

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Campante, Filipe; Depetris-Chauvin, Emilio; Durante, Ruben

Working Paper "Building Nations through Shared Experiences: Evidence from African Football": A Reply to Delpeyrou and Bertoli (2024)

I4R Discussion Paper Series, No. 194

Provided in Cooperation with:

The Institute for Replication (I4R)

*Suggested Citation:* Campante, Filipe; Depetris-Chauvin, Emilio; Durante, Ruben (2024) : "Building Nations through Shared Experiences: Evidence from African Football": A Reply to Delpeyrou and Bertoli (2024), I4R Discussion Paper Series, No. 194, Institute for Replication (I4R), s.l.

This Version is available at: https://hdl.handle.net/10419/307962

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# **INSTITUTE** for

No. 194 DISCUSSION PAPER SERIES

# "Building Nations through Shared Experiences: Evidence from African Football": A Reply to Delpeyrou and Bertoli (2024)

Filipe Campante Emilio Depetris-Chauvin Ruben Durante

# This paper responds to:

Delpeyrou, Léonie, and Simone Bertoli. 2024. "Building Nations through Shared Experiences: Evidence from African Football": a Replication Depetris-Chauvin, Durante and Campante (2020). *IAR Discussion Paper Series No.* 193. Institute for Replication.

December 2024



# **I4R DISCUSSION PAPER SERIES**

I4R DP No. 194

# "Building Nations through Shared Experiences: Evidence from African Football": A Reply to Delpeyrou and Bertoli (2024)

Filipe Campante<sup>1,2</sup>, Emilio Depetris-Chauvin<sup>3</sup>, Ruben Durante<sup>4,5,6,7,8</sup>

<sup>1</sup>Johns Hopkins University, Baltimore/USA <sup>2</sup>National Bureau of Economic Research <sup>(</sup>NBER), Cambridge/USA <sup>3</sup>Pontificia Universidad Católica de Chile, Santiago/Chile <sup>4</sup>National University of Singapore <sup>5</sup>Universitat Pompeu Fabra (ICREA-UPF), Barcelona/Spain <sup>6</sup>CESifo, Munich/Germany <sup>7</sup>IZA – Institute of Labor Economics, Bonn/Germany <sup>8</sup>Centre for Economic Policy Research (CEPR), London/Great Britian

DECEMBER 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and metascientific work in the social sciences. Provided in cooperation with EconStor, a service of the <u>ZBW – Leibniz Information Centre for Economics</u>, and <u>RWI – Leibniz Institute for Economic Research</u>, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

		Eultors	
Abel Brodeur University of Ottawa	Anna Dreber Stockholm Schoo	ol of Economics	Jörg Ankel-Peters RWI – Leibniz Institute for Economic Research
E-Mail: joerg.peters@rwi-essen.de RWI – Leibniz Institute for Econom	ic Research	Hohenzollernstraße 1-3 45128 Essen/Germany	www.i4replication.org

# "Building Nations through Shared Experiences: Evidence from African Football": A Reply to Delpeyrou and Bertoli (2024)\*

Filipe Campante<sup>†</sup> Emilio Depetris-Chauvin<sup>‡</sup> Ruben Durante<sup>§</sup>

August 2024

### Abstract

In their comment, Delpeyrou and Bertoli (2024) make several critiques of our paper (Depetris-Chauvin et al., 2020). In this reply, we show that: (i) Their purported "errors" in data construction are in the vast majority of cases simply the result of different (and, we argue, superior) choices in implementing our identification strategy; (ii) Our specification choices regarding fixed effects are not only justified but are the correct ones for a proper apples-to-apples comparison; and most crucially, (iii) Our key results are essentially robust to their own proposed modifications. Far from overturning our findings, their critique further underscores their robustness.

<sup>\*</sup>We thank Simone Bertoli and Léonie Delpeyrou for their thorough effort in engaging with our work, and for sharing their comment with us in advance of submission. In spite of our strong disagreement with how their results are presented, their effort has helped us improve some aspects of our original work. We also thank co-editor Rema Hanna for her thorough handling of the process, and Guillermo Mondragón for very capable research assistance. All remaining errors are our own.

<sup>&</sup>lt;sup>†</sup>Johns Hopkins University & NBER. Email: fcampante@jhu.edu

<sup>&</sup>lt;sup>‡</sup>Instituto de Economía, Pontificia Universidad Católica de Chile. Email: edepetris@uc.cl

<sup>&</sup>lt;sup>§</sup>National University of Singapore, ICREA-UPF, CESifo, IZA & CEPR. Email: rubendurante@gmail.com

### 1. INTRODUCTION

In a recent comment, Delpeyrou and Bertoli (2024) (henceforth DB) make a number of critiques of our paper (Depetris-Chauvin et al., 2020). Unsurprisingly, given the incentives that often prevail for this type of work, they make some bold claims, indicating that our paper had "major problems," including a number of "errors" that presumably overturn key results. In this reply, we show that these claims are unfortunately built on exaggerations along many dimensions, and fundamentally unfounded. Far from overturning our results, DB end up underscoring their robustness, by showing that they largely survive different modeling and data choices, which happen to be less compelling than the ones that we made in our original contribution.

That said, DB's effort did uncover a few legitimate issues with the original data, and more broadly, it pushed us to provide a more thorough and systematic framework for our sampling procedure, which in hindsight we wish we had done in our original paper. We are in fact sincerely grateful for their help in fixing those issues, and for helping us establish the robustness of our results on an even more solid footing.

Our original analysis is composed of two parts: the individual-level analysis using Afrobarometer data, and the country-level analysis using conflict data. DB claim to have uncovered "major problems" in both of them, so we consider both sets of claims in order.

### 2. INDIVIDUAL-LEVEL ANALYSIS

Before delving into our general take on their comment, we would like to thank DB for spotting two valid issues in the original data used for our individual-level analysis (mentioned in their footnote 8). The first concerns the date of the South Africa vs. Burundi match, played on October 13, 2002 rather than on October 12, 2022. The second concerns the score of the Botswana vs. Mozambique match, played on October 11, 2008, which was 0-1 rather than 1-0. With regard to the first instance, we have verified that the date of October 12, 2022 was reported in the original file we received from the FIFA statistical office; however, other online sources confirm that the actual date of the match was October 13, 2002. The second instance was instead due to a coding oversight on our part, which we regret and for which we take responsibility. We have redone our analysis correcting for both FIFA's and our mistakes. Tables B.1 - B.3 in the Appendix to this reply, the corrections result only in very minor changes in the estimated coefficients. This is reassuring, although not surprising, since our original analysis already checked the robustness of the results to excluding individual matches.

We now turn to DB's broader critique of our individual-level analysis. DB's main point is that the effect of national teams' victories is identified by a limited number of matches – an aspect that was clearly acknowledged in the original article. From this observation they go on to question the validity and solidity of some of our results.

DB discuss three choices that influence the number of effective observations: the inclusion of supposedly too demanding sets of fixed effects; the focus on individuals affected only by one match; and the focus on a limited time-window around each match. Since they are especially critical of the choice of fixed effects, in what follows we discuss this issue in detail, respond to DB's criticism of specific results, and clarify what a limited number of identifying matches may imply for our results.

Our baseline econometric specification includes *country*  $\times$  *match* fixed effects and *ethnic group*  $\times$  *year* fixed effects. Including these fixed effects obviously reduces the number of effective observations. However, it is crucial to make the comparison between treated and control individuals more compelling. Specifically, including *country*  $\times$  *match* FEs ensures that we are comparing respondents from the same country interviewed around the same period, while including *ethnic group*  $\times$  *year* FEs ensures that we are comparing members of the same ethnic group interviewed in the same year.

The inclusion of *country*  $\times$  *match* fixed effects does not appear to substantially affect the results, as DB recognize. DB are instead very critical of the inclusion of *ethnic group*  $\times$  *year* FEs which, they argue "exerts a major influence on the estimates" and represents "a major problem". They also point out that the choice of including those fixed effects was not explicitly motivated in the original paper. In hindsight, we perhaps did not emphasize this motivation explicitly because we thought it would be obvious to most. Let us spell it out here.

Conceptually speaking, our ideal experiment would be to interview the same individuals twice, before and after a national team match. This is however not possible, since the Afrobarometer data do not have a panel dimension. To best approximate the ideal experiment, it is important to compare individuals belonging to the same ethic group interviewed in the same year. Indeed, as DB themselves recognize in footnote 21, "the relative importance of ethnic versus national identity can clearly vary across language groups, and possibly even over time". This is indeed the case in the Afrobarometer data. Looking at the sample of responses used in our original analysis, the between-group standard deviation in ethnic vs. national identity is 0.17, while the overall standard deviation is 0.34. That about one-fourth of the variation in the outcome variable is between groups confirms that time-varying differences between ethnic groups must be controlled for in order to correctly gauge the impact

of national teams' performance. At the same time, the substantial differences within ethnic groups are reassuring that there is enough identifying variation even when *ethnic group*  $\times$  *year* fixed effects are controlled for.

The fact that when *ethnic group*  $\times$  *year* FEs are *not* included both the magnitude and the precision of the coefficient of interest change substantially – as DB report on page 12, and which was already clear from Appendix Table A.8 in the original paper – only confirms this intuition. In other words, this indicates that comparing "apples to apples" uncovers a difference that comparing "apples to oranges" would have obscured.

All the above said, it is worth noting that our original analysis already addressed issues of robustness and sensitivity to these modeling choices and data limitations.<sup>1</sup> In particular, Appendix Table A8 in our original manuscript showed that the results remain largely unchanged when *country* × *match* FEs are not included (column 3), and when both sets of fixed effects are not included (column 2). It also showed that results are robust to including *country* × *year* FEs instead of *country* × *match* FEs, a less demanding approach that we employ when looking at the larger sample of individuals affected by one or multiple matches (Appendix Table A6) considering as treatment variable the share of matches won or the share of available points won by the national team.<sup>2</sup>

That the results are robust to excluding *country*  $\times$  *match* FEs is especially important because it is the inclusion of these fixed effects that is most constraining in terms of the sources of variation. Indeed, while their inclusion contributes to an even cleaner identification – i.e., by comparing individuals from the same country surveyed around the same match – it does impose a steeper trade-off in terms of the number of identifying matches and victories.

Consider instead what happens when we only include the *ethnic group*  $\times$  *year* FEs – which, as argued above, are absolutely crucial for a proper "apples-to-apples" comparison. This leaves us with 39 matches, 15 of which provide variation to identify the post-victory dummy – i.e., the number of victories occurring in years in which individuals of a same ethnic group were interviewed in the winning country – which is substantially more than the 9 matches that DB report as relevant under the full set of fixed effects.<sup>3</sup> First of all, as shown in Table B.4, the sample is properly balanced between treatment and control groups – in fact even more finely balanced than when *country*  $\times$  *match* FEs are included (see Table 1 in the original paper). Second, our main results all hold under this specification, as shown in Tables

6

<sup>&</sup>lt;sup>1</sup> The online appendix for our original manuscript can be found at: https://doi.org/10.1257/aer.20180805.

<sup>&</sup>lt;sup>2</sup> Our analysis also dealt with the threat of overly influential observations, given the small effective sample. In particular, we showed that the results are robust to excluding specific Afrobarometer waves and countries one at a time (Appendix Figure A3), and to excluding country-matches one at a time (Appendix Figure A4).

<sup>&</sup>lt;sup>3</sup> These 39 matches are linked to 210 language-year pairs in 32 country-years.

**B.5-B.7.**<sup>4</sup> In addition, following up on Appendix Table A6 from the original paper, without *country*  $\times$  *match* FEs we can exploit variation coming from individuals exposed to multiple matches. Tables **B.8** and **B.9** shows that the results from the paper are reproduced using the revised data with the aforementioned corrections. They also remain largely unaltered when only exploiting the variation within *ethnic group*  $\times$  *year*.<sup>5</sup> The bottom line is clear: even if one chooses to dial the trade-off away from cleaner identification to incorporate more variation, all of our key results remain unaltered.

With all of this in mind, let us still consider DB's claims on their own terms – namely, the idea that the low number of effective observations in our preferred specification should affect the interpretation of specific results. The first point that they claim is "called into question" is the distinction between the effects of victories versus draws/defeats on ethnic vs. national identities. While the point about draws (as a separate category) is well-taken, when it comes to defeats, it is worth noting that our estimate (Table 2, Column 4) is -0.000, with a standard error of 0.017– contrasted with -0.053 and 0.017 for victories. Even considering power issues, it does not seem unwarranted to interpret this as suggestive that victories seem to have a distinctive effect when compared to defeats, as we do in the original paper.

Their second point is that our statement that the result persists over time is somehow belied by the fact that the largest point estimate for the coefficient of the post-victory dummy in Figure 2 corresponds to 13 to 15 days after the match," and that that particular time window has very few identifying observations. However, even a cursory look at Figure 2 shows that the conclusion about persistence is entirely unaffected by ignoring that last window, which in fact is the largest point estimate" only by a marginal amount. Relatedly, Table A.9 in the Online Appendix shows that the estimated effect is if anything larger when using a shorter time window of 5 days.

Their third point pertains to the results on rivalry matches having a stronger effect. As highlighted in the quote mentioned by DB themselves, our original paper was very explicit in pointing out that this result was relying on a small number of observations. The same is true for the fourth point, about inter-ethnic trust and trust/dislike for foreigners, where our original language is heavily caveated in the same direction.

Finally, it is important to clarify how the limited number of identifying matches should affect the interpretation our results. DB argue that this issue entails a problem of identification and,

<sup>&</sup>lt;sup>4</sup> The only exception is our ancillary result regarding attitudes towards neighbors from other ethnicities (column 3 in Table B.7).

<sup>&</sup>lt;sup>5</sup> Interestingly, in their footnote 3, DB note in passing that in Table A.6 of the original paper we do present the results without *country*  $\times$  *match* FEs. However, they do not point out that this draws variation from more matches, and seem to imply that respondents interviewed before matches should not be coded as 0. It should be obvious that these individuals are indeed part of the relevant control group, and should be coded as such.

in this regard, refer to the paper by Roth et al. (2023). However, the same paper clearly states that this situation primarily affects inference.<sup>6</sup> As Roth et al. (2023) discuss, however, our clustering approach is valid (and potentially conservative), assuming, as it is plausible, that the treatment is independently determined at the country-match level.

To sum up, besides pointing out a couple of data corrections that do not affect any of the results in the paper, the main contribution of DB's effort is to stress the point that the estimates for the individual-level analysis come from a relatively small number of matches, something that was already repeatedly mentioned in the original paper. We have argued that our approach rests on modeling choices that are entirely justified and appropriate, and more compelling than the alternative approach suggested by DB. We have also shown that the results are robust to tweaking those modeling choices in ways that increase the amount of variation available at the cost of a less clean identification. On top of that, even in DB's own terms, what they portray as "major problems" are, in fact, caveats that were abundantly acknowledged in the original paper, and that pertain to ancillary results that were explicitly framed as suggestive evidence.

### 3. COUNTRY-LEVEL ANALYSIS

We must first acknowledge that DB are correct in pointing out that the description of the procedure for choosing which countries to include in our analysis was not sufficiently clear in our original paper. We are sincerely grateful to them for that. It behooves us to provide the clarity that had been lacking and, in the process of so doing, explain why what they describe as "errors" are, in the vast majority of cases, the product of a different and more suitable criterion for inclusion.<sup>7</sup> As we will see, the comparison across results not only underscores the robustness of our findings, but also helps illustrate the intuition behind our empirical strategy – and why DB's approach is less appealing, even while failing to meaningfully affect our results.

 $<sup>^{6}</sup>$  In particular, following the design-based approach in Rambachan and Roth (2022), i.e. "[treating] the units in the population (and their potential outcomes) as fixed rather than drawn from an infinite super-population (...) the usual DiD estimator is unbiased for a finite-population analog to the ATT under a finite-population analog to the parallel trends assumption."

<sup>&</sup>lt;sup>7</sup> That said, this exercise has uncovered a small number (four) cases that should indeed have been excluded from the analysis, and we are again sincerely grateful to DB for helping us spot them. Unsurprisingly – given that, as we will show, even the 41 changes posited by DB have a relatively small impact on our analysis – our results survive unscathed once these four cases are excluded. We discuss these cases below.

### 3.1. WHAT COUNTS AS "BARELY"?

The key endogeneity concern in the context of our analysis is that countries whose national football teams succeed in qualifying to a major tournament, such as the Africa Cup of Nations (ACN) or the World Cup, will be systematically different from those whose teams fail to qualify, in ways that correlate with the subsequent likelihood of conflict. The idea behind our identification strategy is to compare countries that "barely" qualify to those that "barely" fail to do so, with the assumption that what determines qualification or failure, within that set, is as good as random.

Our original analysis, as rightly pointed out by DB, fails to clearly define what counts as "barely" with sufficient precision. In response to that, DB's proposed approach, in a nutshell, is to consider as the relevant threshold a probability of qualification or elimination equal to zero: each and every country that, going into the last round of matches, has a non-zero probability of qualifying (or failing to qualify), as small as it may be, should be included in the sample as treatment or control. It is for this reason, in particular, that they characterize 36 "errors of exclusion": these teams had non-zero probabilities, in the sense that some combination of results could have landed them in or out of the tournament.

Note, however, that our original paper is explicit, if not precise, about the fact that zero probability is not the appropriate threshold for our identification strategy. In fact, we state (p. 1591): "The underlying identification assumption is that if two teams in the same group got to the last match day with *concrete* chances of qualifying, which one would actually qualify will be determined by quasi-random circumstances, such as a goal scored in the final minutes of the last match by one side or the other" (emphasis added).

We contend that DB's proposed approach is not the appropriate benchmark for determining whether a national team had concrete chances of qualifying, and hence "barely" succeeded or failed. Some examples can be helpful to illustrate the point:

**Examples G.3 and G.4: Eritrea and Mozambique, ACN 2000** As DB describe it, "Eritrea was the runner-up in the final table of this group, thus obtaining admission to a real playoff group. Hence, Eritrea could qualify before the last match day." Mozambique was in the same group and finished one point behind, such that with different results, "Mozambique would have ranked second, thus gaining [access] to the real playoff group." While technically true, consider the circumstances: Eritrea was second in the group at the end of the group stage, with a mere four points, because Ghana had been chosen as an emergency host country, midway through the qualifiers, with all games involving the Ghanaian team being at that point disregarded. Eritrea (which had lost to Ghana 5-0) was thus designated to a playoff group, where they proceeded to lose all four matches, with a goal differential of -11, and be promptly eliminated, nine points behind the last qualifying team. Having finished behind Eritrea, it follows that Mozambique was even further away from actual qualification.

**Example G.14: Mozambique, ACN 2004** It is true that Mozambique could have conceivably qualified, but consider the required results. One path involved beating Congo (the match was a 0-0 draw) with leader Burkina Faso losing to last-placed Central African Republic, while squandering a goal difference advantage that started at +12 above Mozambique (Burkina Faso won 3-0). Alternatively, if Mozambique had won, it could conceivably have qualified as the best runner-up across all groups with four teams. That, however, would have required Zimbabwe losing at home to last-placed Eritrea (they won 2-0) while squandering a +10 goal differential over Mozambique. Even then, it would have required negative results by Benin and Libya and Sierra Leone and Togo, in many cases also involving significant goal differential disadvantages. As it happened, Mozambique finished out of the virtual playoffs, where Congo was four points behind qualified Zimbabwe, on top of a 6-goal deficit in goal differential.

The key takeaway is that, broadly speaking, cases like these are ill-suited for identifying teams that were truly of an underlying quality that put them on the verge of qualification – that is to say, in a position from which success or failure are "as good as randomly" assigned. They happened to have a non-zero probability of qualifying (or failing to), because of extraneous circumstances, or simply because in a short round-robin tournament involving a small set of teams, even weak teams can maintain notional possibilities of qualification.

# 3.2. OUR CRITERION

In contrast, our approach aimed at a more constrained definition of "barely" qualified or eliminated, by using the information revealed by the full set of matches. Since our original paper did not describe our inclusion algorithm at length, we do here:

- 1. For each qualification slot defined in a head-to-head format, we assign the winner to the treatment group, and the loser to the control group.
- 2. For each qualification slot defined in a round-robin group format, we include:
  - In the treatment: for each group, any qualifying team that finished 3 points or less ahead of the team in the group that had the most points while still failing to

qualify, as long as the qualifying team in question entered the last matchday of the group with a nonzero probability of elimination.

- In the control: for each group, any non-qualifying team that finished 3 points or less behind the team in the group that had the fewest points while still qualifying, as long as the non-qualifying team in question entered the last matchday of the group with a nonzero probability of qualification.
- 3. For each qualification slot defined in a "virtual playoff" among top ranked not-directlyqualified teams in a round-robin group, we include:
  - In the treatment: any qualifying team in the "virtual playoff"
  - In the control: any team in the "virtual playoff" that finished 3 points or less behind the team in the "virtual playoff" that had the fewest points while still qualifying.

It is important to emphasize why we view this as a more appropriate criterion to implement our identification strategy. When it comes to addressing the identification challenge, the goal is not to pinpoint a specific moment where all residual uncertainty is eliminated; rather, it is to identify a set of countries within which we can plausibly argue that the qualification treatment is as good as randomly assigned. Adding cases where qualification or elimination was notionally possible, but highly unlikely, works against that.

In this context, no criterion can be perfect, in the sense of entirely free of type-I or type-II error. In fact, DB's criterion is certainly not indefensible: for instance, it has the appealing property of a uniform threshold as far as the level of probability at which teams are included or discarded – zero. But it is emphatically not the case that what they flag are "errors" in the data: they are merely the result of a stricterinclusion criteria.

# 3.2.1. REVISITING OUR CRITERION

While having the advantage of simplicity, our criterion did leave aside certain aspects of the qualification process that, in hindsight, seem worth incorporating. We are grateful to DB for prompting us to consider them, as well as for their careful documentation of potential cases, which facilitated this process.

The first key aspect is that, in the context of the round-robin group formats, there are teams that fail to access the "virtual playoffs" for indirect qualification, but finish relatively close to that slot, and would hypothetically have been close to qualification had they acceded. On top of that, the qualification process has occasionally been hit by shocks as it unfolded,

with teams withdrawing or being pulled out, and qualification criteria changing as a result.<sup>8</sup> Our original work was built on looking at final group and playoff tables, and it ended up disregarding some of the complexities brought about by these exceptional features.

In a few cases, taking these features into account makes an important conceptual difference. These encompass four out of the five cases that DB describe as "inclusion errors": these are countries that came into the last round of matches either qualified or eliminated from contention, in light of the playoff possibilities. These actually matter, because it is reasonable to assume that a team's performance in the last match is affected by having nothing to play for, which clearly distorts the idea behind of our identification strategy. In light of that, we agree with DB that what they label as cases F2, F3, F4, and F5, should indeed be excluded.<sup>9</sup> We disagree with case F1 (Uganda 2004), for the simple reason that Uganda did have a shot at qualifying when playing their last match; it only so happened that this match took place before the very last group matchday, because the group had an odd number of teams.

### 3.2.2. Defining Distance

More broadly, however, we agree that it is worth incorporating the full set of playoff possibilities and the occasional shocks into the analysis. For that, we will now describe a measure of distance from qualification in round-robin groups that incorporates the information from the full set of matches, as argued above, while also taking into account the qualification possibilities beyond the group itself.

We start by defining a broad sample that encompasses all the observations proposed by DB – that is to say, without the "inclusion errors" that we acknowledge above – plus Uganda, which that they claim as such an error, but we have argued should be included.<sup>10</sup> Then, for each of the country-qualification cases coming from a round-robin group format, we define a measure of distance from qualification (henceforth *Distance*) as follows:

<sup>&</sup>lt;sup>8</sup> This is illustrated by the case of ACN 2000, which gave rise to the Eritrea and Mozambique examples highlighted above. Zimbabwe was supposed to host the event, but in February 1999 it was announced that the right to host was being withdrawn because of non-compliance with event specifications. The African Football Confederation (CAF) opened a process for new bids, and in March 1999 it was announced that Ghana and Nigeria had won with a joint bid, beating Morocco. This led to the change in the qualification process that we explained above: Ghana and Nigeria qualified as hosts, and the matches they had played were annulled from the classification table. Senegal (from Nigeria's group) and Eritrea (from Ghana's) joined a playoff group with Zimbabwe, with one slot available.

<sup>&</sup>lt;sup>9</sup> Two of these cases were mistakenly included because of an overlooked change in tie-breaking criteria, from goal differential to performance in head-to-head matchups. This meant that South Africa in 2006, and Zambia in 2015, were already qualified even though they could still end up tied in points, because they were set to win the head-to-head comparison.

<sup>&</sup>lt;sup>10</sup> Rather unsurprisingly, this case does not affect our results. That this is the case is clear from the fact that the original paper already checked robustness with respect to excluding each individual country-qualification case.

- 1. For each qualified team, we compute the difference between the team's final number of points and the number of points of the first not-directly-qualified team in the same group, plus the difference between the latter and that of the first non-qualified team in the "automatic playoff".
- 2. For each non-qualified team involved in an "automatic playoff", we compute the difference between the team's final number of points and that of the last qualified team in the "playoff".
- 3. For each non-qualified team not involved in an "automatic playoff", we compute the difference between the team's final number of points and that of the team in the group that qualified for the "playoff", plus the difference between the latter and the last team qualified through the "playoff".

This measure summarizes how far the team was from the qualification threshold, given all the potential opportunities, with success coded as *Distance*  $\geq 0$  and failure as *Distance*  $\leq 0$ . (*Distance* = 0 cases get assigned to treatment or control depending on tiebreaking criteria such as goal differential.) The results are summarized in Table B.10.<sup>11,12</sup>

It is instructive to consider the distribution of *Distance* including all the instances framed as "exclusion errors" by DB, compared to the sample in our original paper. As we can see in Figure 1, DB's proposed sample is much less concentrated around zero than ours; capturing the fact that their criterion selects for cases that are not actually close to qualifying. This is

<sup>&</sup>lt;sup>11</sup> The case of ACN 2000, discussed above, poses complexities. Recall that Senegal, Zimbabwe, and Eritrea were put into a real playoff group after the group stage. Senegal and Zimbabwe finished tied at 9 points and Senegal qualified on goal difference, while Eritrea finished with zero points. We assign *Distance* = 0 to both Senegal and Zimbabwe, and -9 to Eritrea. Then there is Mozambique (from Ghana's group), where it seems clear that the relevant distance is -10 points, as they finished one point behind Eritrea in the regular group stage. The case of Burundi is more complicated, as it finished two points behind Senegal in Nigeria's group. We choose to set *Distance* = -11, under the premise that a team that did not get access to the playoff should be assessed as being behind the teams that did play in it (in this case, Eritrea). An alternative assumption would have Burundi at a distance of 2 points, which would add the distance from Senegal within the group to Senegal's performance in the playoff, as opposed to Eritrea's. Results are essentially unaltered by this alternative assumption (available upon request). A second complication comes from ACN 2012, where the virtual playoff required adjustments to the number of matches played, because different groups had different numbers of teams. This affects especially the case of Tanzania, as discussed in the notes to Table B.10.

<sup>&</sup>lt;sup>12</sup> There are fourteen cases that were not in our original sample, and that are now coded as having distance withing three points or less, under the revisited definition. Two of those come from the ACN 2000 playoff (Senegal and Zimbabwe, as per above); eight could have potentially qualified to a virtual playoff, but did not, yet would have been within three points if they had (Guinea and Sudan in 2004; Mali, Zambia, Tanzania, and Togo in 2008; Zambia in 2012; Ivory Coast in 2015); and three were teams that finished in last place in their groups (with no virtual playoff), with very negative goal differentials (Namibia and Kenya in 2002; Ghana in 2004). The one remaining case is Angola 2015, which finished five points behind the last directly qualified team in their group, but ended up three points behind the DRC in the virtual playoff.

true even leaving aside the extreme outliers to the left, which include some of the cases we have highlighted above.



FIGURE 1: DISTANCE TO QUALIFICATION IN POINTS (SAMPLE COMPARISON)

The figure represents the two histograms for points distance to qualification for our original sample (in light green) and DB(2024)'s proposed sample (in light red).

# 3.3. COMPARING RESULTS

We are now ready for a broad comparison of results, depending on different distance cutoffs that could be used to define what counts as "barely" qualified. Tables B.11- B.32 in the Appendix reproduce every single column of Tables 5-8 in our original paper. Their structure is as follows: Column 1 presents the results from our original paper; Column 2 presents the results from DB's proposed sample, and Columns 3-7 showcase results with the updated sample (DB's plus Uganda 2004), respectively restricting the sample to cases within 8, 4, 3, 2, and 1 points from qualification.

We will not bury the lede here: Every single result we find in the paper is reproduced with

*the new sample, within eight points of distance to qualification.* This is remarkable, and completely belies DB's assessment of a supposed lack of robustness of our results, which is wildly exaggerated even within their own terms. In fact, inspection of Tables B.11- B.32 reveals a clear pattern: broadly speaking, the smaller the distance threshold, the stronger the results. In other words, as we move away from close qualification, the coefficients typically become smaller, underscoring the importance of the idea behind our empirical strategy.<sup>13</sup>

Let us walk through each of DB's alleged "substantial number of results" that are presumably "modified." They claim (p.16) that "the effect of the treatment is not significant on the subsample of the 23 country-qualification campaign pairs that had never qualified for the ACN in Col. (4)" of Table 6 in our paper. Table B.21 shows that this is not the case as soon as we drop the most egregious cases (more than 8 points' distance). Even leaving aside the appropriateness of making as much of a drop in the estimated coefficient, which isn't even particularly large in their sample, the robustness of this result is very clear from our comparison.<sup>14</sup>

DB then call into question the effect on ethnic conflict. We should first note that the definition of "ethnic conflict" we use in the paper is rather strict. As we state on p.1579, "ACLED data do not explicitly distinguish between ethnic and non-ethnic conflict," and the information we use is, as we again note explicitly, "of course vulnerable to substantial measurement error, namely to the risk of coding as non-ethnic episodes that are in fact driven by ethnic motives." As such, we were very clear that this could lead to imprecise estimates. In spite of all these caveats, with their customarily ungenerous framing, DB make a lot of results becoming insignificant in that particular specification.

Yet, once again, that is entirely an artifact of their adding cases that lay far from the close qualification threshold. As Table B.23 makes clear, moving towards smaller points differences recovers larger coefficients – eventually larger than in the paper itself. Granted, the distinction between groups with weak and strong political power becomes smaller than in the original paper, though the coefficient on the former remains uniformly larger in magnitude under all specifications (see Tables B.24 and B.25).

DB's next point is about the results with different fatality thresholds. The pattern is exactly the same as in their other claims: everything hinges entirely on the fact that they incorporate

<sup>&</sup>lt;sup>13</sup> Needless to say, even if the coefficients became larger with distance, this would not invalidate our empirical strategy; it would rather suggest that the bias introduced by considering less comparable cases would have been positive on average. The pattern we find in the data suggests a negative bias introduced by those cases.

<sup>&</sup>lt;sup>14</sup> DB then go on to point out that "the interaction between the treatment and a first-time qualification is not significant in Col. (5)" of Table 6, implying that this is another result that is modified by their approach. But this is also true in our original paper! Even leaving aside the misleading impression left by their text, this should serve as a reminder that focusing on "significant" vs "not significant" is not the best approach.

the most egregious, not-at-all-close-qualification cases. As soon as we stick to even the 8-point threshold, all results are pretty much exactly as in the original paper, as can be seen in Tables B.30-B.32.

Having gone through all DB's claims of lack of robustness, let us turn to the "placebo" test. DB emphasize this, by selectively quoting from our paper: we state that "[i]f conflict was evolving differently in the two groups in the pre-qualification period, which would threaten the validity of our difference-in-differences approach, we would expect the fictitious treatment to display a significant coefficient," as they quote, but we also mention immediately afterwards that "[i]n particular, there is no indication that conflict was decreasing in countries that would eventually qualify, relative to countries that would not." This is evidently true even in their results, as can be seen from inspection of the event-study Figure J.1 in their comment.

Consider the event-study picture in Figure 2 below. It puts together our original version and DB's, as well as the ones that emerge from considering distances within three and one points. There is no evidence of pre-trends in any specification (including theirs), and it is clear that the results are even sharper as we focus on smaller distances.

In spite of that, DB make much of the fact that, in their specification, "the placebo tests in Cols. (3) and (4) of Table J.1 reveal a marginally significant difference in the evolution of conflict intensity in the treated and in the control group in the 12 weeks before the treatment date." Tables B.13 and B.14 make clear that even this goes away when focusing on the closest qualification cases.<sup>15,16</sup>

### 3.4. IN SUM

Any reasonable reading of the results laid out in the previous subsection can only underscore the remarkable robustness of our original analysis. In fact, they clarify that the sharpness of our findings increases as one moves closer to the set of truly close qualification cases. Even DB's claims of results being modified, which on closer inspection are much less stark than they seem to imply, largely rely on including egregious cases that do not fit the spirit or logic of our identification strategy.

<sup>&</sup>lt;sup>15</sup> Moreover, even for the other thresholds, the (marginal) significance is entirely due to the case of Zambia in 2008.

<sup>&</sup>lt;sup>16</sup> It is hard to know what to make of their claim that the results "do not control for any confounding effect that extends beyond the end of the final phase of the ACN." It is clear that, the farther into the future one goes, from the moment of the qualification treatment, the more confounding factors could appear. DB themselves note that we acknowledge these limitations.



### FIGURE 2: NUMBER OF CONFLICT EPISODES BEFORE AND AFTER QUALIFICATION

This figure represents Panel A of Figure 4 in Depetris-Chauvin et al. (2020) plotting coefficients and 95 percent confidence intervals for interactions between the dummy for countries that barely qualified to the ACN and 11 dummies for 4-week period included between 25 weeks before and after the qualification using the original sample in Depetris-Chauvin et al. (2020) and the sample proposed in Delpeyrou and Bertoli (2024) for different point threshold to qualification. The regressions also include week FE, calendar-month FE, and country × qualifier dummies.

I4R DP No. 194

### 4. CONCLUDING REMARKS

Replication work is a key feature of healthy science, and DB's effort illustrates why that is so. That said, it would be remiss of us not to point out that the way they present it also showcases the pitfalls of the currently existing incentives underpinning this type of work.

On the positive side, DB pushed us to be more precise where we had been insufficiently so in our original contribution. In doing so, they uncovered a few real issues with the selection of cases for our country-level analysis. More broadly, they helped us further establish the robustness of our results to different sample configurations. While we have argued that our criteria are superior to theirs, it is nevertheless very reassuring that our results essentially survive even under extreme approaches to the configuration of the sample. By the same token, when it comes to the individual-level analysis, they helped us spot a couple of legitimate mistakes in the construction of our data, which however proved to be immaterial for the results. We are sincerely grateful for their efforts on this regard.

On the other hand, however, DB fall prey to the unfortunate set of incentives that currently exist for replication – or more broadly and aptly, in this case, work that revisits previous contributions. The inescapable fact is that the impact of such work is often predicated on the extent to which it is deemed to "debunk" or find "major problems" with the original piece. As a result, it is disappointing but understandable that DB, at every single juncture, choose to present their points in a maximalist way. Reading their abstract, one would be forgiven the impression that our paper was severely discredited by their findings.

Yet, closer examination shows that this is emphatically not the case. The "major problems" claimed by DB are either exhaustively acknowledged and addressed in our paper – as in the case of the relatively small number of identifying observations in the individual-level analysis – or for the most part, the result not of "errors" but of different (and, as we argue, more compelling) choices in terms of specification and data construction. Moreover, the differences in results are actually remarkably small even under DB's choices, and largely pertain to ancillary results and not to the core of the analysis. Put simply, a rather more reasonable reading of their own findings actually underscores the robustness of our key conclusions.

The crucial problem, for the profession, is that there is very little payoff for putting in the kind of effort that they clearly did, and in the end come to the conclusion that the original paper is not "debunked." It behooves all of us to try and find a better equilibrium, where work like DB's is valued regardless of whether the conclusions it reaches "debunk" the original paper or not. Replicating and revisiting existing findings is a key propeller for scientific progress – again, as their comment indeed illustrates. Hopefully, we will move in a direction of healthier, more positive incentives towards that.

# REFERENCES

- **Delpeyrou, Léonie and Simone Bertoli**, "Building Nations through Shared Experiences: Evidence from African Football': a comment on Depetris-Chauvin, Durante and Campante (2020)," 2024. Unpublished.
- **Depetris-Chauvin, Emilio, Ruben Durante, and Filipe Campante**, "Building Nations through Shared Experiences: Evidence from African Football," *American Economic Review*, 2020, *110*, 1572–1602.
- Rambachan, Ashesh and Jonathan Roth, "Design-Based Uncertainty for Quasi-Experiments," 2022. arXiv:2008.00602.
- Roth, Jonathan, Pedro H. Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 2023, 235, 2218–2244.

# **APPENDIX B: ADDITIONAL TABLES AND FIGURES**

### TABLE B.1: NATIONAL TEAM'S PERFORMANCE AND ETHNIC IDENTIFICATION

Dependent Variable: Ethnic over National Identity (0-1 dummy)										
		Pan	el A·Publ	lished Ver	sion					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)			
Post-Match	-0.026	-0.029	-0.036	-0.001						
	(0.012)	(0.014)	(0.014)	(0.016)						
Post-Victory				-0.052	-0.053	-0.053	-0.243			
				(0.019)	(0.017)	(0.016)	(0.062)			
Post-Draw					-0.026					
					(0.039)					
Post-Defeat					-0.000					
Post Victory Marginal Effect					(0.017)		0.037			
Post-victory Marginar Effect							(0.014)			
Observations	37,060	37,060	37,060	37,060	37,060	37,060	35,305			
R-squared	0.094	0.102	0.104	0.104	0.104	0.104	_			
	Panel B: Correct Version									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)			
Post-Match	-0.026	-0.029	-0.036	-0.017						
	[0.012]	[0.014]	[0.014]	[0.020]						
Post-Victory				-0.030	-0.046	-0.042	-0.186			
Post-Draw				[0.026]	[0.018] -0.026	[0.018]	[0.0/1]			
1 Ost Diaw					[0.039]					
Post-Defeat					-0.016					
					[0.021]					
Post-Victory Marginal Effect							-0.048			
	<b>0-</b> 011	<b>2-</b> 0.1.1			<b>2-</b> 0.1.1	<b>0-</b> 011	(0.012)			
Observations <b>P</b> squared	37,011	37,011	37,011	37,011	37,011	37,011	35,247			
Country × Match FE	0.093	0.102 Vac	0.103 Vac	0.104 Vac	0.104 Vac	0.104 Vac	Vac			
	105	105	105	105	105	105	105			
Language $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Individual Controls	No	Yes	Yes	Yes	Yes	Yes	Yes			
Seasonal FE	No	No	Yes	Yes	Yes	Yes	Yes			

Robust standard errors clustered by country×year in parenthesis. Sample includes respondents interviewed within 15 days before and after an official game. *PostGame*, *PostVictory*, *Post – Draw* and *Post – Defeat* take value 1 if the respondent was interviewed in the 15 days after a game, a victory, a draw or a loss respectively, and 0 otherwise.

_	Dependent Variable: Ethnic over National Identity (0-1 dummy)									
			Panel A	: Published	Version					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Post-Victory	0.019	-0.046	-0.054	-0.054	-0.053	-0.052	-0.058	-0.074		
	(0.015)	(0.016)	(0.016)	(0.015)	(0.017)	(0.015)	(0.019)	(0.016)		
Interaction		-0.061 (0.020)	0.014 (0.031)	-0.012 (0.031)	0.001 (0.010)	0.049 (0.027)	-0.034 (0.157)	-0.280 (0.137)		
Uninteracted Term	_	_	_	_	_	-0.023 (0.012)	_	_		
Observations	40,392	37,060	37,060	37,060	37,060	37,060	34,450	26,186		
R-squared	0.087	0.104	0.104	0.104	0.104	0.104	0.107	0.113		
Panel B: Correct Version										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Post-Victory	0.019	-0.033	-0.043	-0.043	-0.039	-0.040	-0.034	-0.064		
	(0.015)	(0.018)	(0.018)	(0.017)	(0.020)	(0.018)	(0.020)	(0.018)		
Interaction		-0.073	0.023	-0.022	-0.003	0.043	-0.239	-0.363		
		(0.020)	(0.032)	(0.031)	(0.010)	(0.027)	(0.145)	(0.236)		
Uninteracted Term	-	-	-	-	—	-0.022	—	-		
						(0.011)				
Observations	40,392	37,011	37,011	37,011	37,011	37,011	34,401	26,137		
R-squared	0.087	0.104	0.104	0.104	0.104	0.103	0.107	0.113		
Interaction Term	None	Rivalry	Home	Wide	Goals	State	Diversity	Diversity		
			Game	Margin	in Game	Presence	Country	Team		
Sample:	Friendly	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline		
	Matches									
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Seasonal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Language×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Country×Match FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		

### TABLE B.2: NATIONAL TEAM'S VICTORIES AND ETHNIC IDENTITY: STAKES AND HETEROGENEOUS EFFECTS

Robust standard errors clustered at the country×year level in parentheses. Post-Victory takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. To ease the comparison with previous tables, variables in the interaction terms were demeaned. State presence is computed as the mean value of three indicators coded by Afrobarometer's interviewer at the enumeration area: presence of schools, post offices, and paved roads. National diversity is based on the ELF index from Fearon and Laitin (2003). Team diversity is computed as a ELF index based on the ethnic composition of the national team in the same year of the Afrobarometer's wave.

	Trust in Countrymen	Inter-Ethnic Trust	Like Neighbors Other Ethnicities	Dislike Foreign Neighbors					
		Panel A:	Published Version						
	(1)	(2)	(3)	(4)					
Post-Victory	0.063	0.140	0.102	0.019					
	(0.021)	(0.040)	(0.030)	(0.018)					
Observations	9,355	7,973	7,511	7,497					
R-squared	0.140	0.169	0.162	0.153					
		Panel B: Correct Version							
	(1)	(2)	(3)	(4)					
Post-Victory	0.072	0.140	0.102	0.019					
	(0.021)	(0.040)	(0.030)	(0.018)					
Observations	9,355	7,973	7,511	7,497					
R-squared	0.140	0.169	0.162	0.153					
Individual Controls	Yes	Yes	Yes	Yes					
Seasonal FE	Yes	Yes	Yes	Yes					
Language×Year FE	Yes	Yes	Yes	Yes					
Country×Match FE	Yes	Yes	Yes	Yes					

### TABLE B.3: NATIONAL TEAM'S VICTORIES AND TRUST IN OTHERS

Robust standard errors clustered at the country  $\times$  year level in parentheses. Post-Victory takes value 1 if the respondent was interviewed in the 15 days after a victory, 0 otherwise. Trust in Countrymen takes value 1 if respondent trusts other countrymen "somewhat" or "a lot", 0 otherwise. Inter-Ethnic Trust takes value 1 if respondent trusts "somewhat" or "a lot" people of other ethnicities, 0 otherwise. Like Neighbors Other Ethnicities takes value 1 if respondent would "like" or "strongly like" having neighbors from other ethnicities, 0 otherwise. Dislike Foreign Neighbors takes value 1 if respondent would "dislike" or "strongly dislike" having immigrants or foreign workers as neighbors, 0 otherwise.

Covariate

Male Education Age Unemployed Major ethnicity

Rural

Public goods

Same language

Influenced by others

Education Interviewer

Male Interviewer

Age Interviewer

Fixed Effect

Religious group member

0.62

0.41

0.49

0.44

0.04

0.57

7.18

28.70

37.085

36.957

37.085

37,085

37,038

37,085

36,431

37.085

0.168

-0.017

-0.023

-0.022

0.000

-0.003

-0.069

0.189

**Country x Match** 

[0.081]

[0.026]

[0.017]

[0.045]

[0.008]

[0.018]

[0.061]

[0.134]

37.011

36,883

37.011

37,011

36,964

37,011

36,357

37,011

			TAB	le <b>B</b> .4:	BALANC	e Test						
	Panel A: Post-Victory (Original) Panel B: Post-Victory		ctory	Panel C	C: Share of V	ictories	Panel D: Share of Points Won					
Mean Value	Ν	Estimate	Std Errors	Ν	Estimate	Std Errors	Ν	Estimate	Std Errors	Ν	Estimate	Std Errors
0.50	37,085	0.01	[0.005]	37,011	0.007	[0.009]	45,459	0.003	[0.008]	45,459	0.004	[0.008]
3.08	37,085	-0.254	[0.154]	37,011	-0.246	[0.135]	45,459	-0.22	[0.140]	45,459	-0.271	[0.133]
36.82	37,085	1.175	[0.752]	37,011	0.468	[0.524]	45,459	0.314	[0.483]	45,459	0.422	[0.450]
0.29	37,085	-0.008	[0.013]	37,011	0.013	[0.014]	45,459	0.009	[0.013]	45,459	0.005	[0.014]
0.45	37,085	-0.027	[0.040]	37,011	0.04	[0.037]	45,459	0.055	[0.049]	45,459	0.057	[0.052]

[0.056]

[0.022]

[0.034]

[0.026]

[0.008]

[0.021]

[0.071]

[0.332]

0.084

-0.044

0.011

0.002

0.004

0.024

0.098

-0.067

Language x Year

45.459

45,315

45.459

45,459

45,405

45,459

44,592

45,459

[0.058]

[0.024]

[0.035]

[0.029]

[0.007]

[0.025]

[0.080]

[0.388]

45.459

45,315

45,459

45,459

45,405

45,459

44,592

45,459

0.091

-0.041

-0.003

0.002

0

0.02

0.102

-0.165

[0.047]

[0.025]

[0.029]

[0.028]

[0.007]

[0.026]

[0.081]

[0.395]

Each panel presents estimates and standard errors for 13 regressions of a covariate (listed at the left) on different versions of our treatment variable. Post-victory takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. Share of Victories accounts for the fraction of total games won. Share of points Won accounts for the fraction of total possible points obtained (a win, draw, and lose awards 3, 1, and 0 points, respectively) within a 15 days window. All estimates are based on OLS regressions using a set of country-match dummies in Panel A and language-year dummies in the other panels to ensure that the comparison in the covariates is made between respondents in the proximity of the same game and in the same country as well as within the same ethnic group in the same year, respectively.

0.1

-0.037

0.021

0.002

0.005

0.02

0.049

-0.052

Dependent Variable: Ethnic over National Identity (0-1 dummy)									
	(1)	(2)	(3)	(4)	(5)				
Post-Match	-0.015	0.015							
	[0.013]	[0.015]							
Post-Victory		-0.053	-0.037	-0.041	-0.201				
		[0.018]	[0.017]	[0.016]	[0.072]				
Post-Draw			0.026						
			[0.026]						
Post-Defeat			0.010						
			[0.019]						
Post-Victory Marginal Effect					-0.040				
(Probit Model in Column 5)					[0.015]				
Observations	37,011	37,011	37,011	37,011	35,252				
R-squared	0.099	0.099	0.099	0.099	_				
Language $\times$ Year FE	Yes	Yes	Yes	Yes	Yes				
Individual Controls	Yes	Yes	Yes	Yes	Yes				

# TABLE B.5: MAIN RESULTS ONLY EXPLOITING WITHIN LANGUAGE X YEAR VARIATION

Robust standard errors clustered by country×year in parenthesis. These estimates only include Language × Year FE (i.e., no other set of fixed effects is included). Sample includes respondents interviewed within 15 days before and after an official game. Post-victory takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. Trust in Countrymen takes value 1 if respondent trusts other countrymen "somewhat" or "a lot", 0 otherwise. Inter-Ethnic Trust takes value 1 if respondent trusts "somewhat" or "a lot" people of other ethnicities, 0 other- wise. Like Neighbors Other Ethnicities takes value 1 if respondent would "like" or "strongly like" having neighbors from other ethnicities, 0 otherwise. Dislike Foreign Neighbors takes value 1 if respondent would "dislike" or "strongly dislike" having immigrants or foreign workers as neighbors, 0 otherwise.

	Dependent Variable: Ethnic over National Identity (0-1 dummy)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-Victory	-0.039 (0.013)	-0.039 (0.017)	-0.041 (0.015)	-0.041 (0.015)	-0.040 (0.015)	-0.039 (0.015)	-0.059 (0.013)		
Interaction	-0.075 (0.014)	-0.016 (0.024)	-0.006 (0.030)	-0.00 (0.011)	0.034 (0.028)	-0.201 (0.138)	-0.308 (0.106)		
Uninteracted Term	-	-	_	-	-0.024 (0.012)	-	-		
Observations	37,011	37,011	37,011	37,011	37,011	34,401	26,137		
R-squared	0.100	0.099	0.099	0.099	0.099	0.103	0.109		
Interaction Term	Rivalry	Home Game	Wide Margin	Goals in Game	State Presence	Diversity Country	Diversity Team		
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Language×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		

### TABLE B.6: HETEROGENEITY ONLY EXPLOITING WITHIN LANGUAGE X YEAR VARIATION

Robust standard errors clustered at the country×year level in parentheses. These estimates only include Language × Year FE (i.e., no other set of fixed effects is included). Post-Victory takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. To ease the comparison with previous tables, variables in the interaction terms were demeaned. State presence is computed as the mean value of three indicators coded by Afrobarometer's interviewer at the enumeration area: presence of schools, post offices, and paved roads. National diversity is based on the ELF index from Fearon and Laitin (2003). Team diversity is computed as a ELF index based on the ethnic composition of the national team in the same year of the Afrobarometer's wave.

#### TABLE B.7: TRUST RESULTS ONLY EXPLOITING WITHIN LANGUAGE X YEAR VARIATION

Dependent Variable:	Trust in Countrymen (1)	Inter-Ethnic Trust (2)	Like Neighbors Other Ethnicities (3)	Dislike Foreign Neighbors (4)
Post-Victory	0.031 [0.022]	0.159 [0.009]	-0.014 [0.055]	0.032 [0.011]
Observations	9,355	7,962	7,511	7,497
R-squared	0.134	0.158	0.156	0.148
Language $\times$ Year FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes

Robust standard errors clustered by country×year in parenthesis. These estimates only include Language × Year FE (i.e., no other set of fixed effects is included). Sample includes respondents interviewed within 15 days before and after an official game. Post-victory takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. Trust in Countrymen takes value 1 if respondent trusts other countrymen "somewhat" or "a lot", 0 otherwise. Inter-Ethnic Trust takes value 1 if respondent trusts "somewhat" or "a lot" people of other ethnicities, 0 other- wise. Like Neighbors Other Ethnicities takes value 1 if respondent would "like" or "strongly like" having neighbors from other ethnicities, 0 otherwise. Dislike Foreign Neighbors takes value 1 if respondent would "dislike" or "strongly dislike" having immigrants or foreign workers as neighbors, 0 otherwise.

	Ethnic	Trust	Inter-Ethnic	Like neighbors	Dislike Foreign	Trust in	President's
	Identification	Countrymen	Trust	Other Ethnicities	neighbors	Ruling Party	Approval
			Par	el A: Published Ve	rsion		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of Victories	-0.046	0.068	0.144	0.033	0.015	-0.003	0.026
	(0.013)	(0.019)	(0.039)	(0.057)	(0.009)	(0.022)	(0.026)
Observations	45,500	12,342	8,171	10,735	10,710	48,769	48,481
R-squared	0.098	0.154	0.164	0.228	0.151	0.152	0.204
			Pa	anel B: Correct Vers	sion		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of Victories	-0.039	0.076	0.144	0.033	0.015	0.008	0.027
	(0.014)	(0.023)	(0.039)	(0.057)	(0.009)	(0.022)	(0.027)
Observations	45,459	12,342	8,171	10,735	10,710	48,719	48,432
R-squared	0.098	0.154	0.164	0.228	0.151	0.152	0.204
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Seasonal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Pane	l C: Only Expl	oiting Within Lang	uage x Year Variat	ion	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of Victories	-0.032	0.030	0.158	0.018	0.019	-0.000	0.020
	(0.014)	(0.021)	(0.012)	(0.038)	(0.021)	(0.014)	(0.017)
Observations	45,459	12,342	8,171	10,735	10,710	48,719	48,432
R-squared	0.094	0.150	0.155	0.224	0.148	0.151	0.203
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

# TABLE B.8: MULTIPLE GAMES (SHARE OF VICTORIES): PANEL A IN TABLE A.6

Robust standard errors clustered at the country  $\times$  year level in parentheses. Share of Victories accounts for the fraction of total matches won.

	Panel A: Published Version							
	Ethnic	Trust	Inter-Ethnic	Like neighbors	Dislike Foreign	Trust in	President's	
	Identification	Countrymen	Trust	Other Ethnicities	neighbors	Ruling Party	Approval	
			Par	el A: Published Ve	rsion			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Share of Points Won	-0.047	0.049	0.144	0.019	0.003	0.001	0.024	
	(0.013)	(0.031)	(0.039)	(0.063)	(0.016)	(0.023)	(0.026)	
Observations	45,500	12,342	8,171	10,735	10,710	48,769	48,481	
R-squared	0.098	0.154	0.164	0.228	0.151	0.152	0.204	
			Pa	anel B: Correct Ver	sion			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Share of Points Won	-0.040	0.059	0.144	0.019	0.003	0.012	0.025	
	(0.014)	(0.031)	(0.039)	(0.063)	(0.016)	(0.023)	(0.027)	
Observations	45,459	12,342	8,171	10,735	10,710 48,719		48,432	
R-squared	0.098	0.154	0.164	0.228	0.151	0.152	0.204	
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Seasonal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Language $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Country $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
		Pane	l C: Only Expl	oiting Within Lang	uage x Year Variat	ion		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Share of Points Won	-0.030	0.025	0.158	-0.066	0.018	-0.001	0.026	
	(0.014)	(0.022)	(0.012)	(0.050)	(0.013)	(0.016)	(0.020)	
Observations	45,459	12,342	8,171	10,735	10,710	48,719	48,432	
R-squared	0.094	0.149	0.155	0.222	0.148	0.147	0.197	
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Language × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

# TABLE B.9: MULTIPLE GAMES (SHARE OF POINTS WON): PANEL B IN TABLE A.6

Robust standard errors clustered at the country  $\times$  year level in parentheses. Share of points Won accounts for the fraction of total possible points obtained (a win, draw, and lose awards 3, 1, and 0 points, respectively).

### I4R DP No. 194

# TABLE B.10: DISTANCE IN POINTS TO QUALIFICATION

Institute for Replication

Points Distance to Qualification	Year of Tournament	Treatment Group	Control Group
0	2000	Togo, Senegal	Guinea, Liberia, Zimbabwe
	2002		Gabon
	2008	Benin, South Africa, Zambia	Uganda
	2012	Mali, Niger, Sudan	
	2013	Ivory Coast, Ethiopia, Niger, DRC, Burkina Faso, Angola	Malawi, Botswana, Uganda, Sierra Leone, Senegal, Liberia, Cameroon
		Togo, Ghana, Mali, Nigeria, Zambia	Equatorial Guinea, Sudan, Guinea, Mozambique, Zimbabwe, Gabon, CAR
	2015	DRC	
1	1998	Namibia, DRC, Mozambique	Mali, Senegal, Gabon, Liberia, Malawi
	2000	Congo	Mali, Uganda
	2004	DRC	
	2008	Namibia	Equatorial Guinea
	2010	Malawi	Guinea
	2012	Angola, Zambia	
	2015		Nigeria
2	1998	Angola	Zimbabwe
	2000	Ivory Coast	
	2002	DRC	Zimbabwe, Lesotho, Kenya
	2004	Benin, Rwanda, Mali, Zimbabwe	Uganda, Zambia
	2008	Senegal, Mali	Eritrea, Mozambique, DRC, Togo
	2012		Sierra Leone, Nigeria, Malawi, South Africa, Uganda, CAR
	2015	Congo, Mali, Ivory Coast	Malawi, Uganda
3	2002	Zambia, Burkina Faso	Angola, Madagascar, Namibia
	2004	Kenya, South Africa, Guinea	Togo, Sierra Leone, Ivory Coast, Madagascar, Sudan, Ghana
	2006	DRC	Burkina Faso
	2008	Ivory Coast, Guinea	Gabon, Gambia, Tanzania
	2010	Zambia	Rwanda
	2012	Guinea	Zimbabwe, Kenya, Gambia
	2015	Guinea	Mozambique, Togo, Angola
4	1998	Ivory Coast	Tanzania
	2004	Burkina Faso	Ethiopia, Congo
	2008		Botswana, Burundi, Congo
	2010	Mozambique	Kenya
	2012	Ghana	Tanzania (a), Burundi
	2015	Ghana	
5	2004		Mozambique, Eswatini
	2008		Ethiopia
	2010	Togo	
	2015		Ethiopia
9	2000		Eritrea
10	2000		Mozambique (b)
11	2000		Burundi (c)

(a) Tanzania finished in 4th place, 3 points below the qualifying team to the virtual playoff in 2012. In the 2012 edition the virtual playoff required points adjustments to account for the different number of matches played in the various groups. To calculate the points earned in a hypothetical virtual playoff scenario, an assumption needs to be made regarding the team in the fourth position (either Algeria or CAR) and to consider the points gained against that hypothetical team (those points must be subtracted). The optimal scenario for Tanzania would be if Algeria ended in 4th place, resulting in only 2 points subtracted to Tanzania. In the event that CAR occupied the 4th position, 3 points would have been deducted from Tanzania. We therefore assigned a distance of just 1 point within the playoff (b) Mozambique ended one point below Eritrea, the worst team playing a real playoff group (ending 9 points below the qualifying team). We therefore assigned a distance of 9 points within the playoff (the idea being that a team not qualifying to the playoff group, should be farther to qualification than the worst team in the playoff group). (c) Burundi ended two points below Senegal and not qualifying to the playoff group. We therefore assigned a distance of 9 points within the playoff (the idea being that a team not qualifying to the playoff group) should be farther to qualification than the worst team in the playoff group) should be farther to qualification than the worst team in the playoff group) (c) Burundi group should be farther to qualification than the worst team in the playoff group).

	Dependent Variable: Log (1+ Number of Events)							
Sample	AER (2020)	DB (2024)	DB (2024) + Uganda (2004)					
-	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Post-qualification	-0.131	-0.132	-0.135	-0.143	-0.143	-0.201	-0.170	
	(0.064)	(0.059)	(0.057)	(0.057)	(0.061)	(0.069)	(0.072)	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	5,450	7,000	6,900	6,650	5,950	4,400	2,950	
Within R-sq	0.004	0.004	0.004	0.005	0.004	0.008	0.007	

### TABLE B.11: ACN QUALIFICATION AND CONFLICT: COLUMN 1 OF TABLE 5 MAIN EFFECT

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)								
Sample	AER (2020)	DB (2024)		DB (2024) + Uganda (2004)					
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-qualification	-0.090	-0.087	-0.089	-0.093	-0.094	-0.132	-0.109		
	(0.039)	(0.035)	(0.034)	(0.035)	(0.036)	(0.041)	(0.043)		
Long-Run Impact	-0.145	-0.148	-0.147	-0.153	-0.155	-0.214	-0.178		
	(0.065)	(0.061)	(0.058)	(0.059)	(0.062)	(0.071)	(0.072)		
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950		
Within R-sq	0.080	0.087	0.082	0.082	0.085	0.087	0.088		

### TABLE B.12: ACN QUALIFICATION AND CONFLICT: COLUMN 2 OF TABLE 5 MAIN EFFECT WITH LAGS

		Dependent	Variable: L	.og (1+ Nu	mber of Ev	vents)		
Sample	AER (2020)	DB (2024)		DB (2024) + Uganda (2004)				
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
12 Weeks Before Qualification	0.046	0.109	0.102	0.108	0.110	0.104	0.056	
-	(0.062)	(0.058)	(0.058)	(0.058)	(0.061)	(0.070)	(0.083)	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	2,725	3,500	3,450	3,325	2,975	2,200	1,475	
Within R-sq	0.000	0.003	0.002	0.003	0.003	0.002	0.001	

# TABLE B.13: ACN QUALIFICATION AND CONFLICT: COLUMN 3 OF TABLE 5 PLACEBO

Robust standard errors in parentheses clustered at the country× qualification campaign level. Samples cover 25 weeks before the end of qualification process (i.e. pre-treatment period) for 109 country × qualification campaign. The variable 12 Weeks Before Qualification takes value 1 during the 12 weeks immediately before the end of the qualification process for the countries that will eventually qualify to the ACN, 0 otherwise. Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)						
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugano	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
12 Weeks Before Qualification	0.028	0.076	0.072	0.078	0.080	0.075	0.041
	(0.054)	(0.043)	(0.043)	(0.044)	(0.046)	(0.053)	(0.062)
Long-Run Impact	0.033	0.11	0.10	0.11	0.11	0.10	0.058
	(0.067)	(0.060)	(0.061)	(0.061)	(0.064)	(0.072)	(0.087)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	109	140	138	133	119	88	59
Observations	2,725	3,500	3,450	3,325	2,975	2,200	1,475
Within R-sq	0.025	0.041	0.041	0.041	0.042	0.041	0.046

# TABLE B.14: ACN QUALIFICATION AND CONFLICT: COLUMN 4 OF TABLE 5 PLACEBO WITH LAGS

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. Samples cover 25 weeks before the end of qualification process (i.e. pre-treatment period) for 109 country  $\times$  qualification campaign. The variable 12 Weeks Before Qualification takes value 1 during the 12 weeks immediately before the end of the qualification process for the countries that will eventually qualify to the ACN, 0 otherwise. Conflict data comes from the ACLED dataset.

Sampla	Dependent Variable: Log (1+ Number of Events) A ER (2020) DB (2024) $DB$ (2024) + Uganda (2004)								
Sample	AER (2020)	DB (2024)		DB (202	(4) + Ogano	Ja (2004)			
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
1-12 Weeks Post-Qualification (a)	-0.124	-0.126	-0.127	-0.132	-0.129	-0.181	-0.173		
	(0.064)	(0.057)	(0.055)	(0.056)	(0.060)	(0.068)	(0.070)		
13-25 Weeks Post-Qualification (b)	-0.137	-0.138	-0.141	-0.153	-0.157	-0.219	-0.167		
	(0.081)	(0.073)	(0.071)	(0.072)	(0.076)	(0.089)	(0.095)		
Prob > F H0 : a = b	0.86	0.83	0.81	0.74	0.66	0.63	0.95		
Country $ imes$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	5,450	7,000	6,900	6,650	5,950	4,400	2,950		
Within R-sq	0.004	0.004	0.004	0.005	0.005	0.009	0.007		

# TABLE B.15: ACN QUALIFICATION AND CONFLICT: COLUMN 5 OF TABLE 5Persistence

Robust standard errors in parentheses clustered at the country× qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable 13-25 Weeks Post-Qualification takes value 1 starting the 13th week after the end of the qualification process for the countries that barely qualify to the ACN, 0 otherwise. Prob > F H0 : a = b refers to the F-tests with the Null Hypothesis 1-12 Weeks Post-Qualification = 13-25 Weeks Post-Qualification. Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)							
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugano	la (2004)		
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
1-12 Weeks Post-Qualification (a)	-0.098	-0.094	-0.094	-0.097	-0.095	-0.129	-0.124	
	(0.038)	(0.034)	(0.033)	(0.034)	(0.036)	(0.040)	(0.040)	
13-25 Weeks Post-Qualification (b)	-0.083	-0.080	-0.084	-0.090	-0.092	-0.134	-0.095	
	(0.053)	(0.044)	(0.045)	(0.045)	(0.047)	(0.056)	(0.062)	
Prob > F H0 : a = b	0.75	0.71	0.80	0.86	0.95	0.93	0.64	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950	
Within R-sq	0.080	0.087	0.082	0.082	0.085	0.087	0.088	

# TABLE B.16: ACN QUALIFICATION AND CONFLICT: COLUMN 6 OF TABLE 5 PERSISTENCE WITH LAGS

Robust standard errors in parentheses clustered at the country× qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable 13-25 Weeks Post-Qualification takes value 1 starting the 13th week after the end of the qualification process for the countries that barely qualify to the ACN, 0 otherwise. Prob > F H0 : a = b refers to the F-tests with the Null Hypothesis 1-12 Weeks Post-Qualification = 13-25 Weeks Post-Qualification. Conflict data comes from the ACLED dataset.

		Depen	dent Variab	ole: Numbe	r of Events	6	
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-Qualification	-0.307	-0.332	-0.345	-0.373	-0.347	-0.451	-0.396
-	(0.156)	(0.156)	(0.151)	(0.152)	(0.156)	(0.166)	(0.177)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# country $\times$ qual. campaigns	109	140	137	132	118	87	58
Observations	5,450	7,000	6,850	6,600	5,900	4,350	2,900

# TABLE B.17: ACN QUALIFICATION AND CONFLICT: COLUMN 7 OF TABLE 5NEGATIVE BINOMIAL

Robust standard errors in parentheses clustered at the country × qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

TABLE B.18: ACN QUALIFICATION AND CONFLICT: COLUMN 8 OF TABLE 5
NEGATIVE BINOMIAL WITH LAGS

	Dependent Variable: Number of Events								
Sample	AER (2020)	DB (2024)		DB (2024) + Uganda (2004)					
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-Qualification	-0.227	-0.234	-0.245	-0.271	-0.244	-0.323	-0.281		
	(0.127)	(0.127)	(0.122)	(0.123)	(0.125)	(0.134)	(0.143)		
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950		

		Dependent	Variable: L	.og (1+ Nu	mber of Ev	vents)	
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.153	-0.154	-0.163	-0.169	-0.157	-0.216	-0.149
-	(0.070)	(0.057)	(0.057)	(0.057)	(0.059)	(0.062)	(0.053)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	53	74	72	68	58	41	24
Observations	2,438	3,700	3,600	3,400	2,900	2,050	1,200
Within R-sq	0.097	0.087	0.079	0.080	0.084	0.092	0.067

# Institute for Replication ACN QUALIFICATION AND CONFLICT: COLUMN 2 OF TABLE & DP No. 194 OVERDUE QUALIFICATION

Robust standard errors in parentheses clustered at the country $\times$  qualification campaign level. An overdue (firsttime) qualification is defined as reaching the last match-day with chances of qualifying to the ACN finals after 3 or more years (for the very first time). See Appendix Table A.20 in original paper. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)									
Sample	AER (2020)	DB (2024)	DB (2024) + Uganda (2004)							
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)			
Post-qualification	-0.087	-0.087	-0.088	-0.097	-0.095	-0.138	-0.109			
	(0.044)	(0.036)	(0.036)	(0.036)	(0.038)	(0.041)	(0.036)			
Post-qualification x Overdue	-0.148	-0.147	-0.156	-0.150	-0.129	-0.154	-0.048			
	(0.088)	(0.072)	(0.073)	(0.074)	(0.076)	(0.085)	(0.073)			
Prob > F	0.11	0.03	0.02	0.02	0.03	0.00	0.00			
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01			
# country $\times$ qual. campaigns	109	140	138	133	119	88	59			
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950			
Within R-sq	0.081	0.088	0.084	0.083	0.086	0.088	0.089			

# TABLE B.20: ACN QUALIFICATION AND CONFLICT: COLUMN 3 OF TABLE 6OVERDUE QUALIFICATION WITH INTERACTION TERM

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise. An overdue (first-time) qualification is defined as reaching the last match-day with chances of qualifying to the ACN finals after 3 or more years (for the very first time). See Appendix Table A.20 in original paper. Prob > F refers to the F-tests with the Null Hypothesis that coefficients for post-qualification and its interaction with overdue are jointly equal to zero. Interaction terms were demeaned to ease the comparison of uninteracted terms. Conflict data comes from the ACLED dataset.

	vents)						
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.273	-0.220	-0.257	-0.258	-0.255	-0.299	-0.120
	(0.137)	(0.147)	(0.147)	(0.145)	(0.144)	(0.147)	(0.118)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	16	23	25	20	17	13	5
Observations	736	1,150	1,050	1,000	850	650	250
Within R-sq	0.124	0.129	0.093	0.097	0.110	0.118	0.066

# Institute for Replication ACN QUALIFICATION AND CONFLICT: COLUMN 4 OF TABLE & DP No. 194 FIRST-TIME QUALIFICATION

Robust standard errors in parentheses clustered at the country $\times$  qualification campaign level. An overdue (first-time) qualification is defined as reaching the last match-day with chances of qualifying to the ACN finals after 3 or more years (for the very first time). See Appendix Table A.20 in original paper. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise. Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)								
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	la (2004)			
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-qualification	-0.099	-0.097	-0.100	-0.104	-0.103	-0.136	-0.108		
	(0.043)	(0.039)	(0.039)	(0.040)	(0.041)	(0.043)	(0.041)		
Post-qualification x First-Time	-0.226	-0.165	-0.171	-0.171	-0.184	-0.202	0.004		
-	(0.176)	(0.170)	(0.168)	(0.169)	(0.169)	(0.178)	(0.111)		
Prob > F	0.07	0.05	0.04	0.03	0.05	0.01	0.04		
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950		
Within R-sq	0.081	0.088	0.083	0.083	0.086	0.088	0.088		

# TABLE B.22: ACN QUALIFICATION AND CONFLICT: COLUMN 5 OF TABLE 6FIRST-TIME QUALIFICATION WITH INTERACTION TERM

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. An overdue (first-time) qualification is defined as reaching the last match-day with chances of qualifying to the ACN finals after 3 or more years (for the very first time). See Appendix Table A.20 in original paper. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Prob > F refers to the F-tests with the Null Hypothesis that coefficients for post-qualification are jointly equal to zero. Interaction terms were demeaned to ease the comparison of uninteracted terms. Conflict data comes from the ACLED dataset.

		Dependent Variable: Log (1+ Number of Events)							
Sample	AER (2020)	AER (2020) DB (2024)			DB(2024) + Uganda(2004)				
-	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-qualification	-0.024	-0.018	-0.018	-0.019	-0.025	-0.030	-0.040		
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(0.020)		
Long-Run Impact	-0.033	-0.024	-0.024	-0.026	-0.034	-0.041	-0.058		
	(0.019)	(0.018)	(0.018)	(0.018)	(0.0019)	(0.024)	(0.035)		
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950		
Within R-sq	0.057	0.055	0.048	0.047	0.046	0.052	0.065		

# TABLE B.23: ACN QUALIFICATION AND CONFLICT: COLUMN 1 OF TABLE 7ETHNIC CONFLICT

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. The samples include the 25 weeks before and after the close qualification. Ethnic conflict is coded using the procedure described in the main text of the original paper. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)						
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugano	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.003	-0.035	-0.030	-0.031	-0.034	-0.036	-0.020
	(0.011)	(0.013)	(0.013)	(0.013)	(0.015)	(0.018)	(0.014)
Long-Run Impact	-0.003	-0.066	-0.054	-0.054	-0.061	-0.069	-0.031
	(0.013)	(0.035)	(0.032)	(0.032)	(0.036)	(0.046)	(0.024)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	109	140	138	133	119	88	59
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950
Within R-sq	0.042	0.103	0.086	0.086	0.087	0.116	0.124

# TABLE B.24: ACN QUALIFICATION AND CONFLICT: COLUMN 2 OF TABLE 7ETHNIC GROUPS WITH STRONG POLITICAL POWER

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. Strong political power refers to conflict events taking place in locations inhabited by ethnic groups with strong political power (i.e., monopoly or dominant according to the ethnic power relations core dataset -EPR-). The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

TABLE B.25: ACN QUALIFICATION AND CONFLICT: COLUMN 3 OF TABLE 7	
ETHNIC GROUPS WITH WEAK POLITICAL POWER	

~ .	Dependent Variable: Log (1+ Number of Events)						
Sample	AER (2020)	DB (2024)		DB (202	4) + Uganc	la (2004)	
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.057	-0.053	-0.048	-0.050	-0.046	-0.055	-0.073
	(0.026)	(0.024)	(0.024)	(0.025)	(0.027)	(0.030)	(0.032)
Long-Run Impact	-0.089	-0.081	-0.071	-0.075	-0.069	-0.081	-0.11
	(0.038)	(0.039)	(0.037)	(0.038)	(0.041)	(0.046)	(0.054)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	109	140	138	133	119	88	59
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950
Within R-sq	0.073	0.065	0.061	0.062	0.063	0.063	0.075

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. Weak political power refers to conflict events taking place in locations inhabited by ethnic groups with no political power (i.e., discriminated, powerless, or self excluded according to the ethnic power relations core dataset -EPR-). The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)							
Sample	AER (2020)	AER (2020) DB (2024) DB (2024) +				ganda (2004)		
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Post-qualification	-0.030	-0.026	-0.022	-0.022	-0.025	-0.029	0.011	
	(0.022)	(0.018)	(0.018)	(0.018)	(0.019)	(0.022)	(0.017)	
Long-Run Impact	-0.047	-0.042	-0.033	-0.034	-0.040	-0.049	0.014	
	(0.038)	(0.033)	(0.032)	(0.032)	(0.035)	(0.044)	(0.023)	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950	
Within R-sq	0.066	0.065	0.058	0.060	0.066	0.084	0.039	

# TABLE B.26: ACN QUALIFICATION AND CONFLICT: COLUMN 4 OF TABLE 7LOCATION WITH NO LINGUISTIC DIVERSITY

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. No linguistic diversity refers to conflict events taking place in first-level administrative sub-national units wherein only one language is spoken. Language data come from Ethnologue. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise. Conflict data comes from the ACLED dataset.

		Dependent	Variable: I	.og (1+ Nu	mber of Ev	vents)	
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugance	la (2004)	
-	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.062	-0.055	-0.061	-0.065	-0.065	-0.098	-0.113
	(0.031)	(0.028)	(0.028)	(0.029)	(0.030)	(0.034)	(0.040)
Long-Run Impact	-0.10	-0.093	-0.10	-0.11	-0.11	-0.16	-0.19
	(0.050)	(0.048)	(0.047)	(0.048)	(0.050)	(0.056)	(0.070)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	109	140	138	133	119	88	59
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950
Within R-sq	0.078	0.086	0.083	0.080	0.081	0.085	0.104

# TABLE B.27: ACN QUALIFICATION AND CONFLICT: COLUMN 5 OF TABLE 7LOCATIONS WITH HIGH LINGUISTIC DIVERSITY

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. High linguistic diversity refers to conflict events taking place in first-level administrative sub-national units wherein more than 5 different languages are spoken. Language data come from Ethnologue. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise. Conflict data comes from the ACLED dataset.

		Dependent	Variable: L	.og (1+ Nu	mber of Ev	vents)			
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugano	Jganda (2004)			
-	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Post-qualification	-0.092	-0.086	-0.087	-0.091	-0.093	-0.129	-0.121		
	(0.039)	(0.035)	(0.034)	(0.035)	(0.036)	(0.041)	(0.042)		
Long-Run Impact	-0.15	-0.15	-0.14	-0.15	-0.15	-0.21	-0.20		
	(0.064)	(0.062)	(0.059)	(0.060)	(0.063)	(0.071)	(0.072)		
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
# country $\times$ qual. campaigns	109	140	138	133	119	88	59		
Observations	4,715	6,627	6,530	6,292	5,631	4,153	2,777		
Within R-sq	0.078	0.086	0.081	0.080	0.084	0.083	0.085		

# TABLE B.28: ACN QUALIFICATION AND CONFLICT: COLUMN 1 OF TABLE 8 OMITTING ACN WEEKS

		Dependent	Variable: L	.og (1+ Nu	mber of Ev	ents)	
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	la (2004)	
-	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-qualification	-0.091	-0.082	-0.084	-0.089	-0.091	-0.131	-0.114
	(0.038)	(0.035)	(0.034)	(0.034)	(0.036)	(0.040)	(0.041)
Long-Run Impact	-0.15	-0.14	-0.14	-0.15	-0.15	-0.21	-0.19
	(0.064)	(0.061)	(0.058)	(0.059)	(0.062)	(0.071)	(0.071)
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# country $\times$ qual. campaigns	109	140	138	133	119	88	59
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950
Within R-sq	0.080	0.087	0.082	0.082	0.085	0.087	0.088

# TABLE B.29: ACN QUALIFICATION AND CONFLICT: COLUMN 2 OF TABLE 8ACN WEEKS AS INTERACTION TERM

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

Sample	Dependent Variable: Log (1+ Number of Events) AFR (2020) DB (2024) DB (2024) + Uganda (2004)							
Sample	Full	Full	+/- 8 pts	+/- 4 pts	+/-3  pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Post-qualification	-0.032	-0.034	-0.040	-0.041	-0.045	-0.048	-0.074	
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.021)	(0.030)	
Long-Run Impact	-0.041	-0.047	-0.055	-0.056	-0.062	-0.064	-0.102	
	(0.022)	(0.025)	(0.024)	(0.025)	(0.028)	(0.030)	(0.044)	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Prob > F H0: 4 lags jointly = $0.00$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950	
Within R-sq	0.039	0.042	0.042	0.041	0.043	0.051	0.056	

# TABLE B.30: ACN QUALIFICATION AND CONFLICT: COLUMN 3 OF TABLE 8 FATALITY THRESHOLD > 10

		Dependent	Variable: L	.og (1+ Nu	mber of Ev	vents)	s)						
Sample	AER (2020)	DB (2024)		DB (202	4) + Ugand	anda (2004)							
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)						
Post-qualification	-0.030	-0.026	-0.035	-0.035	-0.037	-0.042	-0.061						
	(0.014)	(0.014)	(0.012)	(0.012)	(0.014)	(0.016)	(0.023)						
Long-Run Impact	-0.038	-0.036	-0.046	-0.046	-0.049	-0.054	-0.084						
	(0.019)	(0.019)	(0.018)	(0.018)	(0.020)	(0.022)	(0.034)						
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes						
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00						
# country $\times$ qual. campaigns	109	140	138	133	119	88	59						
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950						
Within R-sq	0.038	0.039	0.040	0.040	0.039	0.043	0.056						

# TABLE B.31: ACN QUALIFICATION AND CONFLICT: COLUMN 4 OF TABLE 8 FATALITY THRESHOLD > 25

Robust standard errors in parentheses clustered at the country  $\times$  qualification campaign level. The samples include the 25 weeks before and after the close qualification. The variable *Post-Qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise.Conflict data comes from the ACLED dataset.

	Dependent Variable: Log (1+ Number of Events)							
Sample	AER (2020)	DB (2024)		DB (2024) + Uganda (2004)				
	Full	Full	+/- 8 pts	+/- 4 pts	+/- 3 pts	+/- 2 pts	+/- 1 pts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Post-qualification	-0.019	-0.012	-0.020	-0.020	-0.021	-0.025	-0.038	
	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)	(0.011)	(0.017)	
Long-Run Impact	-0.022	-0.016	-0.024	-0.024	-0.026	-0.029	-0.047	
	(0.012)	(0.013)	(0.011)	(0.011)	(0.012)	(0.013)	(0.020)	
Country $\times$ Qual. Campaign FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Calendar-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
4 Lags of Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Prob > F H0: 4 lags jointly = 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
# country $\times$ qual. campaigns	109	140	138	133	119	88	59	
Observations	5,014	7,000	6,900	6,650	5,950	4,400	2,950	
Within R-sq	0.036	0.039	0.031	0.032	0.031	0.036	0.034	

# TABLE B.32: ACN QUALIFICATION AND CONFLICT: COLUMN 5 OF TABLE 8 FATALITY THRESHOLD > 50