Online Appendix to 'Convergence vs. The Middle Income Trap: The Case of Global Soccer'

Melanie Krause* Stefan Szymanski†

January 2019

This Online Appendix complements the article 'Convergence vs. the middle income trap: the case of global soccer', forthcoming in *Applied Economics*. It contains supplementary information, additional statistics and further robustness checks of the empirical analysis.

Online Appendix

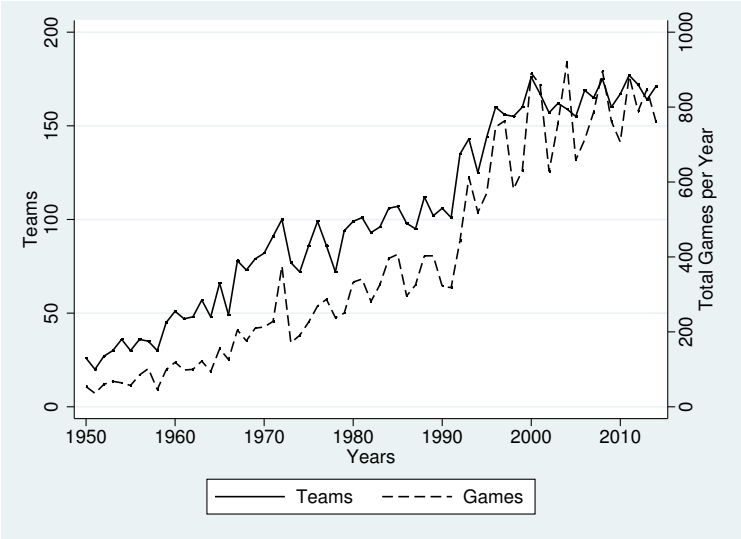
*Department of Economics, Hamburg University, Von-Melle-Park 5, 20146 Hamburg, Germany; email: *melanie.krause@uni-hamburg.de*

†Department of Kinesiology, University of Michigan, 3118 Observatory Lodge 1402 Washington Heights, Ann Arbor, MI 48109-2013, United States; email *stefansz@umich.edu*

A The Growth of International Competition

Association football (soccer) is a game whose rules were first written down in 1863 in England. Originally played only between local clubs, the first "international" match was played between England and Scotland in 1872. The game spread rapidly and by the end of the nineteenth century most European and South American nations had established national associations to administer the game, thus facilitating competition between national teams. In 1904 FIFA was created as an organization to manage soccer relations between countries, and in 1930 the FIFA World Cup was first played, with 13 national teams competing. In the first half of the 20th century, there were still rather few international games; under 2,200 were recorded between 1900 and 1940, an average of 54 per year, and almost all of these involved European and South American countries. But in the second half of the 20th century, this has changed, turning soccer into a truly global industry: Since 1950 there have been over 36,000 games played between men's national soccer teams, an average of over 500 per year, see Figure A-1.

Figure A-1 – The Growth of International Soccer Competition



Notes: The graph shows yearly figures on the number of international games played between national teams as well as the number of internationally active national teams. Apart from the steady increase the graphs exhibit cyclical peaks in the years of a FIFA World Cup.

Table A-1 shows the number of games between the teams from the various continental confederations. Despite the globalized nature of soccer, the vast majority of games take place between teams from the same continent.

Table A-2 lists the years since 1950 in which a FIFA World Cup took place and the number of participating teams from each continental association. Teams from CONMEBOL, the South American association, and UEFA, the European one, where the game first took root, have tended to dominate the World Cup; in fact, no team from outside these associations has ever won the Cup. Teams from outside the big two regional

Table A-1 – Regional Matches Involving Teams from the Various Federations, 1950-2014

	Asia	Africa	America (N,C)	America (S)	Oceania	Europe
Asia	9586	691	161	202	130	788
Africa	691	12524	99	124	9	460
America (N,C)	161	99	4214	666	17	456
America (S)	202	124	666	3454	15	711
Oceania	130	9	17	15	32	26
Europe	788	460	456	711	26	11884

Notes: The table shows the number of international matches pitting Team 1 from the regional federation in the row against Team 2 from the regional federation in the column. The continental confederations are AFC (Asia), CAF (Africa), CONCACAF (North and Middle America and the Caribbean), CONMEBOL (South America), OFC (Oceania) and UEFA (Europe).

confederations have reached the semi-finals twice: the USA in the first World Cup in 1930 (contested by only 13 nations), and South Korea in 2002. But FIFA has consciously tried to expand opportunities for the smaller associations. While each continent controls its own qualifying process, the number of slots allocated to each continental association is agreed centrally. The share allocated to UEFA and CONMEBOL has shrunk considerably over time, largely through expansion of the number of participating teams. A further expansion of 16 teams has been agreed for the 2026 World Cup, which will reduce the European and South American share further, possibly to as little as 46 %. Critics have argued that the distribution remains unfair and should reflect global population shares more accurately. The counter argument is that for a given quality of team it is harder to qualify through UEFA or CONMEBOL than any other federation.

Table A-2 – Number of Countries Qualifying for the FIFA World Cup 1950-2014

World Cup	AFC (Asia)	CAF (Africa)	CONCA-CAF (Central+ North Am.)	CON-MEBOL (South America)	OFC (Oceania)	UEFA (Europe)	Total	UEFA + CONME-BOL share
1950	1	0	2	5*	0	7	15	0.800
1954	1	0	1	2	0	12*	16	0.875
1958	0	0	1	3	0	12*	16	0.938
1962	0	0	1	5*	0	10	16	0.938
1966	1	0	1	4	0	10*	16	0.813
1970	0	1	2*	3	0	10	16	0.813
1974	1	1	1	4	0	9*	16	0.813
1978	1	1	1	3*	0	10	16	0.813
1982	1	2	2	4	1	14*	24	0.750
1986	2	2	2*	4	0	14	24	0.750
1990	2	2	2	4	0	14*	24	0.750
1994	2	3	2*	4	0	13	24	0.708
1998	4	5	3	5	0	15*	32	0.625
2002	4*	5	3	5	0	15	32	0.625
2006	4	5	4	4	1	14*	32	0.563
2010	4	6*	3	5	1	13	32	0.563
2014	4	5	4	6*	0	13	32	0.594

Notes: For each FIFA World Cup, the table lists the number of participating teams by continental federation. The * indicates the host federation. The CONCACAF federation includes Central and North America as well as the Caribbean. Note that the table shows the number of teams that actually qualified; in some cases the final slots were allocated by inter-continental play-offs.

Table A-3 lists the 32 national teams playing in the 2014 World Cup in Brazil. It underlines the internationalization of soccer: It indicates which teams had a foreign coach (14 out of 32) and how many of the 23 players of the squad played, respectively, in their home league and in a (potentially different) European league. In only eight of the 32 countries did more than half of the squad members play for a club in their country, and four of these were countries with top national leagues. At the other extreme, only one player from each of Bosnia-Herzegovina, Uruguay, Ivory Coast and Ghana played for domestic clubs. Cases such as Russia, whose national team players exclusively play domestically, shows the importance of political and institutional factors in player migration, see Leeds and Leeds (2009). Overall, we see that national team players with international club experience is a noticeable phenomenon.

Table A-3 – Squads of 32 National Teams Participating in the 2014 FIFA World Cup

Team	Coach	Players (out of 23)	
	Foreign	Home League	(Other) European League
UEFA (Europe)			
Germany	No	16	7
Spain	No	14	9
Italy	No	20	3
England	No	22	1
France	No	8	15
Portugal	No	8	15
Greece	Yes	14	9
Russia	Yes	23	0
Netherlands	No	10	13
Belgium	No	3	20
Switzerland	Yes	7	16
Croatia	No	2	21
Bosnia & Herzegovina	No	1	22
CONMEBOL (South America)			
Brazil	No	4	18
Argentina	No	3	19
Chile	Yes	5	15
Colombia	Yes	3	16
Uruguay	No	1	16
Ecuador	Yes	8	4
CONCACAV (North/Central American + Caribbean)			
United States	Yes	9	13
Mexico	No	15	8
Costa Rica	Yes	9	11
Honduras	Yes	11	5
AFC (Asia)			
Australia	No	7	13
Japan	Yes	11	12
Iran	Yes	14	6
South Korea	No	6	10
CAF (Africa)			
Nigeria	No	4	19
Cameroon	Yes	2	21
Ivory Coast	Yes	1	22
Ghana	No	1	18
Algeria	Yes	2	19

Notes: Each official squad consists of 23 players. Players which neither play in the home league nor in a European league make up the difference to 23. The data are from http://resources.fifa.com/mm/document/tournament/competition/02/36/33/44/fwc_2014_squadlists_neutral.pdf

B Summary Statistics of the Data Set and Determinants of Soccer Success

This section takes a closer look at our data set and national teams' performance as well as the determinant variables.

Table B-1 provides summary statistics of the outcome and explanatory variables. Overall, the variables look stable over time; only the slight increase in the standard deviation of the population and GDP per capita ratios indicate that in later years larger and richer countries played more often against smaller ones.

Table B-1 – Summary Statistics of the Outcome and Explanatory Variables

	All Years	1950-1966	1967-1982	1983-1998	1999-2014
Game Outcome, Winning Percentage					
Mean	0.5000	0.5000	0.5000	0.5000	0.5000
St.Dev.	0.4336	0.4490	0.4384	0.4297	0.4325
Min	0.0000	0.0000	0.0000	0.0000	0.0000
Max	1.0000	1.0000	1.0000	1.0000	1.0000
Obs	50804	2970	7990	14866	24978
Goal Difference					
Mean	0.0000	0.0000	0.0000	0.0000	0.0000
St.Dev.	2.1868	2.5762	2.2716	2.1455	2.1326
Min	-20.0000	-14.0000	-14.0000	-17.0000	-20.0000
Max	20.0000	14.0000	14.0000	17.0000	20.0000
Obs	50804	2970	7990	14866	24978
Log Population Ratio					
Mean	0.0000	0.0000	0.0000	0.0000	0.0000
St.Dev.	2.0940	1.7661	1.9321	2.0823	2.1849
Min	-9.1152	-6.9764	-8.6362	-9.1152	-8.4066
Max	9.1152	6.9764	8.6362	9.1152	8.4066
Obs	50804	2970	7990	14866	24978
Log GDP per capita Ratio					
Mean	-0.0000	-0.0000	0.0000	-0.0000	-0.0000
St.Dev.	1.2194	0.8994	1.1123	1.2150	1.2861
Min	-5.7318	-3.4041	-5.1160	-4.9244	-5.7318
Max	5.7318	3.4041	5.1160	4.9244	5.7318
Obs	50804	2970	7990	14866	24978
Log Experience Ratio					
Mean	0.0000	0.0000	0.0000	0.0000	0.0000
St.Dev.	1.0290	1.0613	1.0296	1.1804	0.9227
Min	-6.4877	-4.0678	-5.5910	-6.4877	-6.1092
Max	6.4877	4.0678	5.5910	6.4877	6.1092
Obs	50804	2970	7990	14866	24978

Notes: The table presents summary statistics of the match-level data presented in the text. The years from 1950 to 2014 can be divided into 4 four-year World Cup cycles. In terms of observations, every game is counted twice, once from the perspective of country i and once from country j, to capture in the subsequent regressions both the home advantage and the disadvantage of playing in the opponent's country.

Table B-2 repeats the regression of game outcome on explanatory factors from the paper, using the goal difference rather than the outcome in terms of win, draw and loss. The results are very similar. Panel B shows that the R^2 decreases markedly over the last decades, indicating that the explanatory factors have become less decisive in predicting game success, in line with our convergence hypothesis.

Table B-2 – Game Outcome (Goal Difference) Regressed on Explanatory Factors

Panel A: By Types of Games

Dependent Var: Goal Difference	(1) All Games	(2) Friendlies	(3) Competitive	(4) Qualifiers	(5) World + Cont. Cup
home	0.589***	0.465***	0.766***	0.407***	0.774***
	(0.035)	(0.042)	(0.051)	(0.090)	(0.102)
away	-0.629***	-0.561***	-0.675***	-1.042***	-0.582***
	(0.031)	(0.037)	(0.047)	(0.083)	(0.095)
lgdppcratio	0.136***	0.123***	0.147***	0.134***	0.192***
	(0.016)	(0.019)	(0.023)	(0.025)	(0.031)
lpopratio	0.168***	0.146***	0.205***	0.224***	0.111***
	(0.015)	(0.013)	(0.020)	(0.022)	(0.024)
lexpratio	0.657***	0.589***	0.675***	0.637***	0.716***
	(0.031)	(0.030)	(0.041)	(0.042)	(0.070)
Constant	-0.016	0.346***	-0.195***	0.145	-0.557***
	(0.033)	(0.030)	(0.053)	(0.090)	(0.039)
Country Dummies	Yes	Yes	Yes	Yes	Yes
R2	0.274	0.213	0.356	0.388	0.252
Observations	50804	27708	23096	17784	5312
Countries	182	181	182	182	132

Panel B: By Time Period

Dependent Var: Goal Difference	(1) All Games	(2) 1950-1966	(3) 1967-1982	(4) 1983-1998	(5) 1999-2014
home	0.589***	0.693***	0.678***	0.617***	0.538***
	(0.035)	(0.147)	(0.088)	(0.051)	(0.035)
away	-0.629***	-0.694***	-0.853***	-0.633***	-0.538***
	(0.031)	(0.141)	(0.073)	(0.046)	(0.040)
lgdppcratio	0.136***	-0.142*	0.161***	0.198***	0.131***
	(0.016)	(0.082)	(0.041)	(0.023)	(0.019)
lpopratio	0.168***	0.214***	0.160***	0.182***	0.168***
	(0.015)	(0.045)	(0.023)	(0.020)	(0.017)
lexpratio	0.657***	0.892***	0.726***	0.552***	0.748***
	(0.031)	(0.072)	(0.036)	(0.037)	(0.052)
Constant	-0.016	0.380***	0.675***	-0.357***	0.093**
	(0.033)	(0.127)	(0.104)	(0.053)	(0.036)
Country Dummies	Yes	Yes	Yes	Yes	Yes
R2	0.274	0.305	0.317	0.320	0.277
Observations	50804	2970	7990	14866	24978
Countries	182	86	130	175	182

Notes: Analogous to the paper (Section 3), the table presents OLS regression results with the goal difference rather than the points outcome (0, 0.5, 1) as the dependent variable.

C Additional Beta-Convergence Results

Analogous to the test for β-convergence in countries' winning percentages as explained in Section 4.1 in the text, we here conduct the analysis with other performance variables and subsamples. The following tables are all structured similarly and regress the change in performance of country i in cycle t on its past performance:

$$\Delta y_{it} = \alpha + \beta \cdot y_{i,t-1} + \epsilon_{it}, \tag{C-1}$$

Panel A, col. (1) runs this regression for unconditional convergence, col. (2) tests for conditional convergence by including additional controls. Col. (3) includes regional confederation dummies. Col. (4) and Col. (5) test for, respectively, unconditional and conditional convergence using country fixed effects.
Panel B estimates

$$y_{it} = \alpha_i + \underbrace{(\beta + 1)}_{\rho} \cdot y_{i,t-1} + \epsilon_{it}, \tag{C-2}$$

with specific short T dynamic panel data model estimation techniques, Arellano-Bond GMM in col. (1) and col. (2) and Unconditional Quasi-Maximum Likelihood in col. (3) and col. (4).
Panel C conducts weighted regressions. Col. (1) and col. (2) use time weights $w_{it} = (\bar{n}_i/n_{it})^{1/2}$, where n_{it} is the number of games played by country i in cycle t and $\bar{n}_i$ is the average number of games by i over all cycles. In col. (3) and col. (4) dominance weights are used, reflecting how often country i played against an opponent from the two top confederations, Europe and South America.

In particular, we conduct the analysis with different performance variables and subsamples and compare the results to those in the main text. Using the goal difference (Table C-1) yields very similar coefficients as the win percentage. Concerns that convergence results might be driven by stronger teams' anecdotically worse performance at friendlies, when they often give weaker players a chance, can be alleviated by Table C-2: restricting the sample to competitive games gives even stronger convergence results, in line with our previous analysis that 'friendlies' and competitive games are mostly decided by the same factors. In Table C-3 we consider only the teams that were active from the first cycle (1950-1954) onwards, to exclude the effect of newcomers. Obviously, the national teams entering the international stage and catching up has contributed to the overall convergence effect, but we also observe unconditional and conditional convergence among the 42 teams which were present throughout the years. Finally, we split the sample into the time periods 1950-1982 (the first eight cycles, Table C-4) and 1983-2014 (the last eight cycles, Table C-5). While we find significant convergence results throughout time,

there is no indication that they have become stronger in later years. This is confirmed by Table C-6, which shows that the regression coefficients are clearly negative in each four-year cycle but their magnitude has slightly decreased rather than increased.
We conclude from this analysis that our results of β-convergence in national teams' performance is a result that is robust across econometric specifications, performance variables, sub-samples and time periods.

Table C-1 – Beta-Convergence Regression Results, Goal Difference (GD)

Panel A: Panel Data Regression

Dep Var: Δ GD	(1)	(2)	(3)	(4)	(5)
l.GD	-0.456***	-0.587***	-0.594***	-0.796***	-0.859***
	(0.030)	(0.033)	(0.033)	(0.030)	(0.032)
lgdppcratio		0.043	0.048		0.083
		(0.029)	(0.030)		(0.054)
lpopratio		0.095***	0.098***		0.026
		(0.024)	(0.023)		(0.063)
lexpratio		0.337***	0.334***		0.502***
		(0.044)	(0.044)		(0.065)
Constant	-0.049*	-0.032	-0.113*	-0.140***	-0.098***
	(0.026)	(0.024)	(0.066)	(0.008)	(0.011)
Confed Dummies	No	No	Yes	No	No
Country FE	No	No	No	Yes	Yes
R2	0.367	0.453	0.454	0.554	0.600
Observations	1644	1644	1644	1644	1644
Countries	178	178	178	178	178

Panel B: Fixed Effects Short T Dynamic Panel Estimation

Dep Var: GD	(1) (GMM)	(2) (GMM)	(3) (QML)	(4) (QML)
l.GD	0.234***	0.128**	0.267***	0.187***
	(0.057)	(0.058)	(0.045)	(0.038)
lgdppcratio		0.062		0.148**
		(0.074)		(0.057)
lpopratio		0.114*		0.045
		(0.060)		(0.047)
lexpratio		0.597***		0.433***
		(0.079)		(0.066)
Constant	-0.132***	-0.077*	-0.055	-0.046
	(0.047)	(0.041)	(0.045)	(0.040)
AR1	-6.692	-6.072		
AR2	2.596	1.807		
Observations	1484	1484	1372	1372
Countries	176	176	139	139

Panel C: Weighted Regressions

Dep Var: Δ GD	(1) (Time W)	(2) (Time W)	(3) (Dom W)	(4) (Dom W)
l.GD	-0.474***	-0.068***	-0.287***	-0.463***
	(0.030)	(0.005)	(0.037)	(0.043)
lgdppcratio		0.007		0.048
		(0.005)		(0.057)
lpopratio		0.012***		0.149***
		(0.003)		(0.027)
lexpratio		0.046***		0.186***
		(0.007)		(0.060)
Constant	-0.065**	-0.007	0.046	0.092*
	(0.029)	(0.008)	(0.028)	(0.052)
R2	0.381	0.223	0.187	0.307
Observations	1644	1644	599	599
Countries	178	178	56	56

Notes: Analogous to Section 4.1 in the paper, the table presents beta convergence regressions when the goal difference is used as performance variable. See the text in this Online Appendix for more details.

Table C-2 – Beta-Convergence Regression Results, Competitive Games

Panel A: Panel Data Regression

Dep Var: Δ points	(1)	(2)	(3)	(4)	(5)
l.points	-0.453***	-0.609***	-0.617***	-0.918***	-0.947***
	(0.029)	(0.033)	(0.033)	(0.030)	(0.029)
lgdppcratio		0.024***	0.024***		0.015
		(0.006)	(0.006)		(0.010)
lpopratio		0.018***	0.018***		0.009
		(0.005)	(0.005)		(0.008)
lexpratio		0.066***	0.067***		0.075***
		(0.008)	(0.008)		(0.011)
Constant	0.219***	0.294***	0.277***	0.431***	0.447***
	(0.016)	(0.018)	(0.021)	(0.014)	(0.014)
Confed Dummies	No	No	Yes	No	No
Country FE	No	No	No	Yes	Yes
R2	0.276	0.386	0.388	0.527	0.563
Observations	1530	1530	1530	1530	1530
Countries	176	176	176	176	176

Panel B: Fixed Effects Short T Dynamic Panel Estimation

Dep Var: points	(1) (GMM)	(2) (GMM)	(3) (QML)	(4) (QML)
l.points	0.045	0.101*	0.151***	0.116***
	(0.052)	(0.052)	(0.035)	(0.033)
lgdppcratio		0.036***		0.017
		(0.014)		(0.010)
lpopratio		0.016*		0.005
		(0.009)		(0.007)
lexpratio		0.069***		0.068***
		(0.014)		(0.011)
Constant	0.448***	0.427***	0.416***	0.431***
	(0.027)	(0.025)	(0.022)	(0.020)
AR1	-5.742	-6.221		
AR2	-1.130	-0.449		
Observations	1354	1354	1292	1292
Countries	168	168	140	140

Panel C: Weighted Regressions

Dep Var: Δ points	(1) (Time W)	(2) (Time W)	(3) (Dom W)	(4) (Dom W)
l.points	-0.479***	-0.652***	-0.349***	-0.570***
	(0.031)	(0.036)	(0.048)	(0.057)
lgdppcratio		0.023***		0.021
		(0.007)		(0.014)
lpopratio		0.019***		0.036***
		(0.005)		(0.007)
lexpratio		0.073***		0.049***
		(0.010)		(0.014)
Constant	0.230***	0.296***	0.185***	0.301***
	(0.017)	(0.023)	(0.029)	(0.031)
R2	0.287	0.406	0.205	0.349
Observations	1530	1530	579	579
Countries	176	176	56	56

Notes: Analogous to Section 4.1 in the paper, the table presents beta convergence regressions when the sample is restricted only to competitive games, excluding 'friendlies'. See the text in this Online Appendix for more details.

Table C-3 – Beta-Convergence Regression Results, Only National Teams Present Since 1950

Panel A: Panel Data Regression

Dep Var: Δ points	(1)	(2)	(3)	(4)	(5)
l.points	-0.384***	-0.537***	-0.553***	-0.753***	-0.790***
	(0.058)	(0.057)	(0.053)	(0.054)	(0.057)
lgdppcratio		-0.002	-0.001		-0.003
		(0.011)	(0.012)		(0.016)
lpopratio		0.015**	0.020***		-0.005
		(0.006)	(0.006)		(0.019)
lexpratio		0.080***	0.077***		0.096***
		(0.012)	(0.014)		(0.019)
Constant	0.203***	0.262***	0.234***	0.392***	0.394***
	(0.033)	(0.031)	(0.037)	(0.028)	(0.026)
Confed Dummies	No	No	Yes	No	No
Country FE	No	No	No	Yes	Yes
R2	0.234	0.339	0.345	0.433	0.473
Observations	574	574	574	574	574
Countries	42	42	42	42	42

Panel B: Fixed Effects Short T Dynamic Panel Estimation

Dep Var: points	(1) (GMM)	(2) (GMM)	(3) (QML)	(4) (QML)
l.points	-0.006	-0.011	0.265***	0.201***
	(0.059)	(0.056)	(0.056)	(0.045)
lgdppcratio		0.018		0.014
		(0.022)		(0.017)
lpopratio		-0.001		0.001
		(0.015)		(0.015)
lexpratio		0.084***		0.095***
		(0.025)		(0.024)
Constant	0.521***	0.505***	0.397***	0.403***
	(0.037)	(0.037)	(0.037)	(0.029)
AR1	-4.729	-4.885		
AR2	0.116	-0.0618		
Observations	538	538	483	483
Countries	42	42	34	34

Panel C: Weighted Regressions

Dep Var: Δ points	(1) (Time W)	(2) (Time W)	(3) (Dom W)	(4) (Dom W)
l.points	-0.439***	-0.778***	-0.339***	-0.583***
	(0.117)	(0.129)	(0.067)	(0.063)
lgdppcratio		0.083		0.004
		(0.045)		(0.022)
lpopratio		0.069**		0.031***
		(0.020)		(0.006)
lexpratio		0.028		0.077***
		(0.036)		(0.019)
Constant	0.237***	0.295***	0.193***	0.308***
	(0.067)	(0.046)	(0.039)	(0.037)
R2	0.193	0.365	0.187	0.318
Observations	112	112	398	398
Countries	8	8	27	27

Notes: Analogous to Section 4.1 in the paper, the table presents beta convergence regressions when the sample is restricted to the countries which played matches from the first four-year cycle onwards. See the text in this Online Appendix for more details.

Table C-4 – Beta-Convergence Regression Results, Period 1 (1950-1982)

Panel A: Panel Data Regression

Dep Var: Δ points	(1)	(2)	(3)	(4)	(5)
l.points	-0.565***	-0.735***	-0.741***	-0.993***	-1.011***
	(0.043)	(0.038)	(0.040)	(0.044)	(0.045)
lgdppcratio		0.008	0.007		0.049*
		(0.010)	(0.010)		(0.026)
lpopratio		0.021***	0.022***		0.034*
		(0.007)	(0.007)		(0.019)
lexpratio		0.092***	0.093***		0.053**
		(0.012)	(0.013)		(0.021)
Constant	0.274***	0.362***	0.349***	0.474***	0.490***
	(0.023)	(0.021)	(0.027)	(0.020)	(0.022)
Confed Dummies	No	No	Yes	No	No
Country FE	No	No	No	Yes	Yes
R2	0.403	0.532	0.530	0.648	0.667
Observations	474	474	474	474	474
Countries	108	108	108	108	108

Panel B: Fixed Effects Short T Dynamic Panel Estimation

Dep Var: points	(1) (GMM)	(2) (GMM)	(3) (QML)	(4) (QML)
l.points	-0.106	-0.045	0.093**	0.067
	(0.076)	(0.085)	(0.045)	(0.044)
lgdppcratio		0.032		0.038
		(0.031)		(0.029)
lpopratio		0.020		0.021
		(0.018)		(0.019)
lexpratio		0.074***		0.065***
		(0.027)		(0.021)
Constant	0.527***	0.506***	0.441***	0.455***
	(0.038)	(0.040)	(0.027)	(0.025)
AR1	-2.989	-2.998		
AR2	-1.608	-1.081		
Observations	386	386	425	425
Countries	100	100	87	87

Panel C: Weighted Regressions

Dep Var: Δ points	(1) (Time W)	(2) (Time W)	(3) (Dom W)	(4) (Dom W)
l.points	-0.547***	-0.734***	-0.412***	-0.740***
	(0.053)	(0.047)	(0.066)	(0.081)
lgdppcratio		0.020*		0.007
		(0.012)		(0.019)
lpopratio		0.022***		0.039***
		(0.008)		(0.011)
lexpratio		0.085***		0.094***
		(0.014)		(0.019)
Constant	0.265***	0.354***	0.217***	0.375***
	(0.029)	(0.038)	(0.036)	(0.044)
R2	0.388	0.512	0.237	0.420
Observations	346	346	215	215
Countries	78	78	36	36

Notes: Analogous to Section 4.1 in the paper, the table presents beta convergence regressions when the sample period is restricted to 1950-1982, the first eight four-year cycles. See the text in this Online Appendix for more details.

Table C-5 – Beta-Convergence Regression Results, Period 2 (1983-2014)

Panel A: Panel Data Regression

Dep Var: Δ points	(1)	(2)	(3)	(4)	(5)
lagpts	-0.355***	-0.494***	-0.503***	-0.902***	-0.959***
	(0.028)	(0.038)	(0.038)	(0.040)	(0.037)
(mean) lgdppcratio		0.012**	0.013**		0.019**
		(0.005)	(0.005)		(0.009)
(mean) lpopratio		0.015***	0.016***		0.021**
		(0.004)	(0.004)		(0.009)
(mean) lexpratio		0.041***	0.040***		0.066***
		(0.007)	(0.007)		(0.012)
Constant	0.170***	0.241***	0.238***	0.417***	0.456***
	(0.013)	(0.019)	(0.021)	(0.018)	(0.017)
Confed Dummies	No	No	Yes	No	No
Country FE	No	No	No	Yes	Yes
R2	0.223	0.304	0.305	0.516	0.558
Observations	1170	1170	1170	1170	1170
Countries	177	177	177	177	177

Panel B: Fixed Effects Short T Dynamic Panel Estimation

Dep Var: points	(1) (GMM)	(2) (GMM)	(3) (QML)	(4) (QML)
l.points	-0.109	-0.019	0.230***	0.180***
	(0.081)	(0.078)	(0.051)	(0.047)
lgdppcratio		0.024*		0.013
		(0.014)		(0.010)
lpopratio		0.031***		0.018**
		(0.011)		(0.009)
lexpratio		0.049***		0.054***
		(0.016)		(0.013)
Constant	0.510***	0.483***	0.363***	0.393***
	(0.036)	(0.035)	(0.026)	(0.024)
AR1	-4.266	-5.528		
AR2	0.459	1.211		
Observations	897	897	1007	1007
Countries	175	175	161	161

Panel C: Weighted Regressions

Dep Var: Δ points	(1) (Time W)	(2) (Time W)	(3) (Dom W)	(4) (Dom W)
lagpts	-0.353***	-0.493***	-0.259***	-0.417***
	(0.027)	(0.038)	(0.038)	(0.048)
(mean) lgdppcratio		0.012**		0.019
		(0.006)		(0.012)
(mean) lpopratio		0.015***		0.027***
		(0.004)		(0.006)
(mean) lexpratio		0.040***		0.024**
		(0.008)		(0.011)
Constant	0.168***	0.236***	0.138***	0.234***
	(0.013)	(0.022)	(0.021)	(0.031)
R2	0.218	0.292	0.150	0.257
Observations	1170	1170	384	384
Countries	177	177	56	56

Notes: Analogous to Section 4.1 in the paper, the table presents beta convergence regressions when the sample period is restricted to 1983-2014, the last eight four-year cycles. See the text in this Online Appendix for more details.

Table C-6 – Beta-Convergence Regression Results For Each Four-Year Cycle

Dep Var: Δ points	1955-1958	1959-1962	1963-1966	1967-1970	1971-1974	1975-1978	1979-1982	
lagpts	-0.573**	-0.643***	-0.805***	-0.572***	-0.524***	-0.429***	-0.519***	
	(0.208)	(0.105)	(0.087)	(0.065)	(0.081)	(0.093)	(0.073)	
Constant	0.284**	0.319***	0.408***	0.284***	0.258***	0.196***	0.250***	
	(0.127)	(0.056)	(0.050)	(0.037)	(0.045)	(0.044)	(0.034)	
R2	0.214	0.537	0.582	0.431	0.343	0.248	0.428	
Observations	29	39	50	74	91	92	99	
Countries	29	39	50	74	91	92	99	
	1983-1986	1987-1990	1991-1994	1995-1998	1999-2002	2003-2006	2007-2010	2011-2014
lagpts	-0.309***	-0.336***	-0.436***	-0.369***	-0.395***	-0.326***	-0.297***	-0.339***
	(0.076)	(0.068)	(0.061)	(0.051)	(0.073)	(0.067)	(0.049)	(0.060)
Constant	0.153***	0.160***	0.217***	0.178***	0.190***	0.150***	0.138***	0.160***
	(0.040)	(0.036)	(0.033)	(0.025)	(0.035)	(0.034)	(0.024)	(0.031)
R2	0.134	0.214	0.335	0.261	0.238	0.190	0.172	0.191
Observations	105	110	119	155	170	169	170	172
Countries	105	110	119	155	170	169	170	172

Notes: The table presents the unconditional beta regression results for each four-year cycle separately.

D Additional Sigma Convergence Results

In the paper we conduct the test for σ-convergence for longer time horizons, as is standard in the literature (Carree and Klomp, 1997). Here we repeat it for shorter time horizons, namely within each four-year cycle. D-1 presents the results. Although σ-convergence is less likely to materialize at shorter horizons, the table shows many significant results, in particular in the 1960s and 1980s/1990s. The lack of significance within the latest four-year cycles is mirrored in the flattening of the standard deviation graphs in the paper.

Table D-1 – Ratio Test Statistics for σ-Convergence in Win Percentage and Goal Difference Within 4-year Cycles

Period	N	Win Percentages			Goal Difference		
		$\hat{\beta}$	$\hat{\sigma}_1^2$	R-stat	$\hat{\beta}$	$\hat{\sigma}_1^2$	R-stat
1955-1958	26	-0.4829	0.0393	-0.0473	-0.4169	1.6862	0.8472
1959-1962	29	-0.2975	0.0406	0.1944	-0.3812	1.2688	1.7653**
1963-1966	44	-0.7092	0.0274	1.6663**	-0.6390	0.6218	3.8155***
1967-1970	61	-0.5349	0.0316	0.2524	-0.5413	0.9812	1.6188**
1971-1974	80	-0.4409	0.0279	0.8135	-0.3567	0.9621	0.7557
1975-1978	88	-0.3801	0.0344	-0.2552	-0.2816	1.2167	0.0799
1979-1982	95	-0.4344	0.0268	1.8086**	-0.4320	0.8802	2.7708***
1983-1986	103	-0.2962	0.0310	-0.7965	-0.2788	1.1284	-1.5246
1987-1990	107	-0.2749	0.0307	1.0206	-0.2902	0.8378	5.5443***
1991-1994	111	-0.3816	0.0287	0.3875	-0.1547	1.5515	-4.3673
1995-1998	146	-0.3783	0.0249	1.3643*	-0.4231	0.9565	4.4460***
1999-2002	165	-0.4177	0.0260	0.2007	-0.3885	1.0789	1.3555*
2003-2006	169	-0.3263	0.0246	1.3327*	-0.3764	0.9171	4.5074***
2007-2010	169	-0.3059	0.0233	0.5907	-0.2509	0.8761	0.5310
2011-2014	172	-0.3390	0.0227	0.2989	-0.2992	0.7239	2.0603**

Notes: The table presents the variables and results of the sigma-convergence test by Carree and Klomp (1997) described in the paper, computed for the respective periods. *** p<0.01, ** p<0.05, * p<0.1.

E Histograms and Kernel Densities

We plot the histograms and kernel densities of both win percentage and goal difference for each four-year cycle. The scale is the same for comparison. As Figure E-1 and Figure E-2 show, the histograms mostly seem unimodal. Over time, they become taller and thinner, which is in accordance with our finding on σ-convergence. Note that the number of countries varies. For a complete distributional analysis with balanced samples of countries, see Section 4.3 in the paper.

Figure E-1 – Histograms and Kernel Density Plots: Win Percentage per World Cup Cycle (varying numbers of countries)

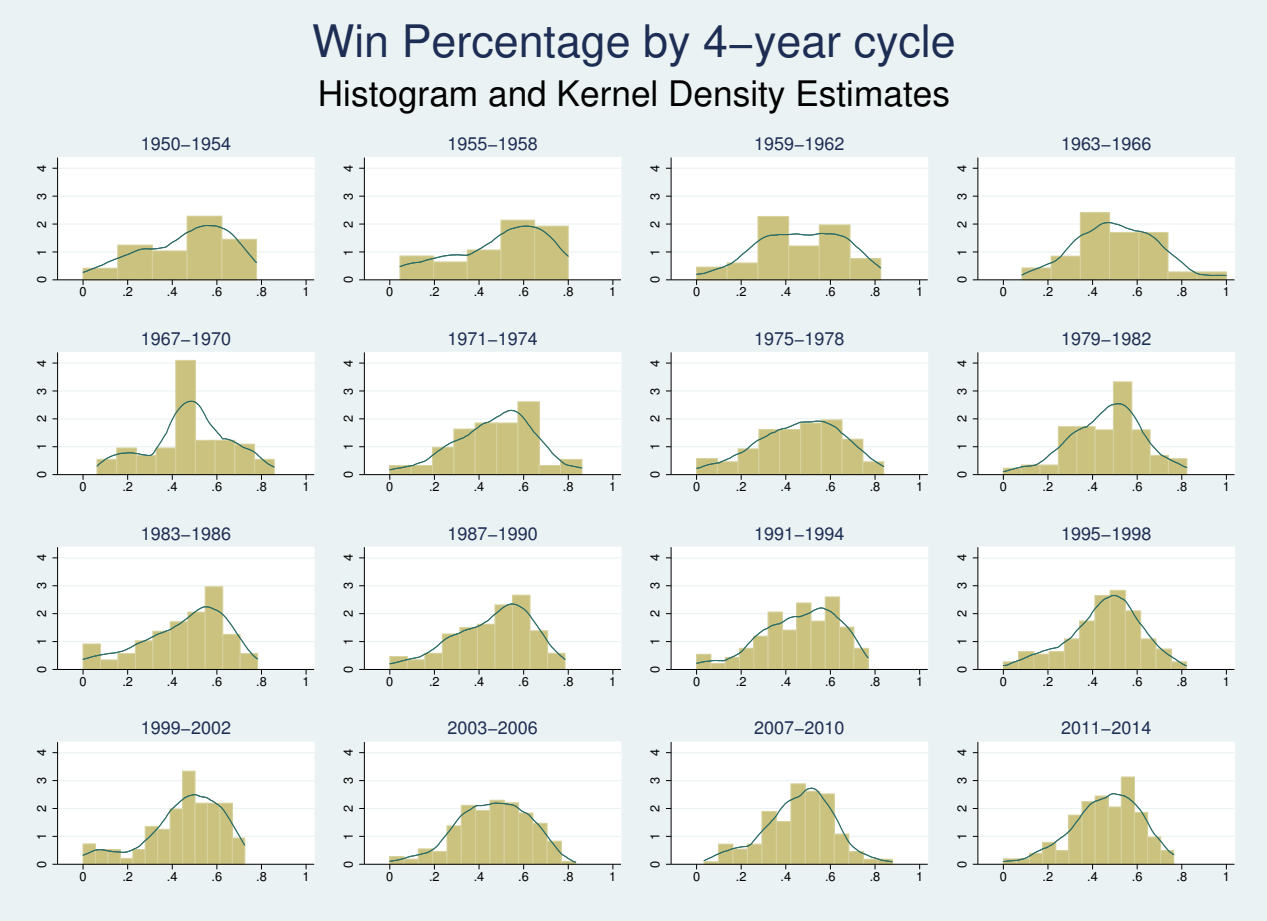
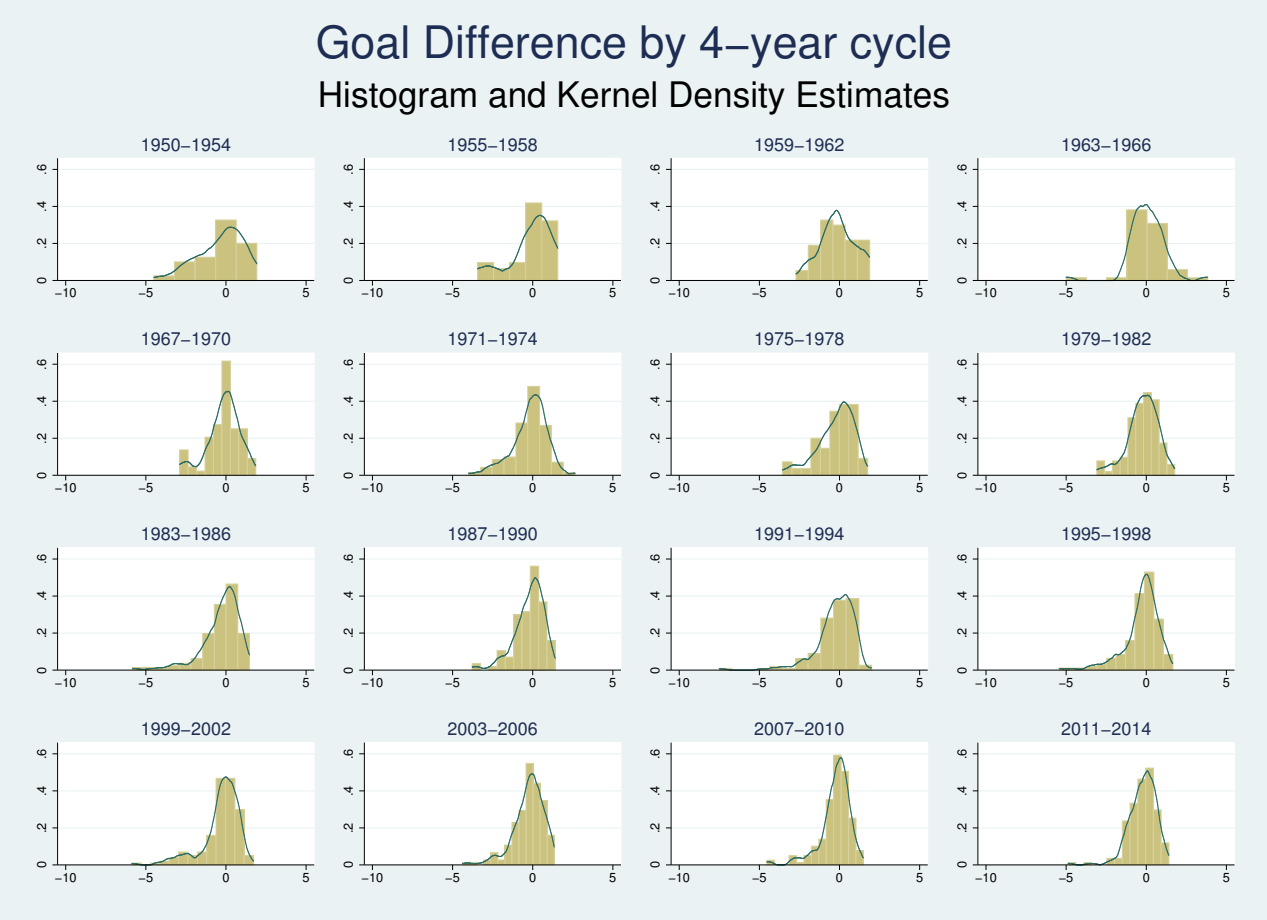


Figure E-2 – Histograms and Kernel Density Plots: Goal Difference per World Cup Cycle (varying numbers of countries)



F Distributional Analysis with Different Samples

Here we repeat the distributional analysis, which the main text conducted with Sample 1 (76 countries and 10 four-year cycles, 1975-2014). We consider the shorter Sample 2 (127 countries and 6 four-year cycles, 1990-2014) as well as an extended Sample 3 (Sample 1 including countries with less than 1m inhabitants, in total 86 countries).

Table F-1 and Table F-2 describe the evolution of the distribution of win percentages and goal differences for both samples according to various characteristics. While Sample 2 behaves very similarly to Sample 1 from the main text in terms of the reduction of standard deviation, skewness and kurtosis, we see that the higher moments remain high for Sample 3. The distribution including tiny countries remains relatively skewed and long-tailed so that the Jarque-Bera null hypothesis of Gaussianity is rejected. This is also visible in the kernel densities Figure F-1. Still, we have observed convergence across all countries, and also within Sample 3, there is a clear decrease in performance inequality in terms of the Gini coefficient (last column of Table F-2). Our conclusion is therefore that very small football nations face significant obstacles due to scarce resources in terms of population and wealth. This effect is, however, not strong enough to affect the overall result of worldwide convergence in performance.

Table F-1 – Distribution of Points and Goal Difference Sample 2 (127 countries)

Panel a) Distribution of Win Percentage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	St.Dev.	Skew	Kurt	JB pvalue	Unimod pvalue	CC Ind.	Pola	Gini
1991-94	0.4752	0.1668	-0.5967	2.9474	0.0280	0.1433	0.3313	0.1482	0.1963
1995-98	0.4858	0.1480	-0.4899	3.3679	0.0460	0.9567	0.1948	0.1071	0.1686
1999-02	0.4986	0.1356	-0.7987	3.5606	0.0062	0.3633	0.3473	0.1110	0.1498
2003-06	0.4959	0.1394	-0.2458	2.2911	0.0941	0.3667	0.3403	0.1328	0.1602
2007-10	0.5007	0.1310	0.0276	3.2144	0.5000	0.5067	0.2732	0.1073	0.1459
2011-14	0.5003	0.1301	-0.2388	2.5107	0.2149	0.5300	0.2967	0.1168	0.1474

Panel b) Distribution of Goal Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	St.Dev.	Skew	Kurt	JB pvalue	Unimod pvalue	CC Ind.
1991-94	-0.1545	1.0823	-1.4467	6.0484	0.0010	0.3267	0.3006
1995-98	-0.0451	0.8217	-0.7569	3.8306	0.0057	0.2700	0.3563
1999-02	0.0427	0.7578	-1.0645	5.1369	0.0010	0.4567	0.2709
2003-06	-0.0177	0.7381	-0.5354	3.2609	0.0379	0.8633	0.2246
2007-10	0.0188	0.6426	-0.5112	3.6708	0.0255	0.7667	0.2219
2011-14	0.0020	0.6497	-0.1382	2.3739	0.2141	0.4900	0.2899

Notes: The analysis is based on a balanced sample of 127 countries (Sample 2) with more than 1m inhabitants throughout the sample period. Columns 1-4 report the distributional moments mean, standard deviation, skewness and kurtosis. Column 5 contains the p-values of the Jarque Bera test with the null hypothesis being the Gaussian distribution. Column 6 shows the p-values of Silverman's (1981) multimodality test with the null hypothesis being a unimodal distribution. Column 7 present the club convergence indicator by Krause (2017), Column 8 the bi-polarization index by Wolfson (1994) and Column 9 the Gini coefficient as a measure of inequality. Due to the presence of negative values in the goal differences, Wolfson's (1994) bi-polarization index and the Gini coefficient cannot be computed for this data.

Table F-2 – Distribution of Points and Goal Difference Sample 3 (86 countries, including those with less than 1m inhabitants)

Panel a) Distribution of Win Percentage

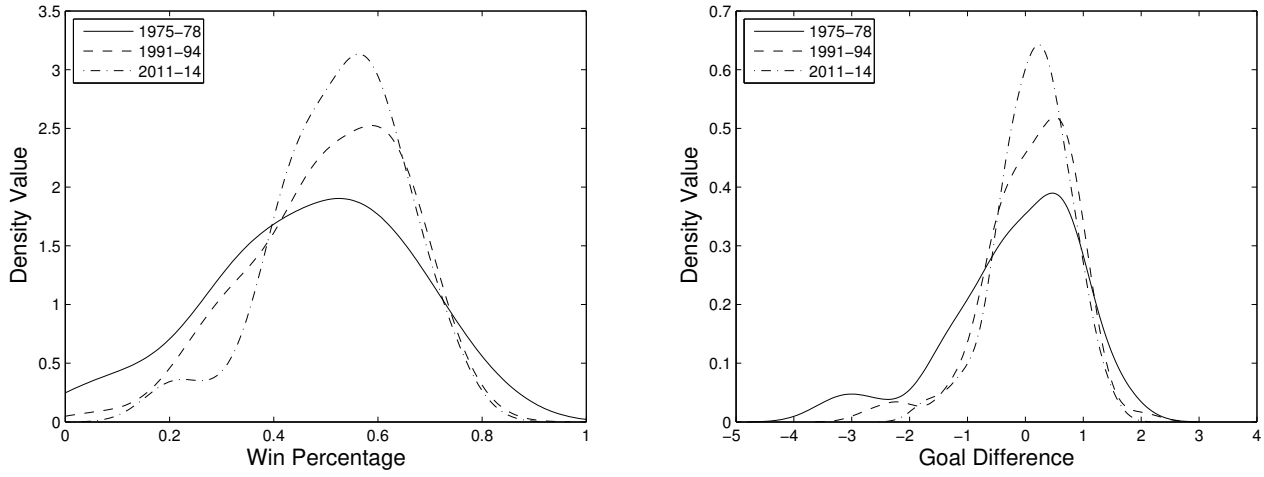
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	St.Dev.	Skew	Kurt	JB pvalue	Unimod pvalue	CC Ind.	Pola	Gini
1975-78	0.4690	0.1888	-0.3993	2.7193	0.1802	0.7333	0.2848	0.1807	0.2260
1979-82	0.4856	0.1573	-0.4490	3.4461	0.0988	0.6300	0.2738	0.1263	0.1781
1983-86	0.5045	0.1537	-0.9286	3.6328	0.0082	0.8933	0.2383	0.1183	0.1651
1987-90	0.4970	0.1582	-0.6722	3.0560	0.0359	0.4567	0.3186	0.1321	0.1757
1991-94	0.5074	0.1443	-0.6202	2.9253	0.0473	0.3700	0.3535	0.1385	0.1584
1995-98	0.5159	0.1341	-0.4774	3.2503	0.1045	0.9733	0.2086	0.1059	0.1437
1999-02	0.5292	0.1160	-0.8369	4.7137	0.0033	0.1967	0.3952	0.1061	0.1199
2003-06	0.5253	0.1360	-0.7245	3.7247	0.0182	0.1300	0.4101	0.1232	0.1431
2007-10	0.5222	0.1356	-0.3753	3.7377	0.0839	0.9900	0.1797	0.1048	0.1422
2011-14	0.5272	0.1232	-0.5948	3.4196	0.0450	0.4667	0.3361	0.1020	0.1289

Panel b) Distribution of Goal Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	St.Dev.	Skew	Kurt	JB pvalue	Unimod pvalue	CC Ind.
1975-78	-0.1622	1.1141	-0.9871	3.8775	0.0053	0.2033	0.4277
1979-82	-0.0947	0.9184	-0.7554	3.9015	0.0130	0.3000	0.3348
1983-86	0.0685	0.8227	-1.0952	4.9705	0.0011	0.6333	0.2617
1987-90	-0.0241	0.7702	-0.8152	3.5144	0.0147	0.4600	0.3250
1991-94	0.1020	0.7827	-0.9667	4.7886	0.0020	0.7833	0.2518
1995-98	0.1257	0.7142	-0.5727	3.7947	0.0317	0.9633	0.1934
1999-02	0.1973	0.6291	-0.7496	4.9710	0.0028	0.2033	0.3495
2003-06	0.1614	0.6953	-1.0305	5.2054	0.0010	0.8167	0.2348
2007-10	0.1124	0.6660	-0.7816	4.0398	0.0099	0.4967	0.2991
2011-14	0.1359	0.6138	-0.5612	3.6389	0.0415	0.5800	0.2843

Notes: The analysis is based on a balanced sample of 86 countries (Sample 3), which, in contrast to Sample 1 includes those with less than 1m inhabitants. See Table F-1 for more details.

Figure F-1 – Densities of Win Percentage and Goal Differences in Various Years, Sample 3 (86 Countries)



G Additional Results on Countries' Performance Evolution

This section provides additional results on the performance evolution of some countries and continents over time, complementing the analysis in the main text.

Table G-1 shows the Theil-Index of inequality in win percentages within the continental confederation in each time period. We note a strong decrease of performance inequality within nearly all continents. For example, within Europe performance inequality decreased by 75% between 1979 and 2014.

Table G-2 shows the correlation of countries' ranks in the performance distribution over time. While countries with a strong performance in one four-year cycle are also likely to do well next period, the correlation of 0.4-0.6 is not as strong as for measures of economic welfare, such as GDP per capita. There is more mobility in the football performance distribution.

Table G-1 – Theil-Index of Inequality in Win Percentage Within Continental Confederations, Sample 1 (76 countries)

	Asia	Africa	America (N,C)	America (South)	Europe
1975-1978	0.1430	0.0439	0.0081	0.0764	0.0358
1979-1982	0.0805	0.0259	0.0140	0.0423	0.0233
1983-1986	0.0437	0.0155	0.0290	0.0683	0.0311
1987-1990	0.0630	0.0160	0.0764	0.0809	0.0310
1991-1994	0.0509	0.0233	0.0122	0.0540	0.0254
1995-1998	0.0249	0.0199	0.0180	0.0459	0.0207
1999-2002	0.0121	0.0127	0.0104	0.0334	0.0114
2003-2006	0.0165	0.0218	0.0431	0.0216	0.0135
2007-2010	0.0206	0.0237	0.0123	0.0301	0.0217
2011-2014	0.0139	0.0175	0.0137	0.0334	0.0097

Notes: In this sample Oceania only consists of one country (New Zealand), so that within-continental inequality in performance is zero.

Table G-2 – Correlation of Countries' Ranks in the Win Percentage Distribution over Four-Year Cycles, Sample 1 (76 countries)

Variables	1975-78	1979-82	1983-86	1987-90	1991-94	1995-98	1999-02	2003-06	2007-10	2011-14
1975-78	1.00									
1979-82	0.54	1.00								
1983-86	0.54	0.51	1.00							
1987-90	0.50	0.36	0.61	1.00						
1991-94	0.39	0.27	0.47	0.62	1.00					
1995-98	0.53	0.36	0.53	0.43	0.61	1.00				
1999-02	0.43	0.22	0.39	0.46	0.57	0.57	1.00			
2003-06	0.52	0.33	0.52	0.57	0.60	0.57	0.73	1.00		
2007-10	0.41	0.17	0.46	0.45	0.57	0.53	0.70	0.73	1.00	
2011-14	0.48	0.37	0.48	0.59	0.63	0.58	0.55	0.68	0.65	1.00

Additional references for Online Appendix

Carree, M. and L. Klomp 1997. Testing the Convergence Hypothesis: A Comment. *Review of Economics and Statistics 79*(4), 683–686. doi: 10.1162/003465397557114.

Krause, M. 2017. The Millennium Peak in Club Convergence - A New Look at Distributional Changes in the Wealth of Nations. *Journal of Applied Econometrics 32*(3), 621–642. doi: 10.1002/jae.2542.

Leeds, M. and E. Leeds 2009. International Soccer Success and National Institutions. *Journal of Sports Economics 10*(4), 369–390. doi: 10.1177/1527002508329864.

Silverman, B. 1981. Using Kernel Density Estimation to Investigate Multimodality. *Journal of the Royal Statistical Society, Series B (Methodological) 43*(1), 97–99.

Wolfson, M. 1994. When Inequalities Diverge. *American Economic Review, Papers and Proceedings 84*(2), 401–421.