2023 IN NUMBERS.” https://publications.fifa.com/en/annual-report-2022/tournaments-and-events/fifa-world-cup-quatar-2022/fifa-world-cup-qatar-2022-in-numbers/ (April 2, 2024).

Findley, Michael G., Kyosuke Kikuta, and Michael Denly. 2021. “External Validity.” *Annual Review of Political Science* 24(1): 365–93.

Fowler, Anthony, and Andrew B. Hall. 2016. “Do Shark Attacks Influence Presidential Elections? Reassessing a Prominent Finding on Voter Competence.” *The Journal of Politics* 80(4): 1423–37.

Fowler, Anthony, and B. Pablo Montagnes. 2015. “College Football, Elections, and False-Positive Results in Observational Research.” *Proceedings of the National Academy of Sciences* 112(45): 13800–804.

Fowler, Anthony, and B. Pablo Montagnes. 2023a. “Distinguishing between False Positives and Genuine Results: The Case of Irrelevant Events and Elections.” *The Journal of Politics* 85(1): 304–9.

Fowler, Anthony, and B. Pablo Montagnes. 2023b. *On the Importance of Independent Evidence: A Reply to Graham et Al.* https://jop.blogs.uni-hamburg.de/on-the-importance-of-independent-evidence-a-reply-to-graham-et-al/ (March 20, 2024).

Fukumoto, Kentaro, and Hirofumi Miwa. 2018. “Share the Name, Share the Vote: A Natural Experiment of Name Recognition.” *The Journal of Politics* 80(2): 726–30.

Gallagher, J. 2021. *Retrospective Voting and Natural Disasters That Cause No Damage: Accounting for the Selective Reporting of Weather Damage*. https://www.semanticscholar.org/paper/Retrospective-Voting-and-Natural-Disasters-that-No-Gallagher/a1a7d166f4960bde38c6371ed59331167ea49d77 (March 20, 2024).

Gasper, John T., and Andrew Reeves. 2011. “Make It Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters.” *American Journal of Political Science* 55(2): 340–55.

Ge, Qi. 2018. “Sports Sentiment and Tipping Behavior.” *Journal of Economic Behavior & Organization* 145: 95–113.

Ge, Qi, Ignacio Sarmiento Barbieri, and Rodrigo Schneider. 2021. “Sporting Events, Emotional Cues, and Crime: Spatial and Temporal Evidence from Brazilian Soccer Games.” *Economic Inquiry* 59(1): 375–95.

Gkorezis, Panagiotis, Victoria Bellou, Despoina Xanthopoulou, Arnold B. Bakker, and Angelos Tsiftsis. 2016. “Linking Football Team Performance to Fans’ Work Engagement and Job Performance: Test of a Spillover Model.” *Journal of Occupational and Organizational Psychology* 89(4): 791–812.

Glaßel, Christian, Adam Scharpf, and Pearce Edwards. 2024. *Does Sportswashing Work? First Insights from the 2022 World Cup in Qatar*. http://bitly.ws/AyBE (March 20, 2024).

Goldsmith, Benjamin E., Yusaku Horiuchi, and Kelly Matush. 2021. “Does Public Diplomacy Sway Foreign Public Opinion? Identifying the Effect of High-Level Visits.” *American Political Science Review* 115(4): 1342–57.

Goodman-Bacon, Andrew. 2021. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics* 225(2): 254–77.

Graham, Matthew H., Gregory A. Huber, Neil Malhotra, and Cecilia Hyunjung Mo. 2023a. “How Should We Think about Replicating Observational Studies? A Reply to Fowler and Montagnes.” *The Journal of Politics* 85(1): 310–13.

Graham, Matthew H., Gregory A. Huber, Neil Malhotra, and Cecilia Hyunjung Mo. 2023b. “Irrelevant Events and Voting Behavior: Replications Using Principles from Open Science.” *The Journal of Politics* 85(1): 296–303.

Gries, Peter Hays, H. Michael Crowson, and Todd Sandel. 2010. “The Olympic Effect on American Attitudes towards China: Beyond Personality, Ideology, and Media Exposure.” *Journal of Contemporary China* 19(64): 213–31.

Grimmer, Justin, Solomon Messing, and Sean J. Westwood. 2012. “How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation.” *American Political Science Review* 106(4): 703–19.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. “How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice.” *Political Analysis* 27(2): 163–92.

Healy, Andrew, and Neil Malhotra. 2009. “Myopic Voters and Natural Disaster Policy.” *American Political Science Review* 103(3): 387–406.

Healy, Andrew, and Neil Malhotra. 2010. “Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence.” *Quarterly Journal of Political Science* 5(2): 193–208.

Healy, Andrew, Neil Malhotra, and Cecilia Hyunjung Mo. 2010. “Irrelevant Events Affect Voters’ Evaluations of Government Performance.” *Proceedings of the National Academy of Sciences* 107(29): 12804–9.

Healy, Andrew, Neil Malhotra, and Cecilia Hyunjung Mo. 2015. “Determining False-Positives Requires Considering the Totality of Evidence.” *Proceedings of the National Academy of Sciences* 112(48): E6591–E6591.

Heersink, Boris, Brenton D. Peterson, and Jeffery A. Jenkins. 2017. “Disasters and Elections: Estimating the Net Effect of Damage and Relief in Historical Perspective.” *Political Analysis* 25(2): 260–68.

Hirshleifer, David, and Tyler Shumway. 2003. “Good Day Sunshine: Stock Returns and the Weather.” *The Journal of Finance* 58(3): 1009–32.

Ho, Daniel E., and Kosuke Imai. 2008. “Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: The California Alphabet Lottery, 1978-2002.” *The Public Opinion Quarterly* 72(2): 216–40.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3): 199–236.

Hyde, Susan D., and Nikolay Marinov. 2012. "Which Elections Can Be Lost?" *Political Analysis* 20(2): 191–210.

Imai, Kosuke, and In Song Kim. 2019. "When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?" *American Journal of Political Science* 63(2): 467–90.

Jiménez, Bruno. 2023. "The Political Economy of the Minimum Wage." *Labour Economics* 85: 102463.

Jiménez, Luis F. 2016. "Politics on the Pitch: The Impact Political Regimes Have on FIFA World Cup Success at Younger and Senior Levels." *Social Science Quarterly* 97(5): 1295–1314.

Kalist, David E., and Daniel Y. Lee. 2016. "The National Football League: Does Crime Increase on Game Day?" *Journal of Sports Economics* 17(8): 863–82.

Kikuta, Kyosuke, and Mamoru Uesugi. 2023. "Do Politically Irrelevant Events Cause Conflict? The Cross-Continental Effects of European Professional Football on Protests in Africa." *International Organization* 77(1): 179–216.

Kim, Sang-Hyun, and Fernanda L Lopez de Leon. 2019. "In-Group and out-Group Biases in the Marketplace: A Field Experiment during the World Cup." *Oxford Economic Papers* 71(3): 528–47.

Kirby, Stuart, Brian Francis, and Rosalie O'Flaherty. 2014. "Can the FIFA World Cup Football (Soccer) Tournament Be Associated with an Increase in Domestic Abuse?" *Journal of Research in Crime and Delinquency* 51(3): 259–76.

Kitamura, Shuhei. 2022. “UFOs: The Political Economy of Unidentified Threats.” https://osf.io/tme8f/ (October 28, 2023).

Lazarev, Egor, Anton Sobolev, Irina V. Soboleva, and Boris Sokolov. 2014. “Trial by Fire: A Natural Disaster’s Impact on Support for the Authorities in Rural Russia.” *World Politics* 66(4): 641–68.

Leitner, Christoph, Achim Zeileis, and Kurt Hornik. 2010. “Forecasting Sports Tournaments by Ratings of (Prob)Abilities: A Comparison for the EURO 2008.” *International Journal of Forecasting* 26(3): 471–81.

Lohmann, Susanne. 1998. “An Information Rationale for the Power of Special Interests.” *American Political Science Review* 92(4): 809–27.

Lowe, Matt. 2021. “Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration.” *American Economic Review* 111(6): 1807–44.

Malhotra, Neil, and Alexander G. Kuo. 2008. “Attributing Blame: The Public’s Response to Hurricane Katrina.” *The Journal of Politics* 70(1): 120–35.

Marie, Olivier. 2016. “Police and Thieves in the Stadium: Measuring the (Multiple) Effects of Football Matches on Crime.” *Journal of the Royal Statistical Society. Series A: Statistics in Society* 179(1): 273–92.

Meier, Armando, Lukas Schmid, and Alois Stutzer. 2016. *Rain, Emotions and Voting for the Status Quo*. https://papers.ssrn.com/abstract=2868316 (April 10, 2017).

Miguel, Edward, Sebastián M. Saiegh, and Shanker Satyanath. 2011. “Civil War Exposure and Violence.” *Economics & Politics* 23(1): 59–73.

Miller, Dale T., and Michael Ross. 1975. “Self-Serving Biases in the Attribution of Causality: Fact or Fiction?” *Psychological Bulletin* 82(2): 213–25.

Miller, Michael K. 2013. “For the Win! The Effect of Professional Sports Records on Mayoral Elections.” *Social Science Quarterly* 94(1): 59–78.

Mo, Cecilia Hyunjung, Jon M. Jachimowicz, Jochen I. Menges, and Adam D. Galinsky. 2023. “The Impact of Incidental Environmental Factors on Vote Choice: Wind Speed Is Related to More Prevention-Focused Voting.” *Political Behavior*.

Mousa, Salma. 2020. “Building Social Cohesion between Christians and Muslims through Soccer in Post-ISIS Iraq.” *Science* 369(6505): 866–70.

Müller, Stefan, and Liam Kneafsey. 2023. “Evidence for the Irrelevance of Irrelevant Events.” *Political Science Research and Methods* 11(2): 311–27.

Muñoz, Jordi, Albert Falcó-Gimeno, and Enrique Hernández. 2020. “Unexpected Event during Survey Design: Promise and Pitfalls for Causal Inference.” *Political Analysis* 28(2): 186–206.

Muraoka, Taishi. 2021. “The Electoral Implications of Politically Irrelevant Cues under Demanding Electoral Systems.” *Political Science Research and Methods* 9(2): 312–26.

Nielsen Sports. 2018. *World Football Report 2018*. https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/world-football-report-2018.pdf (February 24, 2021).

Otto, A. Ross, Stephen M. Fleming, and Paul W. Glimcher. 2016. “Unexpected but Incidental Positive Outcomes Predict Real-World Gambling.” *Psychological Science* 27(3): 299–311.

Pinto, Gabriele. 2024. “Sports Nationalism and Xenophobia: When Cheering Turns into Violence.” *Journal of Peace Research*.

Priks, Mikael. 2010. “Does Frustration Lead to Violence? Evidence from the Swedish Hooligan Scene.” *Kyklos* 63(3): 450–60.

Ramos, Roberto, and Carlos Sanz. 2020. “Backing the Incumbent in Difficult Times: The Electoral Impact of Wildfires.” *Comparative Political Studies* 53(3–4): 469–99.

Rapeli, Lauri, and Peter Söderlund. 2022. “Does Sports Success Increase Government Support? Voter (Ir)Rationality in a Multiparty Context.” *Research & Politics* 9(3).

Reeves, Andrew. 2011. “Political Disaster: Unilateral Powers, Electoral Incentives, and Presidential Disaster Declarations.” *Journal of Politics* 73: 1142–51.

Remmer, Karen L. 2014. “Exogenous Shocks and Democratic Accountability Evidence From the Caribbean.” *Comparative Political Studies* 47(8): 1158–85.

Reynolds, Andrew, and Marco Steenbergen. 2006. “How the World Votes: The Political Consequences of Ballot Design, Innovation and Manipulation.” *Electoral Studies* 25(3): 570–98.

Rosenzweig, Leah R., and Yang-Yang Zhou. 2021. “Team and Nation: Sports, Nationalism, and Attitudes Toward Refugees.” *Comparative Political Studies* 54(12): 2123–54.

Sances, Michael W. 2017. “Attribution Errors in Federalist Systems: When Voters Punish the President for Local Tax Increases.” *The Journal of Politics* 79(4): 1286–1301.

Scharpf, Adam, Christian Glabel, and Pearce Edwards. 2023. “International Sports Events and Repression in Autocracies: Evidence from the 1978 FIFA World Cup.” *American Political Science Review* 117(3): 909–26.

Schwarz, Norbert, Fritz Strack, Detlev Kommer, and Dirk Wagner. 1987. “Soccer, Rooms, and the Quality of Your Life: Mood Effects on Judgments of Satisfaction with Life in General and with Specific Domains.” *European Journal of Social Psychology* 17(1): 69–79.

Seo, TaeJun, and Yusaku Horiuchi. 2023. “Natural Experiments of the Rally ’Round the Flag Effects Using Worldwide Surveys.” *Journal of Conflict Resolution*.

Simonsohn, Uri, Joseph P. Simmons, and Leif D. Nelson. 2020. “Specification Curve Analysis.” *Nature Human Behaviour* 4(11): 1208–14.

Todorov, Alexander, Anesu N. Mandisodza, Amir Goren, and Crystal C. Hall. 2005. "Inferences of Competence from Faces Predict Election Outcomes." *Science* 308(5728): 1623–26.

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics: A Modern Approach*. Boston, MA: Cengage Learning.

Wunderlich, Fabian, and Daniel Memmert. 2018. "The Betting Odds Rating System: Using Soccer Forecasts to Forecast Soccer." *PLOS ONE* 13(6).

Xu, Yiqing. 2023. "Causal Inference with Time-Series Cross-Sectional Data: A Reflection." In *The Oxford Handbook of Methodological Pluralism in Political Science*, eds. Janet M. Box-Steffensmeier, Dino P. Christenson, and Valeria Sinclair-Chapman. Oxford, New York: Oxford University Press.

Online Appendix for

"The Global Evidence for the Relevance of Irrelevant Events:

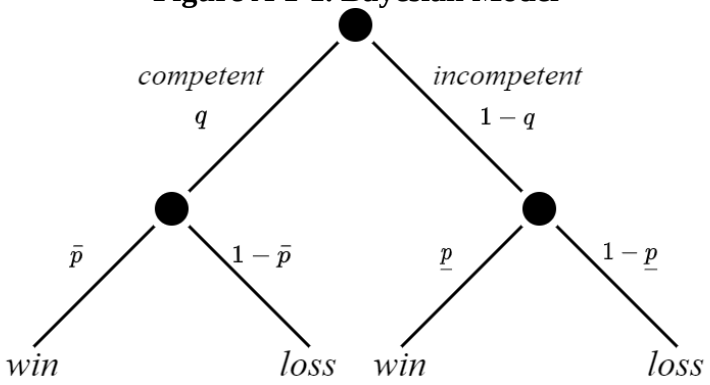
International Soccer Games and Leader Approval"

Table of Contents

A1. Theory: Bayesian Model

We use a stylized model to demonstrate that the random results of international soccer games should not update the belief of a rational voter when s/he has access to a more accurate indicator (e.g., FIFA ranking). To this end, suppose a situation where a voter evaluates the competence of a leader, who is competent with probability q and incompetent with $1-q$. When the leader is competent, his/her national team wins with probability $p = \overline{p}$. By contrast, when the leader is incompetent, the team wins with a lower probability, $p = \underline{p} < \overline{p}$.[1] The voter does not know whether the leader is competent and can only observe the result of the soccer game. For simplicity, we assume that the result of a soccer game is dichotomous (win or loss), but the conclusion is the same even when we consider a draw. Figure A 1-1 summarizes the model.

Figure A 1-1. Bayesian Model



Then, the soccer game updates the voter's belief when the voter lacks access to alternative indicators of the team's performance. The posterior beliefs after a win and loss are $\Pr(competent|win) = \frac{\overline{p}q}{\overline{p}q+\underline{p}(1-q)}$ and $\Pr(competent|loss) = \frac{(1-\overline{p})q}{(1-\overline{p})q+(1-\underline{p})(1-q)}$, respectively. It is shown that $\Pr(competent|win) > \Pr(competent|loss)$ and thus the voter more positively evaluates the leader's competence after a victory than after a loss. However, when the voter

[1] When the leader's competence does not affect the team's performance, $\overline{p} = \underline{p}$, and thus the soccer game does not update the voter's belief.

knows the team's average performance and thus can directly observe p, the soccer game does not update the voter's belief. Because the voter knows whether $p = \overline{p}$ or $\underline{p}$, s/he can accurately evaluate the leader's competence even without knowing the result of the soccer game.

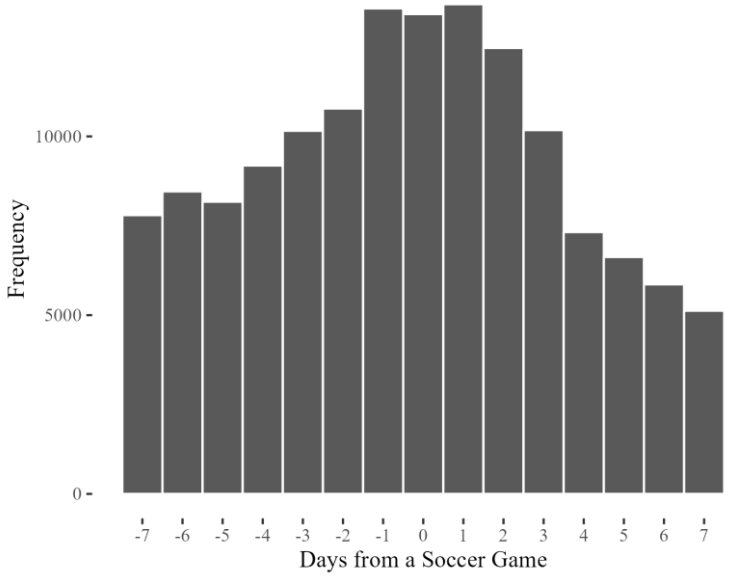
A2. Summary Statistics

Table A 2-1. Summary Statistics

	Mean	SD	Min	Median	Max	N	Histogram
Leader Approval	0.51	0.50	0.00	1.00	1.00	83915	
After	0.27	0.44	0.00	0.00	1.00	75298	
Win	0.42	0.49	0.00	0.00	1.00	90565	
Loss	0.38	0.49	0.00	0.00	1.00	90565	
Matched (Win)	0.52	0.50	0.00	1.00	1.00	90565	
Matched (Loss)	0.46	0.50	0.00	0.00	1.00	90565	
Female	0.54	0.50	0.00	1.00	1.00	90565	
Age	43.16	18.32	15.00	40.00	100.00	90561	
Educ. (2nd)	0.53	0.50	0.00	1.00	1.00	90565	
Educ. (higher)	0.19	0.40	0.00	0.00	1.00	90565	
Married	0.48	0.50	0.00	0.00	1.00	90560	
Income	2.73	0.96	1.00	3.00	4.00	88644	
Christian	0.65	0.48	0.00	1.00	1.00	90081	
Muslim	0.15	0.36	0.00	0.00	1.00	90081	

The table shows the summary statistics in the sample without matching. The main samples include observations for which the games are matched, and *Leader Approval* and *After* are available. *Leader Approval* is missing if it was not asked or respondents answered "don't know" or refused to answer. *After* is missing if respondents answered the survey between 8AM and 8PM on the days of the soccer games (indeterminate treatment status).

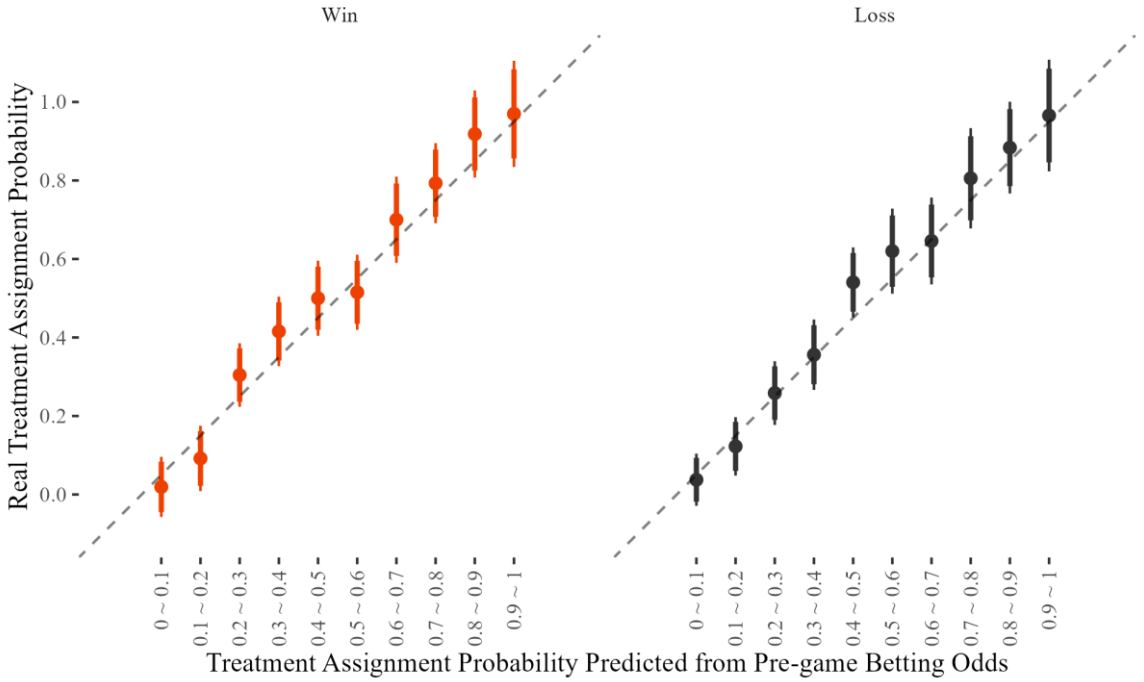
Figure A 2-1. Histogram of Interview Dates



The figure shows the histogram of interview dates relative to the dates of international soccer games. There are more respondents around the dates of soccer games as we focus on soccer games that happened to overlap the survey periods of the GWP.

A3. Validity Checks

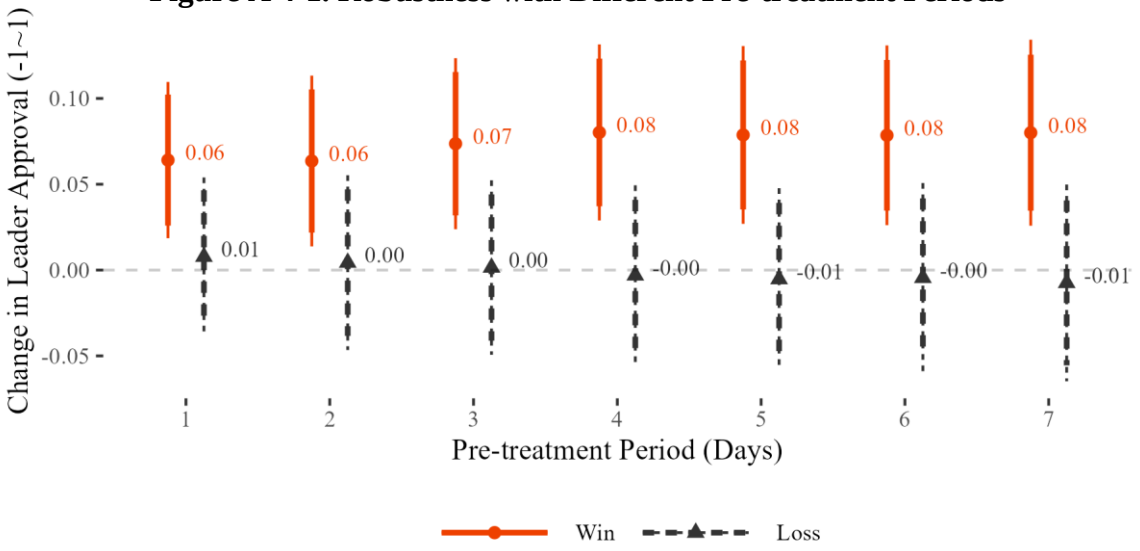
Figure A 3-1. Predictive Accuracy



The figure plots the proportions of winning and losing games for each range of probabilities of victories and losses predicted from the pre-game betting odds. The horizontal and vertical axes show the predicted and real probabilities of victories (left panel) and losses (right panel), respectively. As shown in the figure, the point estimates are close to the diagonal lines (dashed line), indicating the accuracy of the predictions.

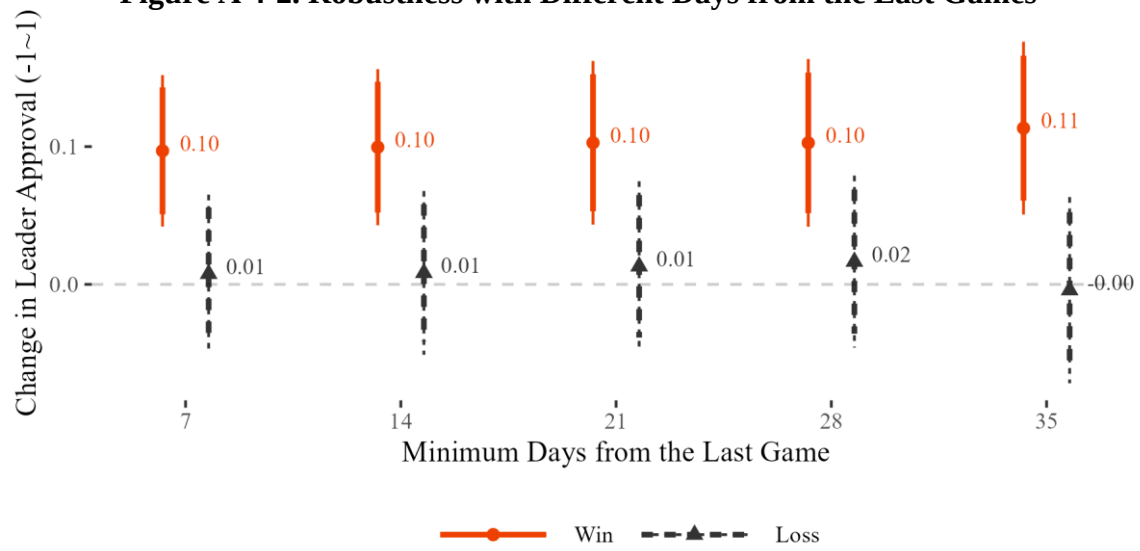
A4. Robustness Checks

Figure A 4-1. Robustness with Different Pre-treatment Periods



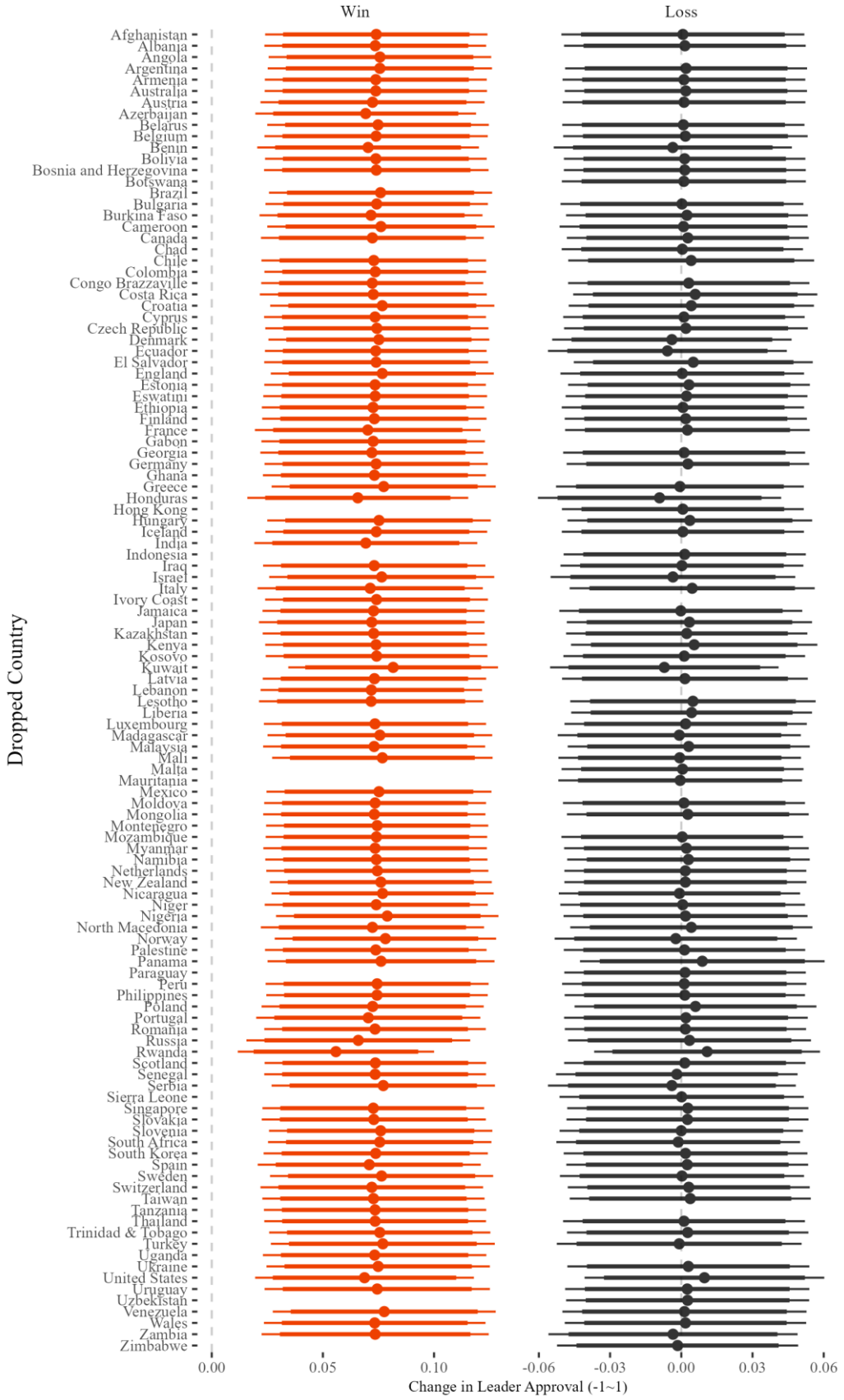
The figure shows the results when the pre-treatment time period is changed to those in the horizontal axis. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 4-2. Robustness with Different Days from the Last Games



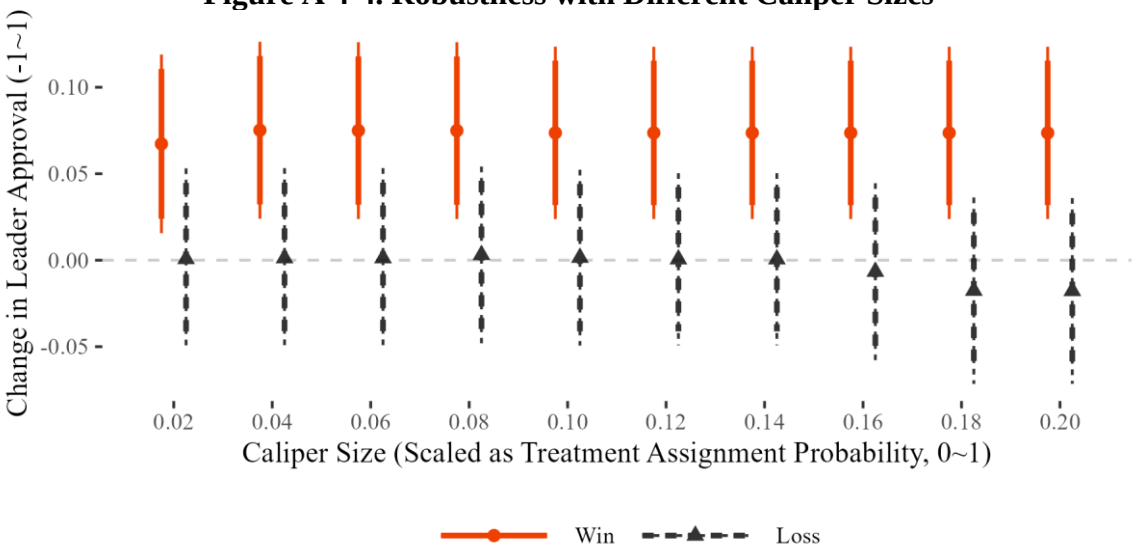
The figure shows the results when we drop the games that were held within h days after the last games. This accounts for potential effects of the last games. The horizontal axis shows the values of h. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 4-3. Leave-one-country-out Tests



The figure shows the results when we omit each country in the sample. The omitted countries are in the vertical axis. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 4-4. Robustness with Different Caliper Sizes



The figure shows the results when we use different caliper sizes in the matching. The caliper sizes are in the vertical axis. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Table A 4-1. Other Robustness Checks (Victories)

	Trichotomous Outcome			No Over-sampled Countries	Excluding t=0	Regression Control	Doubly Robust	Control for Covariates	Country-Game FE	Calendar FEs	Differential Time Trends	Two-way clustering
	Approve	DK/Refuse	Disapprove									
	1	2	3	4	5	6	7	8	9	10	11	12
Win × After	0.06*	0.03*	-0.08**	0.05*	0.08**	0.03†	0.08**	0.07**	0.03*	0.07**	0.06*	0.07**
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)
Win	-0.03	0.01	0.02	-0.02	-0.03	-0.01	-0.03	-0.03		-0.03	0.00	-0.02
	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
After	-0.05*	0.00	0.04**	-0.04*	-0.05**	-0.02	-0.05**	-0.04*	-0.01	-0.04*	-0.03	-0.05*
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Pr(Win)						0.35**	0.44†					
						(0.12)	(0.25)					
Pr(Loss)						0.46**	0.58†					
						(0.12)	(0.31)					
Female								0.00				
								(0.01)				
Age								0.00				
								(0.00)				
Educ. (2nd)								-0.11**				
								(0.02)				
Educ. (higher)								-0.10**				
								(0.02)				
Married								0.02†				
								(0.01)				
Income								0.08**				
								(0.01)				
Christian								0.00				
								(0.03)				
Muslim								0.03				
								(0.04)				
(Intercept)	0.47**	0.07**	0.46**	0.51**	0.51**	0.21*	0.14	0.34**			0.00	0.51**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.09)	(0.20)	(0.05)			(0.34)	(0.03)
N	39,057	39,057	39,057	32,432	34,273	69,749	36,037	35,453	36,037	36,037	60,537	36,037

$**\ p < 0.01$; $*\ p < 0.05$; $\dagger\ p < 0.1$. Model 1-3: The outcome variable is changed to a trichotomous variable, and "don't know" and refusals are not treated as missing values. Model 4: Over-sampled countries (Honduras, Italy, and Russia) are dropped. Model 5: All respondents interviewed on the day of soccer games are dropped. Model 6: Matching is not used, and the pre-game expectations of results ($p_{win,j}$ and $p_{loss,j}$) are controlled in regressions. Model 7: The matching is used as described in the manuscript, and the pre-game expectations of results ($p_{win,j}$ and $p_{loss,j}$) are also controlled in regressions. Model 8: The demographic covariates are controlled in regressions. Model 9: Country-game fixed effect is added. Model 10: Year, month, weekday, and day fixed effects are added. Model 11: The pre-treatment time trends are estimated separately for treated and control cases, and controlled in regressions. Model 12: The standard errors are two-way clustered by team and game.

Table A 4-2. Other Robustness Checks (Losses)

	Trichotomous Outcome			No Over-sampled Countries	Excluding t=0	Regression Control	Doubly Robust	Control for Covariates	Country-Game FE	Calendar FEs	Differential Time Trends	Two-way clustering
	Approve	DK/Refuse	Disapprove									
	1	2	3	4	5	6	7	8	9	10	11	12
Loss × After	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00	-0.01	0.01	0.00
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Loss	0.07*	0.00	-0.07*	0.07†	0.08*	0.05	0.08**	0.07*		0.07*	0.00	0.08*
	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)	(0.09)	(0.03)
After	-0.01	0.01	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Pr(Win)						0.36**	0.60*					
						(0.11)	(0.28)					
Pr(Loss)						0.42**	0.68*					
						(0.13)	(0.30)					
Female								0.01				
								(0.01)				
Age								0.00				
								(0.00)				
Educ. (2nd)								-0.09**				
								(0.02)				
Educ. (higher)								-0.10**				
								(0.03)				
Married								0.02*				
								(0.01)				
Income								0.08**				
								(0.01)				
Christian								0.03				
								(0.03)				
Muslim								0.10*				
								(0.04)				
(Intercept)	0.43**	0.07**	0.50**	0.48**	0.47**	0.20*	-0.01	0.27**			0.00	0.47**
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.09)	(0.21)	(0.04)			(0.55)	(0.03)
N	34,463	34,463	34,463	29,622	30,269	69,749	32,006	31,158	32,006	32,006	52,725	32,006

See the note of Table A 4-1.

A5. Specification Checks

Table A 5-1 shows the list of specifications used in the specification check (p.24), which mostly correspond to those in the robustness checks (see Table 3). The leave-one-country-out tests, and the analyses without Honduras, Italy, and Russia are not conducted as we did not consider those specifications before conducting the analysis and, thus, they do not constitute forking paths. We also exclude the analysis with two-way clustered standard errors as we added it based on feedback in a workshop (that said, the results are nearly identical and will not sway the results). The caliper sizes, pre-treatment periods, and days from the last games are thinned to save computational time.

Table A 5-1. List of Specifications

Outcome variable	1. "Don't know" and refusals as missing (default). 2. "Don't know" and refusals as 0.
Treatment variable	1. Some of the responses at $t = 0$ as non-missing (default).[1] 2. All responses at $t = 0$ as missing.
Method	1. Regression control for $p_{win,j}$ and $p_{loss,j}$. 2 ~ 4. Matching with calipers of .05, .1 (default), or .2. 5 ~ 7. Matching with calipers of .05, .1, or .2 and regression control for $p_{win,j}$ and $p_{loss,j}$.
Pre-treatment period	1. 1 day. 2. 3 days (default). 3. 5 days. 4. 7 days.
Subset	1. No subset (default). 2. No another game within a week. 3. No another game within three weeks.
Control	1. No control variables (default). 2. Control for demographic covariates.
Fixed effects	1. No fixed effect (default). 2. Team-game fixed effect. 3. Year, month, weekday, and day fixed effects. 4. Team-game, year, month, weekday, and day fixed effects.

Note 1: Respondents are assigned to a control group if a game started after 8PM, treated group if a game started before 8AM, and otherwise they are dropped.

A6. Mechanism Checks

Table A 6-1. Survey Items Used in the Mechanism Checks

Misperception	Leader: Do you approve or disapprove of the way the leader/head/President of this country is handling his/her job as leader/head/President? Opponent country: Do you approve or disapprove of the job performance of the leadership of [country name]? Mayor: Do you approve or disapprove of the leadership of the city or area where you live?
Mood	Did you experience the following feelings during a lot of the day yesterday? How about [enjoyment/sadness/anger]?
Immigrant	Please tell me whether you, personally, think each of the following is good thing or a bad thing? [Immigrants living in (respondent's country) / An immigrant becoming your neighbor / An immigrant marrying one of your close relatives].

All questions are binary choices with options for "don't know" and refusals. "don't know" and refusals are treated as missing.

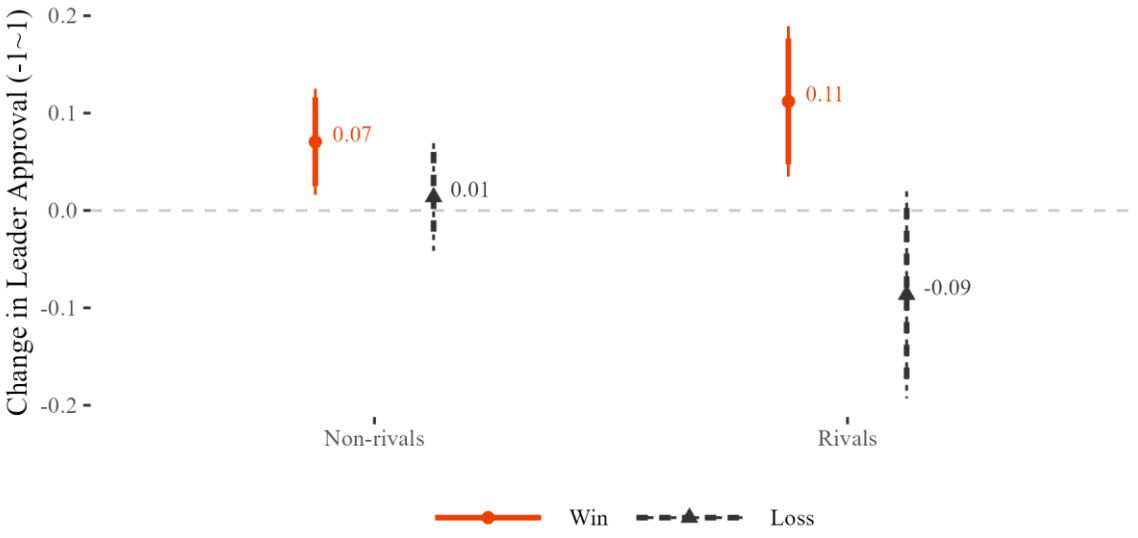
Table A 6-2. Mechanism Checks (Losses)

	Misperception			Mood			Immigrants		
	Leader	Opponent Country	Mayor	Enjoy	Anger	Sad	In my Country	Neighbor	Relatives Married
	1	2	3	4	5	6	7	8	9
Loss × After	0.02	0.09	0.04	0.00	0.00	-0.02	-0.06	-0.08	-0.10
	(0.03)	(0.08)	(0.04)	(0.02)	(0.01)	(0.02)	(0.08)	(0.08)	(0.07)
Loss	0.07†	0.20	0.05	-0.01	-0.01	-0.01	0.14†	0.13	0.10
	(0.04)	(0.12)	(0.04)	(0.02)	(0.01)	(0.01)	(0.08)	(0.08)	(0.09)
After	-0.02	-0.04	-0.02	0.01	0.01	0.02	0.00	-0.01	-0.01
	(0.02)	(0.05)	(0.03)	(0.02)	(0.01)	(0.02)	(0.06)	(0.06)	(0.05)
(Intercept)	0.54**	0.43**	0.56**	0.73**	0.19**	0.23**	0.52**	0.53**	0.51**
	(0.03)	(0.09)	(0.03)	(0.01)	(0.01)	(0.01)	(0.06)	(0.06)	(0.07)
N	28,266	1,905	10,886	30,004	30,115	30,123	4,914	4,904	4,853

The table shows the effects of losses on the outcome variables noted in the first row. The details of the survey items are available in Table A 6-1 of Appendix A6. The standard errors clustered by games are in parentheses. $** \; p < 0.01$; $* \; p < 0.05$; $\dagger \; p < 0.1$.

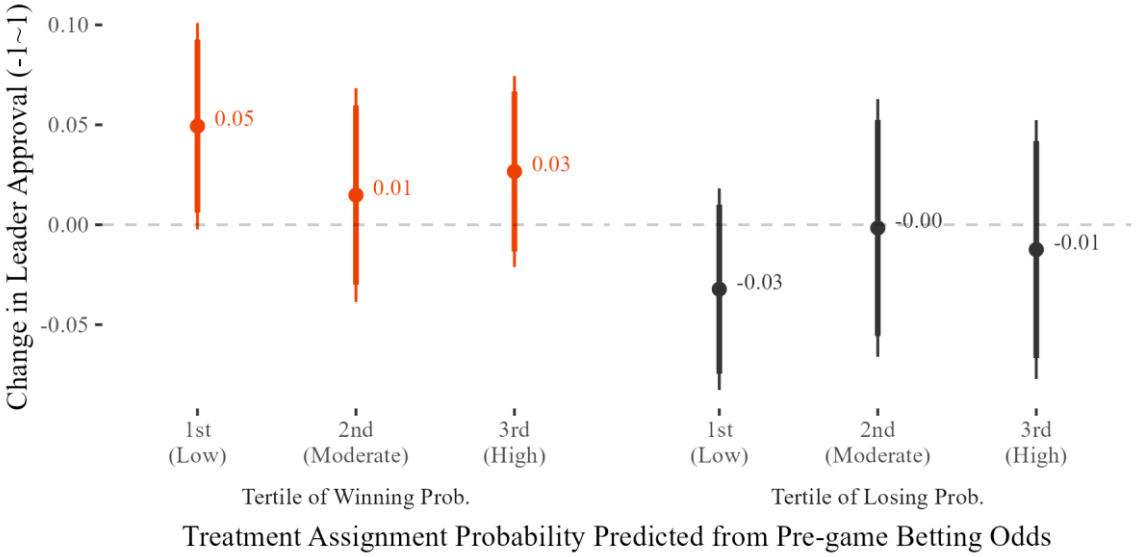
A7. Heterogeneity I: Salience

Figure A 7-1. Effect Heterogeneity by Soccer Rivalries



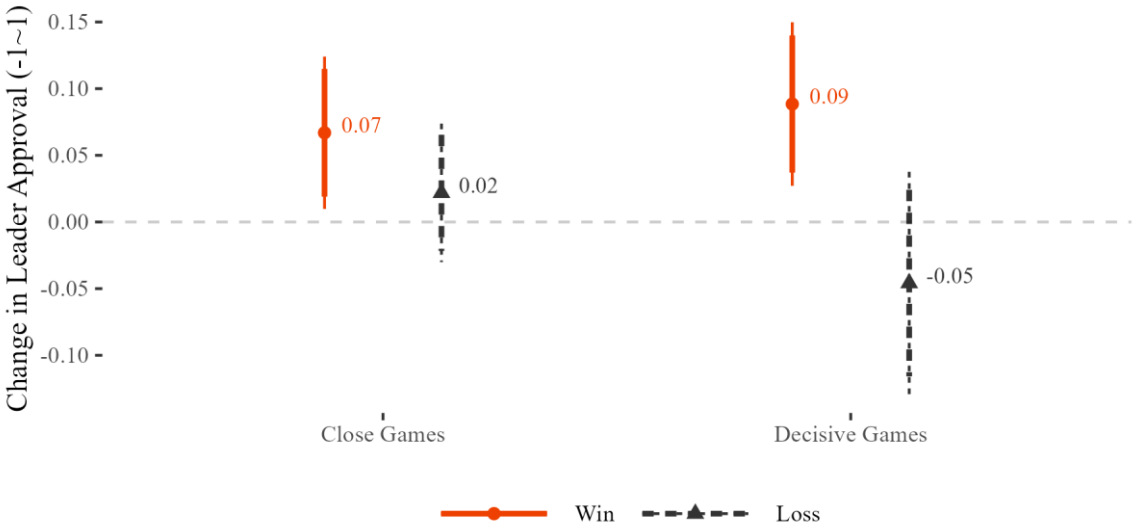
The figure shows the effects of victories and losses in international soccer games between historical rivals (right panel) and between other teams (left panel). The list of historical rivals in international soccer are derived from Wikipedia (https://web.archive.org/web/20240430024434/https://en.wikipedia.org/wiki/List_of_association_football_rivalries#International). To the best of our knowledge, this is the most comprehensive list of soccer rivalries. To ensure the reliability of information, we use the data only if there are references to the external sources or dedicated webpages. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 7-2. Effect Heterogeneity by Pre-game Expectation



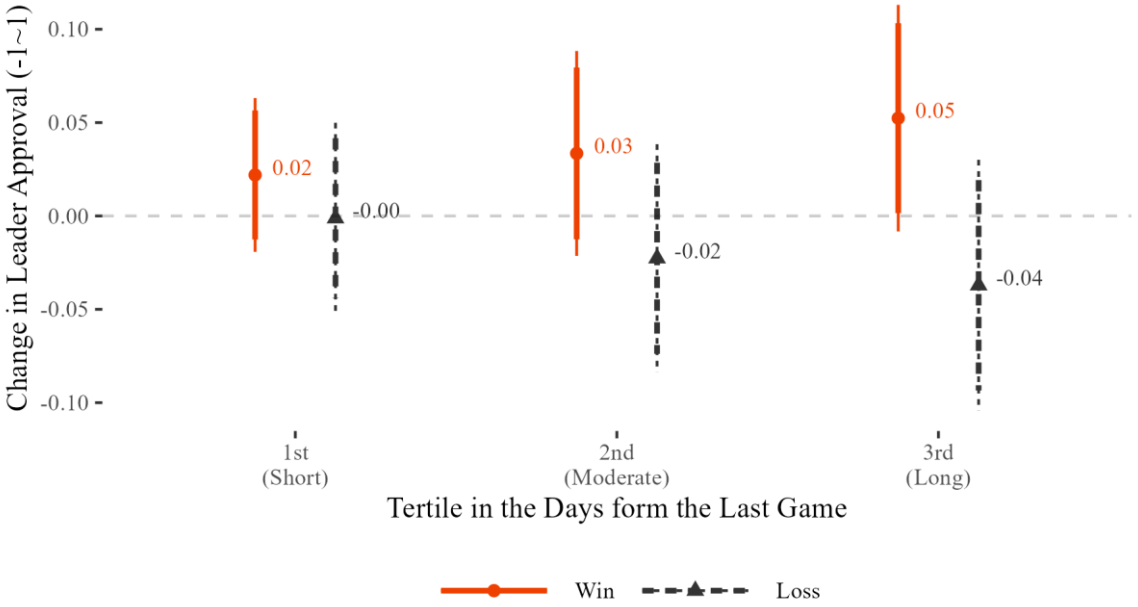
The figure shows the effects of victories and losses in international soccer games on leader approval by the tertiles of pre-game expectation of the results. The left panel shows the effects of victories from the low to high expectations of victories. The right panel shows the effects of losses from low to high expectations of losses. The pre-game expectations $p_{win,j}$ and $p_{loss,j}$ are derived from the pre-game betting odds. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 7-3. Effect Heterogeneity between Close and Decisive Games



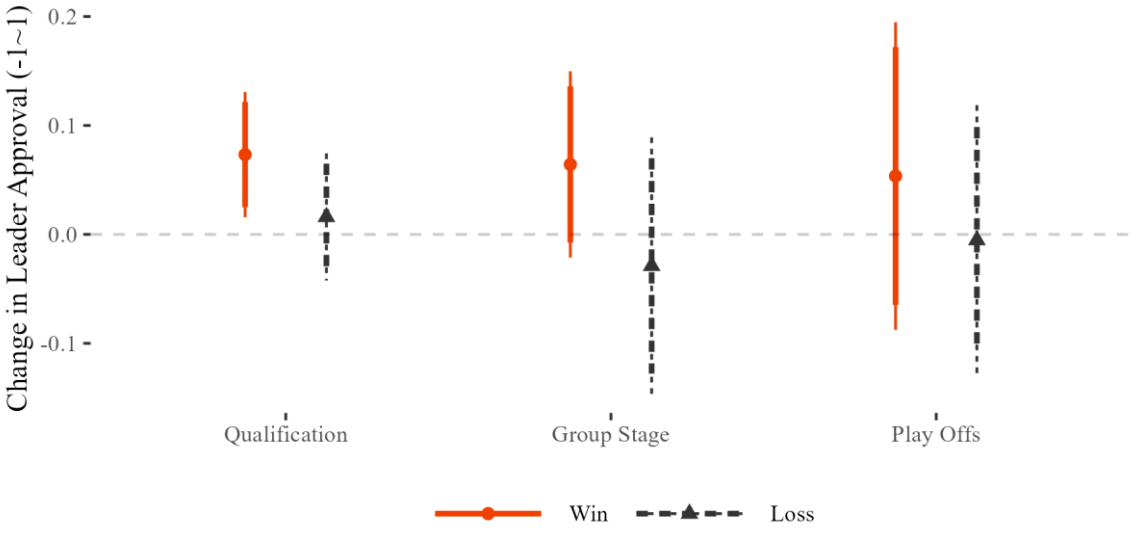
The figure shows the effects of victories and losses in close and decisive international soccer games on leader approval. When a final score difference is 1 or smaller, the game is considered close, and otherwise it is considered decisive. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 7-4. Effect Heterogeneity by the Days from the Last Game



The figure shows the effects of victories and losses in international soccer games on leader approval by the tertiles of the days from the last game. The thick and thin intervals are 90% and 95% confidence intervals, respectively.

Figure A 7-5. Effect Heterogeneity by Game Stages



The figure shows the effects of victories and losses in international soccer games on leader approval by the stages of the international soccer games (qualification, group, and playoff stages). The thick and thin intervals are 90% and 95% confidence intervals, respectively.

A8. Heterogeneity II: Countries

Table A 8-1. Ecuadorian and French Games in the Sample

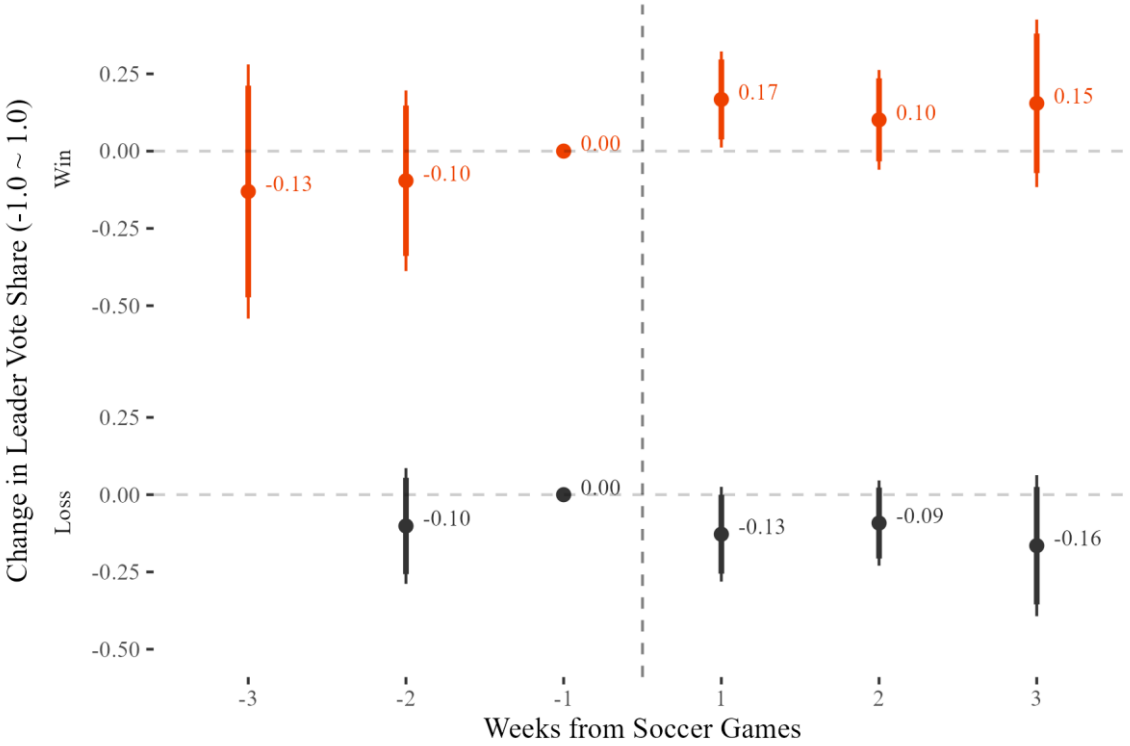
Date	Team	Opponent	Result	$p_{win,j}$	$p_{lose,j}$	Approval (before)	Approval (after)	Approval (diff.)	N
2011-11-11 (Fri)	Ecuador	Paraguay	1-2	0.16	0.59	0.65	0.81	0.17	318
2020-10-08 (Thu)	Ecuador	Argentina	0-1	0.08	0.77	0.15	0.20	0.05	90
2020-10-13 (Tue)	Ecuador	Uruguay	**4-2**	0.35	0.35	0.21	0.17	-0.05	20
2008-06-13 (Fri)	France	Netherland	1-4	0.34	0.35	0.40	0.40	0.00	270
2012-06-15 (Fri)	France	Ukraine	**2-0**	0.45	0.24	0.59	0.67	0.08	191
2017-06-09 (Fri)	France	Sweden	1-2	0.55	0.18	0.34	0.23	-0.11	155
2021-06-15 (Tue)	France	Germany	**1-0**	0.35	0.37	0.44	0.71	0.28	71

The table shows the soccer games involving the Ecuadorian and French national teams within the sample. These are the games whose dates coincided with the GWP surveys. The columns show the dates, teams, opponent teams, results, pre-game expectations, leader approval within three days before the game, leader approval within a day after the games, their differences, and the number of respondents.

A9. Macro-level Analysis: Event Study

The available countries in the sample are Argentina, Austria, Bolivia, Botswana, Bulgaria, Canada, Cape Verde, Croatia, Czech Republic, Denmark, Georgia, Ghana, Greece, Guatemala, Guyana, Iceland, Latvia, Liberia, Lithuania, Luxembourg, Madagascar, Morocco, Mozambique, New Zealand, Nicaragua, Poland, Portugal, Romania, Russia, Slovenia, Spain, Switzerland, Tanzania, Tunisia, Ukraine, U.S., Uruguay, and Venezuela.

Figure A 9-1. Event Study (Macro-level Analysis)



The figure shows the results of event study, where the average incumbent vote share within one week before international soccer games is compared to those in other time periods. The days from international soccer games are binned to weeks. The model includes country-fixed effects. The effect of losses three weeks before soccer games is missing as no election was held within the time window. The thick and thin intervals are 90% and 95% confidence intervals, respectively.