

Correction to “Building Nations through Shared Experiences: Evidence from African Football”*

Filipe Campante[†]
Emilio Depetris-Chauvin[‡]
Ruben Durante[§]

November 2024

Abstract

We present corrections to the data used in Depetris-Chauvin et al (2020), and a detailed description of the criteria used to build the sample of countries analyzed in Section III of the paper, for greater clarity. We provide updated results with the corrected data, showing that the findings in the paper are qualitatively unaltered, and quantitatively very similar.

*We thank Simone Bertoli and Léonie Delpeyrou for pointing out many of the issues addressed in this correction, thus prompting us to thoroughly review the data and clarify ambiguities. We also thank coeditor Rema Hanna for helpful comments, and Alessandro Izzo and Guillermo Mondragón for excellent research assistance. All remaining errors are our own.

[†]Johns Hopkins University & NBER. Email: fcampante@jhu.edu

[‡]Instituto de Economía, Pontificia Universidad Católica de Chile. Email: edepetris@uc.cl

[§]National University of Singapore, ICREA-UPF, CESifo, IZA & CEPR. Email: ruben-durante@gmail.com

1 Introduction

It has been pointed out to us that our paper “Building Nations through Shared Experiences: Evidence from African Football” (Depetris-Chauvin et al 2020) contained some data errors, and that the sample selection criteria for part of the analysis were not described with sufficient precision.¹ In this correction, we describe the errors, and provide an explicit description of the criteria in question. We update all tables in Depetris-Chauvin et al (2020) using the suitably corrected data, and show that results are qualitatively unaffected. Moreover, the magnitudes of the key effects in the paper remain similar: the 37 percent drop in ethnic identification, 30 percent increase in inter-ethnic trust, and 9 percent decline in conflict become 29, 30, and 8, respectively.

This correction is structured as follows. Section 2 discusses the updates to the individual-level analysis presented in Section II of Depetris-Chauvin et al (2020); Section 3 does the same for the country-level analysis in Section III of the original paper. (All tables and figures in the original paper are reproduced here with the revised data, and we also provide a revised version of the Online Appendix.)

2 Individual-Level Analysis

There were two issues in the original data used for our individual-level analysis. The first concerns the date of the match South Africa vs. Burundi, played on October 13, 2002 rather than on October 12, 2022. This affects the number of observations within the relevant window of observation. The second concerns the score of the match Botswana vs. Mozambique, played on October 11, 2008, which was 0-1 rather than 1-0. For 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 the actual date of the match was October 13, 2002. The second instance was instead due to a coding error 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, and the results can be seen in Tables 1 - 4. The updates result in relatively minor changes in the estimated coefficients. The one important caveat is

¹We are very thankful to Simone Bertoli and Léonie Delpeyrou for flagging those issues.

that, while it is still the case that only victories have an impact significantly different from zero (Table 2, column 5), the difference with respect to defeats and especially to draws becomes less pronounced. In fact, comparing the pattern over time for victories only, as in the original Figure 2 (now Figure 2, Panel A), to alternative formulations where we jointly consider victories, draws, and defeats (Panel B) and victories and draws separate from defeats (Panel C), the pattern seems clearest for the latter. This suggests that the effect of football matches on ethnic identification might be meaningfully attenuated in the case of defeats, but not as much in the case of draws.

That said, we choose to maintain subsequent specifications using victories, to minimize the departure with respect to the original version of the paper. For a sense of the comparison in terms of magnitudes, the coefficient in column 6 of Table 2 is now -0.042 (-0.053 in the paper), corresponding to a 29 percent drop in the likelihood of ethnic identification (37 percent in the paper). (The result from Table 4, on the measure of interethnic trust, remains unaltered as the matches in question were not used in that estimation.)

3 Country-Level Analysis

In describing the empirical strategy based on close qualification to the ACN finals, we did not provide a sufficiently clear description of how we selected the sample of close qualification cases. Our algorithm is as follows:

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-directly-qualified 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.

In reviewing the case selection, we identified six cases of unwarranted inclusion into the sample, based on the criteria as spelled out. Four of these cases (South Africa 2006, Sudan 2008, Cameroon 2012, Zambia 2015) ought to be excluded because the countries in question came into the last round of matches either qualified or eliminated from contention, in light of playoff and tiebreaking possibilities.² An additional two cases (Kenya 2012 and Togo 2015) also fall outside of the criteria, though arguably constituting close cases due to idiosyncratic quirks of the qualification process.³ These

²These were overlooked because of idiosyncratic tiebreaking criteria over time. In 2006 and 2015, the criterion was performance in head-to-head matchups: South Africa in 2006 and Zambia in 2015 had already won the head-to-head against Burundi and Mozambique, respectively, even though they could be caught in total number of points and goal differential. Similarly, in 2012, the virtual playoff was affected by Mauritania withdrawing in the middle of the competition: to adjust for that, all second-place teams in four-team groups had to discard their results against the last-placed team in their groups. Cameroon, which could no longer reach Senegal as top of the group, was set to drop 6 points (from two wins against Mauritius) as a result of that, which already made it impossible to qualify in the virtual playoff. In 2008, Sudan overtook Tunisia with a last-round win, to obtain direct qualification, yet would have qualified as one of the top 2nd-place teams in the virtual playoff with any combination of results.

³Kenya in 2012 finished four points behind the last qualifying team in its group, which was Angola, but only three points behind Uganda, which went to the virtual playoff but failed to qualify. However, Kenya would have qualified with a simple victory over Uganda in the last match (which finished 0-0). This is because it would have gone into the virtual playoff, by winning the head-to-head matchup, and would then have qualified, because its results against last-placed Guinea-Bissau would have been discarded (as per the previous footnote), and the resulting 7 points would have been enough to qualify given Kenya’s goal difference. Uganda, in contrast, had better results against Guinea-Bissau, hence dropped down to 5 points in the virtual table, and was thus eliminated. Togo in 2015 was in a very similar situation, though qualification was arguably not as close, as it would have required winning the last match by a substantial margin to reverse its goal differential.

six cases are balanced out by another six (Kenya 2002, Namibia 2002, Ghana 2004, Guinea 2004, Mali 2008, Angola 2015) that had been excluded, but fit the specific inclusion criteria as spelled out here – again, typically because of idiosyncrasies of the tie-breaking criteria for qualification.⁴

We have redone our analysis adjusting for these cases, and the results are presented in Tables 5-8.⁵ We can again see that the updates result in relatively minor changes in the estimated coefficients. For a sense of magnitudes in comparison with the original version of the paper, the coefficient in column 2 of Table 5 is now -0.079 (-0.090 in the paper), corresponding to a 7.6 percent drop in the likelihood of ethnic identification (8.6 percent in the paper).

⁴Namibia and Kenya, in 2002, finished within three and two points of the last qualified team, respectively. They were overlooked in all likelihood because they finished in fourth place, and with negative goal differentials they would have had to reverse. (There were no virtual playoffs that year.) Guinea in 2004 finished first in the group, three points ahead of second-place Niger, which failed to qualify in the virtual playoff. It was likely overlooked because of its much superior goal difference of $+7$, versus -1 for Niger. Ghana failed to qualify in 2004, but finished but three points behind the last team to qualify (Rwanda). It was likely overlooked because it finished last in the group, behind Uganda, and three-team groups had no access to the virtual playoff in that year. In 2008, Mali finished three points ahead of the first non-qualifying team, which was Togo. Again, it was probably overlooked because of its superior goal differential of $+9$, versus -2 for Togo. Finally, Angola in 2015 was three points behind the last qualifying team in the virtual playoff (DR Congo), and was probably overlooked because there were three teams sitting between them.

⁵The results incorporate a few additional adjustments. First, the list of first/overdue qualifications (used in Table 6) now excludes the Democratic Republic of the Congo in 1998, which had been mistakenly included because its participation in 1996 had been recorded under the name Zaire. Second, we adjusted qualification/elimination dates in three cases (Uganda 2004 and 2008 and Benin 2008), because of discrepancies between the date when the teams in question played their last match and when qualification was decided in their group or virtual playoff. We now use the date at which the teams in question actually qualified or were eliminated. Finally, we adjusted the code for classifying ethnic groups according to their political power (used in Table 7), to assign it based on all available observations up to the year preceding our first qualification process. The previous version generated an unstable merging procedure, as the code arbitrarily picked a year for the classification (see details in the revised replication package).

References

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

Table 1: BALANCE IN COVARIATES

	Observations	Mean	<i>Panel A. Post-game</i>		<i>Panel A. Post-victory</i>	
			Estimate	Standard errors	Estimate	Standard errors
Male	37,085	0.503	0.006	0.004	0.010	0.005
Education	37,085	3.075	-0.135	0.150	-0.254	0.154
Age	37,085	36.929	0.796	0.702	1.175	0.753
Unemployed	37,085	0.299	0.000	0.013	-0.008	0.013
Major ethnicity	37,085	0.456	-0.017	0.046	-0.027	0.040
Rural	37,085	0.608	0.099	0.074	0.168	0.081
Religious group member	36,957	0.419	-0.025	0.016	-0.017	0.026
Public goods	37,085	0.480	-0.000	0.020	-0.023	0.017
Same language	37,085	0.465	-0.038	0.032	-0.022	0.045
Influenced by others	37,038	0.045	-0.004	0.006	0.000	0.008
Male interviewer	37,085	0.558	-0.002	0.011	-0.003	0.018
Education interviewer	36,431	7.107	-0.022	0.046	-0.069	0.061
Age interviewer	37,085	28.624	0.010	0.115	0.189	0.134

Notes: Robust standard errors in parentheses clustered at country-match level. Each panel presents point estimates and standard errors for 13 regressions of a covariate (listed at the left) on *Post-game* (panel A) and *Post-victory* (panel B). *Post-game* takes value 1 if the respondent was interviewed within 15 days after a game (regardless of the result), 0 otherwise. *Post-victory* takes value 1 if the respondent was interviewed within 15 days after a victory, 0 otherwise. All estimates are based on OLS regressions using 69 country-match dummies to ensure that the comparison in the covariates is made between respondents in the proximity of the same game and in the same country.

Table 2: NATIONAL TEAM'S PERFORMANCE AND ETHNIC IDENTIFICATION

	Ethnic over national identity (0-1 dummy)						
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Probit (7)
Post-game	-0.026 (0.012)	-0.029 (0.014)	-0.036 (0.014)	-0.017 (0.020)			
Post-victory				-0.030 (0.026)	-0.046 (0.018)	-0.042 (0.018) [0.068]	-0.186 (0.071)
Post-draw					-0.026 (0.039)		
Post-defeat					-0.016 (0.021)		
Post-victory marginal effect							-0.037 (0.014)
Country \times match FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
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
Observations	37,011	37,011	37,011	37,011	37,011	37,011	35,247
R ²	0.093	0.102	0.103	0.104	0.104	0.104	-

Notes: Robust standard errors clustered by country \times year in parentheses. False discovery rate (FDR) adjusted p -value for *Post-victory* is reported in square brackets (Anderson 2008). The outcomes (all dummies) accounted for in the p -value adjustment are: ethnic over national identity, trust in countrymen, interethnic trust, like neighbors from other ethnicities, dislike foreign neighbors, trust ruling party, president's approval, and four indicators for the assessment of present and future own and country's economic conditions. Sample includes respondents interviewed within 15 days before and after an official game. *Post-game*, *Post-victory*, *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.

Table 3: NATIONAL TEAM VICTORIES AND ETHNIC IDENTITY: STAKES AND HETEROGENEOUS EFFECTS

	Ethnic over national identity (0-1 dummy)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-victory	0.019 (0.015)	-0.033 (0.018)	-0.043 (0.018)	-0.043 (0.017)	-0.039 (0.020)	-0.040 (0.018)	-0.034 (0.020)	-0.064 (0.018)
		[0.13]	[0.074]	[0.074]	[0.115]	[0.074]	[0.181]	[0.011]
Interaction		-0.073 (0.020)	0.023 (0.032)	-0.022 (0.031)	-0.003 (0.010)	0.043 (0.027)	-0.239 (0.145)	-0.363 (0.236)
		[0.011]	[0.616]	[0.616]	[0.807]	[0.181]	[0.181]	[0.204]
Uninteracted term						-0.022 (0.011)		
Interaction term	None	Rivalry	Home game	Wide margin	Goals in game	State presence	Diversity country	Diversity team
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Seasonal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language \times year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country \times match FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Friendly matches	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Observations	40,392	37,011	37,011	37,011	37,011	37,011	34,401	26,137
R ²	0.087	0.104	0.104	0.104	0.104	0.103	0.107	0.113

Notes: Robust standard errors clustered at the country \times year level in parentheses. False Discovery Rate (FDR) adjusted p -values are reported in square brackets (Anderson 2008). In addition to all the interacted models presented in this table, the following alternative heterogeneous effects were also accounted for in the p -value adjustment (results in the online Appendix): rural status, unemployment status, gender, education, age, and being part of the ethnic majority. *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 4: NATIONAL TEAM’S VICTORIES AND TRUST IN OTHERS

	Trust in countrymen (1)	Interethnic trust (2)	Like neighbors other ethnicities (3)	Dislike foreign neighbors (4)
Post-victory	0.072 (0.021) [0.062]	0.140 (0.040) [0.062]	0.102 (0.030) [0.062]	0.019 (0.018) [0.387]
Individual controls	Yes	Yes	Yes	Yes
Seasonal FE	Yes	Yes	Yes	Yes
Language \times year FE	Yes	Yes	Yes	Yes
Country \times match FE	Yes	Yes	Yes	Yes
Observations	9,355	7,973	7,511	7,497
R ²	0.140	0.169	0.162	0.153

Notes: Robust standard errors clustered at the country-year level in parentheses. False Discovery Rate (FDR) adjusted p -values are reported in square brackets (Anderson 2008). The outcomes (all dummies) accounted for in the p -value adjustment are: ethnic over national identity, trust in countrymen, interethnic trust, like neighbors from other ethnicities, dislike foreign neighbors, trust ruling party, president’s approval, and 4 indicators for the assessment of present and future own and country’s economic conditions. *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. *Interethnic 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. For ordered probit estimates, see online Appendix Table A.19.

Table 5: ACN QUALIFICATION AND CONFLICT

	Dependent variable							
	log(1 + number of events)					Number of events		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-qualification	-0.116 (0.063)	-0.079 (0.038)					-0.262 (0.157)	-0.178 (0.127)
12 weeks before qualification			0.065 (0.060)	0.056 (0.054)				
1-12 weeks post-qualification (a)					-0.110 (0.064)	-0.083 (0.038)		
13-25 weeks post-qualification (b)					-0.122 (0.078)	-0.076 (0.050)		
Long-run impact	-0.116 (0.063)	-0.126 (0.062)	0.065 (0.060)	0.069 (0.066)	-	-	-	-
Pr > F H ₀ : a = b	-	-	-	-	0.861	0.886	-	-
Regression method	OLS	OLS	OLS	OLS	OLS	OLS	Negative binomial	
Country × qualification campaign	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4 lags of dependent variable	No	Yes	No	Yes	No	Yes	No	Yes
Pr > F H ₀ : 4 lags jointly = 0	-	0.000	-	0.000	-	0.000	-	-
Observations	5,450	5,014	2,725	2,289	5,450	5,014	5,450	5,014
Within R ²	0.003	0.075	0.001	0.026	0.003	0.075	-	-

Notes: Robust standard errors in parentheses clustered at the country × qualification campaign level. Sample in columns 1–2 covers 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. The sample for columns 3–8 includes the 25 weeks before and after the close qualification for 109 country × qualification campaign. The variable *Post-qualification* takes value 1 for the team that qualified for the weeks after the qualification and 0 otherwise. The variable *13–25 weeks post-qualification* takes value 1 starting the thirteenth week after the end of the qualification process for the countries that barely qualify to the ACN, 0 otherwise. Pr > 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 come from the ACLED dataset.

Table 6: OVERDUE AND FIRST QUALIFICATION EFFECTS

	log(1 + number of events)				
	(1)	(2)	(3)	(4)	(5)
Post-qualification	-0.079 (0.038)	-0.138 (0.065)	-0.075 (0.040)	-0.273 (0.137)	-0.088 (0.042)
Post-qualification \times overdue			-0.120 (0.082)		
Post-qualification \times first time					-0.245 (0.178)
Pr $> F$	-	-	0.121	-	0.110
Sample	Full	Overdue Qualification	Full	First Qualification	Full
Country \times qualification campaign FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Calendar-month FE	Yes	Yes	Yes	Yes	Yes
4 lags of dependent variable	Yes	Yes	Yes	Yes	Yes
Pr $> F$ H_0 : 4 lags jointly = 0	0.000	0.000	0.000	0.000	0.000
Observations	5,014	2,484	5,014	736	5,014
Within R ²	0.075	0.070	0.076	0.124	0.077

Notes: Robust standard errors in parentheses clustered at the country \times qualification campaign level. Sample covers \pm 25 weeks around the end of qualification process. Post-qualification takes value 1 during the 25 weeks following the qualification to ACN, 0 otherwise. Conflict data come from the ACLED dataset. 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 online Appendix Table A.20. Pr $> F$ refers to the F-tests with the null hypothesis that coefficients for post-qualification and its interaction with overdue (column 3) or first-time qualification (column 5) are jointly equal to 0. Interaction terms were demeaned to ease the comparison of uninteracted terms.

Table 7: ETHNIC CONFLICT, ETHNIC POLITICAL POWER, AND LINGUISTIC DIVERSITY

	log(1 + number of events)				
	(1)	(2)	(3)	(4)	(5)
Post-qualification	-0.021 (0.012)	-0.013 (0.025)	-0.062 (0.032)	-0.030 (0.022)	-0.056 (0.030)
Long-run impact	-0.027 0.018	-0.017 0.033	-0.104 0.056	-0.046 0.039	-0.089 0.047
Event definition	Ethnic	Strong Political power	Weak Political power	No linguistic diversity	High linguistic diversity
Country \times qualification campaign FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Calendar-month FE	Yes	Yes	Yes	Yes	Yes
4 lags of dependent variable	Yes	Yes	Yes	Yes	Yes
Pr $> F$ H_0 : 4 lags jointly = 0	0.001	0.000	0.000	0.000	0.000
Observations	5,014	5,014	5,014	5,014	5,014
Within R ²	0.046	0.049	0.083	0.067	0.079

Notes: Robust standard errors in parentheses clustered at the country \times qualification campaign level. Sample covers ± 25 weeks around the end of the qualification process. *Post-qualification* takes value 1 during the 25 weeks following the qualification to ACN, and 0 otherwise. Each column presents point estimates and standard errors for a regression of the baseline specification using different definitions of conflict events as the dependent variable. Ethnic conflict is coded using the procedure described in the main text. *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-). *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-). *No linguistic diversity* (*High linguistic diversity*) refers to conflict events taking place in first-level administrative sub-national units wherein only one language is (more than 5 different languages are) spoken. Language data come from *Ethnologue*. All conflict data are from the ACLED dataset.

Table 8: POTENTIAL INCAPACITATION EFFECT AND NEWS CROWDING OUT

	log(1 + number of events)				
	(1)	(2)	(3)	(4)	(5)
Post-qualification	-0.077 (0.038)	-0.079 (0.037)	-0.035 (0.018)	-0.029 (0.014)	-0.017 (0.010)
Long-run impact	-0.122 0.062	-0.125 0.062	-0.045 0.024	-0.038 0.019	-0.021 0.012
Omitted observations	ACN weeks	None	None	None	None
Model Specification	Baseline	Treatment interacted ACN weeks	Baseline	Baseline	Baseline
Fatality threshold	None	None	> 10 fat.	> 25 fat.	> 50 fat.
Country \times qualification campaign FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Calendar-month FE	Yes	Yes	Yes	Yes	Yes
4 lags of dependent variable	Yes	Yes	Yes	Yes	Yes
Pr > F H_0 : 4 lags jointly = 0	0.000	0.000	0.000	0.000	0.000
Observations	4,733	5,014	5,014	5,014	5,014
Within R ²	0.071	0.075	0.039	0.041	0.036

Notes: Robust standard errors in parentheses clustered at the *country* \times *qualification* campaign level. Sample covers ± 25 weeks around the end of the qualification process. *Post-qualification* takes value 1 during the 25 weeks following the qualification to ACN, and 0 otherwise. All conflict data are from the ACLED dataset.

Figure 1: ETHNIC IDENTIFICATION OVER TIME AND ACROSS SELECTED COUNTRIES

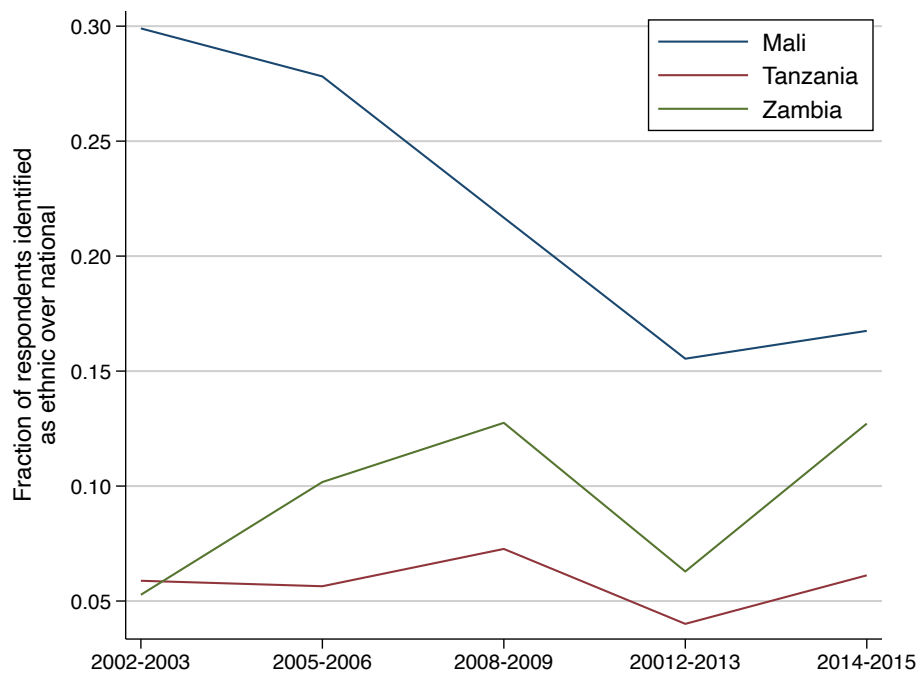
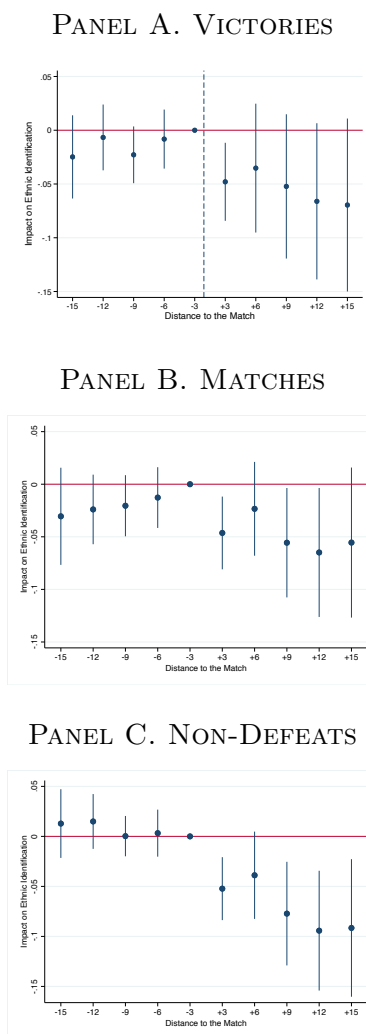









Figure 2: ETHNIC IDENTITY BEFORE AND AFTER NATIONAL TEAM PERFORMANCE



Notes: The figure plots the coefficients and the 95 percent confidence intervals for nine dummies indicating 3-day blocks from 15 days before to 15 days after a victory (Panel A), match (Panel B), or match without defeat (Panel C) by the national football team. The coefficient for the period between 3 to 1 days before the match is normalized to 0. Confidence intervals are based on heteroskedasticity-robust standard errors clustered by country \times year. The coefficients are estimated from a unique regression in which we control for individual characteristics, seasonal dummies, country \times match and language group \times year fixed effects.

Figure 3: EXAMPLE OF CLOSE QUALIFICATION: GROUP A, ACN 2012

Team	Pld	W	D	L	GF	GA	GD	Pts
 Mali	5	3	0	2	7	4	3	9
 Zimbabwe	5	2	2	1	6	3	3	8
 Cape Verde	5	2	1	2	5	6	-1	7
 Liberia	5	1	1	3	5	10	-5	4

08/10/2011	 Liberia	2 - 2	 Mali
	 Cape Verde	2 - 1	 Zimbabwe





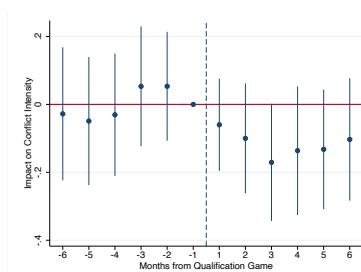
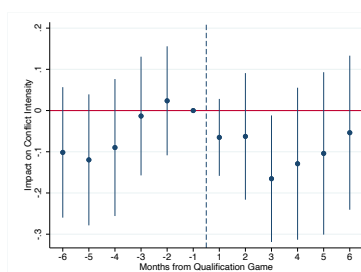
Team	Pld	W	D	L	GF	GA	GD	Pts
 Mali	6	3	1	2	9	6	3	10
 Cape Verde	6	3	1	2	7	7	0	10
 Zimbabwe	6	2	2	2	7	5	2	8
 Liberia	6	1	2	3	7	12	-5	5

Figure 4: NUMBER OF CONFLICT EPISODES BEFORE AND AFTER QUALIFICATION

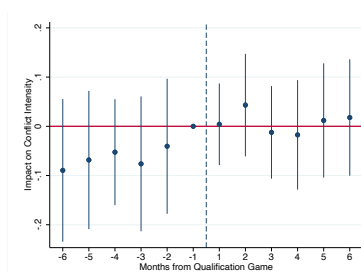
PANEL A. POOLED (4-WEEK BANDWIDTHS)



PANEL B. TREATMENT COUNTRIES



PANEL C. CONTROL COUNTRIES



Notes: Panel A plots coefficients and 95 percent confidence intervals for interactions between the dummy for countries that barely qualified to the ACN and 11 dummies for \times 4-week period included between 25 weeks before and after the qualification. The regressions also include week FE, calendar-month FE, and country \times qualifier dummies. Panel B plots coefficients and 95 percent confidence intervals for 11 dummies for 4-week period included between 25 weeks before and after the qualification for the groups of countries that barely qualified to the ACN. The regressions calendar-month FE and country \times qualifier dummies (week FE are omitted to avoid perfect multicollinearity). Panel C replicates Panel B for the groups of countries that barely did not qualify to the ACN. The dependent variable in all regressions is $\log(1 + \text{number of conflict events})$. The coefficients for the 4 weeks immediately before the end of the qualification process are normalized to 0. Confidence intervals are based on heteroskedasticity-robust standard errors clustered by country \times qualifier