Divergence in Beauty: What Contributed to the Great Divergence of the Beautiful Game?

Guillermo Parra

University of British Columbia (UBC)
Vancouver School of Economics (VSE)

April 28, 2024

Abstract

This study explores the influence of institutional and cultural factors on the international success of soccer teams, motivated by the disparity in performance among countries with similar cultural interests in soccer. By constructing and analyzing two specialized indexes—the Institutional Index and the Cultural Index—this research assesses their effects on the All-Time Rank of national soccer teams. The findings demonstrate that institutional and cultural factors significantly impact a nation's soccer performance. Specifically, an increase in the Institutional Index is associated with an improvement of 1.8 ranks, and the Cultural Index is 1.1.

1 Introduction

1.1 Research Motivation

The motivation for this research comes from an observation of international soccer success among countries with similar cultural affinities for the sport. The example that came to mind was between Colombia and Argentina. Despite its larger population, Colombia does not have the same soccer success as Argentina, as evidenced by its 1 Copa America and 0 World Cups to Argentina's 9 and 3. This discrepancy raises intriguing questions about the factors contributing to national soccer teams' success. We can also see the large difference in their "All Time Rank," which reflects the average historical performance of each country. The study is driven by a desire to uncover why some nations have always been great at soccer while others, sharing seemingly similar conditions and enthusiasm for the sport, have never been. (3). We can see the stark differences in success over time in figure 1.

1.2 Objective of the Study

The primary objective of this study is to explore the impact of institutional and cultural factors on the international success of soccer teams. This is achieved by constructing two specialized indexes:

- The **Institutional Index** quantifies structural support for soccer across different nations by evaluating aspects such as the number of coaches and participation in youth World Cups.

- The **Cultural Index** measures the cultural environment's impact on soccer by assessing factors such as the sport's popularity and professional player density alongside an evaluation of player traits.

1

These indexes are integrated into multiple regression models to analyze their effects on the All-Time Rank of national soccer teams. The methodological approach includes:

1. Normalizing various metrics using the Min-Max Scaler to ensure comparability across different scales.

2. Calculating weighted averages of normalized variables to construct the Institutional and Cultural Indexes.

3. Using Ordinary Least Squares (OLS) regressions to estimate the influence of these indexes on soccer success while controlling for economic factors like GDP per capita.

My results indicate that cultural and institutional factors significantly influence a country's soccer success.

1.3 Significance of the Study

This research aims to provide empirical evidence on the importance of cultural and institutional dynamics in sports success. By understanding these factors, policymakers and sports administrators can better foster environments that promote the development and success of soccer—or any sport—within their respective countries.

2 Existing Literature

2.1 Institutional Literature

Batarfi and Reade (2021) examine various institutional, demographic, and macroeconomic factors influencing how well countries perform in soccer. They found that a higher GDP per capita generally benefits national soccer success, suggesting that economic prosperity allows for better sporting infrastructure and resources. Conversely, a larger population tends to dilute this positive effect, possibly due to increased resource competition. The authors also highlight the importance of a developed domestic soccer structure, which correlates with better national team performance. (4).

Hoffmann, Ging, and Ramasamy (2002) study the socio-economic determinants of international soccer performance, analyzing how economic conditions, cultural influences, and demographic traits impact national teams' success. They discover an inverted U-shaped relationship between wealth and soccer success (10).

Pancotto, Addessi, and Auteri (2024) examine soccer player registration prices, hypothesizing the presence of a speculative bubble in this market. They examine the dramatic rise in player prices since 2012, using structural break tests to identify shifts in pricing trends potentially linked to regulatory changes like the Fair Play regulations. Their findings confirm significant changes in the market dynamics of soccer player transfers, suggesting these were influenced by institutional efforts to regulate financial practices within sports organizations. This study is relevant as it highlights the influence of institutional regulations on the economic aspects of soccer (15).

Daniel Weimar (2019) explores the economics of professional soccer, analyzing how various economic and regulatory aspects influence the sport at different levels. His work categorizes the existing literature into six main streams of soccer economics research, covering topics from team productivity and soccer finance to the labour market within the sport (23).

Boto-García, Varela-Quintana, and Muñiz (2023) investigate the effects of team composition, specifically the number of foreign players, on the technical efficiency of European soccer teams. Their study uses a stochastic frontier analysis to explore how team diversity impacts performance. The researchers found that while diversity might bring potential skill improvements, it often leads to increased inefficiencies due to communication costs. It illustrates the complexity of incorporating different languages into teams like in Africa and highlights the institutional challenge of managing communication (5).

Gregory T. Papaniko's (2015) study examines the economic, population, and political determinants of match outcomes in the 2014 FIFA World Cup. By analyzing goal differences and considering the relative strength of teams, Papanikos finds that economic factors like GDP per capita, population size, and political conditions significantly influence the performance of national teams. The research shows that wealthier nations with larger populations and favourable political conditions tend to perform better(16).

Leeds and Leeds (2007) analyze the factors influencing international soccer success, focusing on the roles of political regimes, colonial heritage, and national institutions. They employ a mix of OLS and negative binomial regressions to explore how these factors affect FIFA points and rankings. The study finds that political regime types and past colonial relationships influence soccer performance, with institutional factors like domestic soccer infrastructure playing a significant role in shaping national teams' success. (13).

2.2 Cultural Literature

Depetris-Chauvin, Durante, and Campante explore the nation-building effects of soccer in sub-Saharan Africa, demonstrating how national soccer team successes can significantly decrease ethnic identity and inter-ethnic violence while increasing trust across different ethnicities. These effects are particularly big when the national team's victory is unexpected or against a historical rival. This study underlines the profound cultural and social impact that soccer can have beyond just the playing field, highlighting how sports successes can foster national unity and potentially ease ethnic tensions (6).

The paper by Juan Pedro Ronconi investigates the relationship between soccer rivalries and social cohesion in Latin America, mainly focusing on how these rivalries can enhance community bonds despite their inherently divisive nature. Using quasi-experimental methods, Ronconi finds that social cohesion often improves following soccer matches unless marred by violence or unethical behaviour from players. This effect extends beyond just soccer fans, suggesting a broader societal impact (18).

Kikuta and Ono (2024) explore the global impact of international soccer games on political leader approval in over 100 countries. Their study identifies a significant effect where unexpected victories in soccer enhance national leaders' approval by 7 percentage points. This phenomenon is attributed to the perception of soccer games as benchmarks for leader performance, despite the non-political nature of the events. This paper empirically demonstrates the cultural impact of soccer on national sentiment and leader perception across a diverse global sample (11).

Macmillan and Smith (2007) address the statistical challenges in other studies on international soccer rankings, notably issues with sample selection bias and non-normal errors. They expanded their dataset to include 100 additional countries and introduced new variables like population size and historical soccer involvement. Utilizing both FIFA and alternative ranking systems, they developed findings about the impact of long-standing soccer traditions and broader national engagement, highlighting the deep-seated cultural roots that could enhance or

impede national soccer performance. (14).

Fischer and Haucap (2021) explore the impact of crowd support on home advantage in professional soccer by analyzing matches played without spectators during the COVID-19 pandemic, termed "ghost games." Their findings indicate a diminished home advantage in the first division of German soccer, suggesting that crowd presence significantly contributes to the psychological and competitive edge usually enjoyed by home teams. This research provides a perspective on the cultural aspects influencing soccer performance, mainly through the lens of spectator support. By demonstrating the tangible effects of crowd absence on game outcomes, Fischer and Haucap's findings reinforce the component of my research that examines cultural significance in soccer (9).

Rossing and Skrubbeltrang explore how different actions in soccer games relate to national culture, suggesting that these actions form distinct "dialects" within the broader "language" of soccer. Their analysis focuses on the Brazilian and Italian soccer cultures, using document and video analysis of World Cup matches from 2010 and 2014. They identify that while there is some alignment between national soccer cultures and the behaviour of national teams, there are also significant discrepancies. This aligns with my decision to create a cultural index based partly on a country's grouped leadership teamwork and other ratings from the Football Manager data set (19).

3 Historical Background

3.1 Early Adoption and Spread

The history of soccer, or Football as it is known outside North America, can be traced back to various forms of folk soccer played in medieval Britain. However, the modern version of soccer began to take shape in the 19th century when it was formalized in England. In 1863, the establishment of the Football Association marked the standardization of the game's rules, making organized competitions and league play possible. This formalization facilitated the spread of the game internationally. By the late 19th century, the British had introduced the sport to many places, setting up clubs in Brazil, Uruguay, Italy, and Germany, among others. Each region adapted the sport to its local contexts, which led to regional diversification of playing styles (12).

3.2 Cultural Significance and National Identity

Soccer has grown to represent far more than a game. It has become a symbol of national identity and pride. In Brazil, the game is intertwined with national identity, where the expressive style of play mirrors the country's famed carnival culture, emphasizing beauty, rhythm, and flair. In European countries like Spain and Italy, soccer clubs like Real Madrid, Barcelona, and Juventus embody regional identities and become arenas where political and cultural tensions play out. Similarly, international tournaments such as the World Cup have seen moments where soccer becomes a stage for historical and political significance. For example, the infamous 1969 "soccer War" between Honduras and El Salvador showed how deeply the sport could influence national sentiments and international relations. (20).

3.3 Economic Transformation

The transformation of soccer into a global economic powerhouse represents one of the most significant aspects of its history. The sport's ability to generate substantial economic activity has been recognized by countries and corporations alike. With its global broadcasting rights, the

English Premier League attracts viewership in the hundreds of millions, generating significant revenue through advertising, broadcasting rights, and merchandise. This economic aspect is not just confined to Europe; in Asia, leagues in China, Japan, and South Korea have invested heavily in developing their domestic leagues and attracting international talent. The economic impact of major tournaments like the FIFA World Cup is significant, often spurring infrastructural development, increasing tourism, and boosting local businesses in host countries. The 2014 World Cup in Brazil, for example, contributed an estimated $30 billion to the Brazilian economy (17).

4 Data

4.1 Data Description

This section provides detailed descriptions of the datasets utilized in this study.

4.1.1 FIFA U-17 World Cup Performance Data

This dataset has the performance of national teams that participated in the FIFA U-17 World Cup from 1985 to 2023. It includes data on 368 teams, detailing goals scored, conceded, tournament points, and overall tournament results. The data was collected by scraping Wikipedia pages, which sourced their information from the official FIFA website (1). This data can be summarized in figure 6.

4.1.2 FIFA U-20 World Cup Performance Data

This dataset covers national teams' FIFA U-20 World Cup performances from 1977 to 2023. It consists of 424 observations and features similar variables as the U-17 dataset. The method of data collection also involved web scraping Wikipedia (2). This data can be summarized in figure 7.

4.1.3 Sport Popularity by Country

This dataset includes the most popular sports in each country from a Babbel survey. It compiles information from the early 2000s and more recent data from 2021 to track changes in sports popularity over time. A dummy variable was created to determine whether soccer is a favourite sport in my master aggregated dataset (7).

4.1.4 World Bank Economic Indicators

This dataset provides economic data from the World Bank, featuring the latest observations on each country's population and GDP per capita. The data is crucial for my analysis since a country's economic conditions are likely relevant to its performance in international soccer competitions. The data is updated annually, and the version used in this study is from 2022 (World Bank). The relationship between this data and all-time rank can be summarized in figure 8.

4.1.5 FIFA Game Coaches Database

This dataset is from the latest FIFA video game, with a list of the top 1000 soccer coaches based on a points-per-game system and total games managed. It includes over 1,000 entries detailing each coach's name, country of origin, and date of birth. This will allow us to see how many coaches in the top ˜1000 are from specific countries. The data was extracted from the FIFA

video game's database through web scraping (Electronic Arts). This data can be summarized in figure 5.

4.1.6 Football Manager Soccer Players Database

This dataset includes detailed attributes and metrics for nearly 160,000 professional soccer players scraped from the Football Manager video game (Sports Interactive). This dataset was obtained through web scraping. The level of detail in player attributes makes this dataset particularly valuable for analysis. This data will end up being aggregated at the country level; it includes:

- **Physical Attributes:** Height, Weight, Natural Fitness, Stamina, Strength, Pace, Jumping, Balance.

- **Technical Skills:** First Touch, Dribbling, Finishing, Heading, Long Shots, Freekicks, Passing, Crossing, Tackling, Technique, Penalty Taking.

- **Goalkeeping Skills:** Aerial Ability, Command of Area, Handling, Kicking, One on Ones, Reflexes, Rushing Out, Tendency to Punch, Throwing.

- **Mental Attributes:** Aggression, Anticipation, Bravery, Composure, Concentration, Vision, Decisions, Determination, Flair, Leadership, Off The Ball, Positioning, Teamwork, Workrate.

- **Personality Traits:** Consistency, Dirtiness, Important Matches Performance, Injury Proneness, Versatility, Adaptability, Ambition, Loyalty, Pressure Handling, Professionalism, Sportsmanship, Temperament, Controversy.

- **Positional Roles:** Goalkeeper, Sweeper, Striker, Attacking Midfielder (Central, Left, Right), Defender (Central, Left, Right), Defensive Midfielder, Midfielder (Central, Left, Right), Wing Back (Left, Right).

4.1.7 FIFA Rankings and All-Time Rank Calculation

This dataset has historical FIFA rankings of national soccer teams from their inception in 1992 through April 2023. The rankings are updated every 70 days, reflecting the teams' performance in international fixtures. Additionally, an 'All-Time Rank' metric is constructed within this dataset, offering a weighted average of all historical rankings for each country. Calculating the All-Time Rank involves summing the ranks across all available data points for a country and dividing by the number of observations. This method provides an overview of a country's soccer success over time, showing how their performance trends change. The dataset was also obtained by scraping FIFA's official website. This data can be summarized in figure 2.

4.1.8 Aggregated Soccer Data by Country

This dataset aggregates individual player statistics from the Football Manager player dataset, which I grouped by country. It provides an overview of the number of professional soccer players each country has produced and their average age, height, and the other averaged metrics in the dataset. This dataset is replicated but only includes players with at least 10 international appearances (caps). The number of pro players a country produces can be seen in figure 3.

4.1.9   Master Aggregated Soccer Data by Country

The Master Aggregated Dataset integrates various datasets, including FIFA rankings, population statistics, GDP per capita, and participation in youth World Cups, to provide a different view of each country's soccer landscape. This dataset includes variables like the logarithm of professionals per population, all-time FIFA ranking, and most popular sport soccer dummy. This dataset will be used for my analyses to examine the relationship between a country's culture and institutions and soccer success. This data set also includes two columns, the cultural and institutional indexes, whose calculations will be presented below.

4.2   Data Transformations

This subsection details the specific transformations applied to the data to prepare it for analysis.

4.2.1   Aggregation of Data

To analyze the performance trends over time, I calculate the cumulative moving average of the ranks for each country. This method provides what I refer to as the "All Time Rank," which reflects the average historical performance of each country based on available data. From here, I took just the latest observations to give us the latest overall rank. I merged them with the data set containing all the information about individual countries' players' qualities.

4.2.2   Log Transformation of Population Proportionality

The proportionality between the number of professional soccer players and the population in each country is calculated using a logarithmic transformation. The formula applied is:

$$\text{Proportion} = \frac{\log(\text{Number of Professional Soccer Players})}{\log(\text{Population})}$$

This transformation is needed for normalizing the data, particularly to address the skewness caused by large populations. It ensures that the data scales uniformly, allowing for a more accurate and fair comparison across different countries. I noticed that countries such as Brazil had the most professional soccer players in South America, but it also had the biggest population. Hence, their pro over population numbers were very small. We now have a more normal distribution in this data, which can be summarized in figure 4.

4.3   Creation of the Cultural Index

The Cultural Index quantifies the cultural environment towards soccer in each country. This index uses normalized data and a combination of positively and negatively weighted player attributes to reflect the socio-cultural dynamics that impact soccer success.
Initial variables are normalized using the Min-Max Scaler from the `sklearn.preprocessing` package, ensuring a uniform scale from 0 to 1 for comparability. The initial variables considered are:

- Soccer popularity dummy, reflecting the relative popularity of soccer in the country (normalized as `norm_soccer_dummy`).

- Logarithm of professionals per population, indicating the density of professional soccer players relative to the country's population (normalized as `norm_pro_per_pop`).

Following the normalization of these variables, player traits are averaged at the country level, using the 10 cap minimum data as I want only to include players that affect the success of a country. These are also normalized. A list of these includes Communication, Eccentricity,

Bravery, Aggression, Composure, Concentration, Decisions, Determination, Flair, Leadership, Teamwork, Work rate, Consistency, Dirtiness, Important Matches, Versatility, Adaptability, Ambition, Loyalty, Pressure, Professionalism, Sportsmanship, Temperament, and Controversy.

Specific traits such as aggression, dirtiness, temperance, controversy, and eccentricity are assigned negative weights due to their potential negative impact on team dynamics and player conduct. These negatively weighted traits are multiplied by -1 in the calculation.

The Cultural Index is computed using a formula that integrates the normalized values of soccer popularity, professional density, and an aggregate score of player traits. The formula for the Cultural Index is as follows:

$$\begin{aligned}
\text{Cultural Index} = \ & 0.40 \times \texttt{norm\_soccer\_dummy} \\
& + 0.24 \times \texttt{norm\_pro\_per\_pop} \\
& + 0.015 \times (\text{average of all weighted player traits})
\end{aligned}$$

This approach allows for a measurement that considers the positive and negative aspects of soccer-related cultural traits. The calculation ensures that the Cultural Index reflects player characteristics at the country level.

4.4 Creation of the Institutional Index

The Institutional Index calcifies the structural support provided to soccer in different countries. This index is constructed by normalizing and weighting several variables indicative of institutional support for soccer. The normalization process uses the Min-Max Scaler from the `sklearn.preprocessing` package to transform the variables into a range between 0 and 1, ensuring that each variable contributes equally to the index.
The variables included are:

- Number of coaches (normalized as `norm_coaches`)

- U20 World Cup participation (normalized as `norm_U20`)

- U17 World Cup participation (normalized as `norm_U17`)

The weighted average of these normalized variables is calculated with weights of 0.3, 0.35, and 0.35, respectively. The formula used is:

$$\text{Institutional Index} = 0.3 \times \text{norm\_coaches} + 0.35 \times \text{norm\_U20} + 0.35 \times \text{norm\_U17}$$

Finally, these calculated indexes are multiplied by 100 and saved into the dataset, serving as my variables of interest for analyzing the cultural and institutional impact on soccer's popularity and success in different countries.

5 Empirical Approach

The analysis uses multiple ordinary least squares (OLS) regressions to assess the direct effects of various predictors on soccer success.

5.1 Model Specification

5.1.1 Baseline Institutional Model

The baseline model investigates the relationship between the Institutional Index and the Rank of national soccer teams. The model is specified as:

$$\text{Percentage Rank} = \beta_0 + \beta_1 \text{Institutional Index} + u$$

5.1.2 Baseline Cultural Model

Similarly, another baseline model assesses the influence of the Cultural Index on the Percentage Rank. This model is formulated as:

$$\text{Percentage Rank} = \beta_0 + \beta_1 \text{Cultural Index} + u$$

5.1.3 Institutional Model with Economic Control

To account for economic influences that might confound the relationship between institutional support and soccer success, GDP per capita is included as a control variable in an extended model. This adjustment allows for a more precise interpretation of the institutional effect, isolating it from the country's economic conditions. The model is extended as follows:

$$\text{Percentage Rank} = \beta_0 + \beta_1 \text{Institutional Index} + \beta_2 \text{GDP per capita} + u$$

5.1.4 Cultural Model with Economic Control

Similarly, GDP per capita is also considered in the cultural model. This inclusion helps to mitigate any overlapping effects between economic conditions and the cultural environment. However, caution is advised due to potential multicollinearity, especially since the Cultural Index incorporates the natural logarithm of population and GDP per capita has population as its denominator. The extended cultural model is:

$$\text{Percentage Rank} = \beta_0 + \beta_1 \text{Cultural Index} + \beta_2 \text{GDP per capita} + u$$

5.2 Combined Institutional and Cultural Model

The final regression combines the effects of both institutional and cultural indexes on soccer success. The model is specified as:

$$\text{Percentage Rank} = \beta_0 + \beta_1 \text{Cultural Index} + \beta_2 \text{Institutional Index} + \beta_3 \text{GDP per capita} + u$$

6 Results

6.1 Interpretation of Results

My regression analyses provide evidence of the influence of institutional and cultural factors on nations' success in international soccer. The findings suggest that these factors are statistically significant predictors of a nation's FIFA ranking, even after adjusting for GDP per capita.

6.2 Institutional Quality and Soccer Success

The first set of analyses focused on the Institutional Index. Results from the basic OLS regression Table 1 indicate that an increase in the Institutional Index by one unit is associated with an improvement of 2.188 units in the national soccer team's rank, holding other factors constant. This effect remains stable at 2.157 after controlling for GDP per capita Table 2, suggesting that the impact of institutional quality on soccer success is significant and unchanged when controlling for economic conditions.

6.3 Cultural Significance and Soccer Popularity

Turning to the Cultural Index, the regression results without controls in Table 3 show that each unit increase in the Cultural Index leads to an improvement of 1.464 in the ranking. When GDP per capita is included as a control in Table 4, the coefficient slightly adjusts to -1.477, which remains statistically significant. However, the model's presence of GDP per capita introduces potential multicollinearity. Despite this, the significance and magnitude of the Cultural Index suggest a strong cultural influence on soccer performance.

Further analysis combining the Cultural and Institutional Indexes in Table 5 indicates a cumulative and significant impact, with coefficients of -1.098 for cultural and -1.859 for institutional factors. This combined model, even when extended to include GDP per capita Table 6, maintains the significance and strength of these predictors with similar coefficients, indicating that both cultural and institutional dimensions significantly shape the success of national soccer teams across the globe. The model exhibits good explanatory power, with R-squared values of 0.576 and 0.583, respectively, indicating that institutional and cultural factors can explain a considerable portion of the variance in soccer success.

7 Discussion and Conclusions

This research explores the roles that institutional and cultural aspects play in determining the international success of soccer teams. However, despite these advancements, my study has limitations, which give rise to several important avenues for future research.

7.1 Limitations of the Study

1. **Identification Issues**: One of the major limitations of this study is the difficulty in isolating the causal impacts of institutional and cultural variables on soccer success. The study's observational nature limits my ability to draw definitive causal inferences. For instance, the relationship between institutional quality and soccer success could be influenced by unobserved variables, such as internal policy changes or international soccer governance, that are not captured in my current models.

2. **Potential endogeneity of Instrumental Variable**: The initial plan to use players born in one country but played for another as an instrumental variable for institutional quality faced challenges. This variable may not meet the exclusion restriction, as the decision of players to represent different countries can be influenced by the very institutional factors I aim to measure. In my estimation, this endogeneity introduces potential biases.

3. **Limited Data and Data Collection**: The inability to scrape data limited my analysis. Enhanced data collection from sources like the Football Manager databases or Transfermarket could provide more comprehensive data, improving the robustness and precision of my findings. Ideally, I would have been able to web scrape all of Transfermarket.com as a football manager. However, unfortunately, this was far harder than I imagined and far past my abilities (22).

4. **Simplification of Complex Constructs**: The construction of the Cultural and Institutional Indexes, while insightful, grossly oversimplifies complex cultural and institutional dynamics into single composite scores. This simplification can mask how individual components of these indexes influence soccer success.

5. **Timing of Variables**: A notable limitation of this study is the potential temporal misalignment between the measured performance of national soccer teams and the factors

used to predict them. While GDP might logically act as a leading variable, the popularity of soccer could be considered a lagging variable. It is plausible that the socioeconomic conditions from a decade prior, which influence the developmental phase of players, have a more significant impact on current success than present values. Similarly, while popularity and success are closely linked, it appears equally probable that success could drive popularity rather than vice versa. Utilizing lagged GDP values and popularity or examining GDP growth rates might yield more insightful results into the dynamics influencing soccer success.

7.2 Implications for Future Research

7.2.1 Attribute Analysis of Soccer Success

A promising direction for future research would be to look into the specific player attributes that drive soccer success. Given the comprehensive dataset on player traits available from the Football Manager database, it would be possible to conduct an analysis to uncover distinctive patterns. For instance, exploring whether players from top-ranking countries consistently exhibit higher ratings in traits such as Passion and Versatility could provide insights into the characteristics that drive national team success. This approach adds depth to understanding the micro-level attributes that contribute to the macro success of national soccer teams.

7.2.2 Robustness Checks

To strengthen the conclusions drawn from this study, future research should focus on several robustness checks:

1. **Adjusting Index Weights**: The weighting scheme used to create the Cultural and Institutional Indexes was based on initial assumptions about the relative importance of various factors. Future studies could explore alternative weighting methods to assess the sensitivity of the results to these assumptions. This includes using methods to optimize the weights based on robust criteria.

2. **Component-wise Regression Analysis**: An in-depth analysis regressing the All-Time Rank on each individual component of the indexes could reveal which factors are most influential. This disaggregated approach would help understand the specific contributions of different cultural and institutional elements to soccer success.

3. **Different Econometric Techniques**: Employing more sophisticated econometric techniques, such as IV, Fixed Effects or a structural model, could help address some of the identification problems noted. These methods would clarify the causal relationships between cultural and institutional factors and soccer performance.

7.2.3 Expanding Data Collection

Further research should also consider expanding the data collection efforts:

1. **Enhanced Web Scraping**: Developing the ability to scrape and integrate data from databases like Transfermarkt and Football Manager would enable researchers to access a richer dataset. This would include player statistics, coaching data, and other relevant metrics over time.

2. **Longitudinal and Cross-Sectional Expansion**: Incorporating more longitudinal data to capture changes over time and expanding the cross-sectional scope to include more countries could provide a more detailed picture of global soccer.

7.2.4   Exploring New Variables and Instruments

Finally, future studies should explore the potential of new variables and instruments:

1. **Alternative Instruments**: Identifying and testing new instruments that better satisfy the exclusion restriction could strengthen the instrumental variable approach. This might involve exploring socio-political changes, policy shifts, or international soccer regulations as potential instruments.

2. **In-depth Cultural Analysis**: Further qualitative and quantitative research into the specific cultural aspects that influence soccer, such as national identity, historical rivalries, or public interest in sports, could provide insights into how these factors affect soccer success.

Addressing these limitations will help us better understand the cultural, institutional, and economic factors that shape soccer teams' international success.

References

[1] FIFA U-17 World Cup. `https://en.wikipedia.org/wiki/FIFA_U-17_World_Cup`. 2024.

[2] FIFA U-20 World Cup. `https://en.wikipedia.org/wiki/FIFA_U-20_World_Cup`. 2024.

[3] (2024). Fifa world ranking.

[4] Batarfi, M. and Reade, J. J. (2020). Why are we so good at football, and they so bad? institutions and national footballing performance. *De Economist*, 169(1).

[5] Boto-García, D., Varela-Quintana, C., and Muñiz, A. (2023). Foreign players, team production, and technical efficiency: Evidence from european soccer. *Bulletin of Economic Research*, 75(4):1209–1241.

[6] Depetris-Chauvin, E., Durante, R., and Campante, F. (2020). Building nations through shared experiences: Evidence from african football. *American Economic Review*, 110(5):1572–1602.

[7] Devlin, T. M. (2021). A world tour of athletics: The most popular sports by country. *Babbel Magazine*. Accessed: Insert date here.

[Electronic Arts] Electronic Arts. FIFA. `https://www.ea.com/games/fifa`. 2024.

[9] Fischer, K. and Haucap, J. (2021). Does crowd support drive the home advantage in professional football? evidence from german ghost games during the covid-19 pandemic. *Journal of Sports Economics*, 22(8):982–1008.

[10] Hoffmann, R., Lee, C. G., and Ramasamy, B. (2002). The socio-economic determinants of international soccer performance. *Journal of Applied Economics*, 05(2):20.

[11] Kikuta, K. and Ono, Y. (2024). Global evidence for the relevance of irrelevant events: International soccer games and leader approval. SocArXiv.

[12] Kitching, G. (2015). The origins of football: History, ideology and the making of 'the people's game'. *History Workshop Journal*, No. 79:127–153.

[13] Leeds, E. M. and Leeds, M. (2007). International soccer success and national institutions. *Journal of Sports Economics*, 10(4):1–17.

[14] Macmillan, P. and Smith, I. (2007). Explaining international soccer rankings. *Journal of Sports Economics*, 8(2):202–213.

[15] Pancotto, F., Addessi, G., and Auteri, N. (2024). Soccer bubble: Is there a speculative bubble in the price of international soccer players? *Journal of Sports Economics*.

[16] Papanikos, G. (2015). Economic, population and political determinants of the 2014 world cup match results. *Soccer and Society*, 18(4):1–17.

[17] Review, B. E. (Year of publication). The tale of two tournaments: A football story.

[18] Ronconi, J. P. (2022). Divided for good: Football rivalries and social cohesion in latin america. Bravo Center.

[19] Rossing, N. and Skrubbeltrang, L. (2017). The language of football: a cultural analysis on selected world cup nations. *Sport in Society*, 20(5-6):599–611.

[20] Rouquié, A. and Vale, M. (1973). Honduras–el salvador, the war of one hundred hours: A case of regional "disintegration". *International Journal of Politics*, 3(3):17–51.

[Sports Interactive] Sports Interactive. Football Manager. `https://www.footballmanager.com/`. 2024.

[22] Transfermarkt GmbH & Co. KG (2024). Transfermarkt. `https://www.transfermarkt.com/`. Accessed: April 28, 2024.

[23] Weimar, D. (2019). The economics of professional soccer. In *Handbook of Sports Economics*. Sage Publications.

[World Bank] World Bank. Population, total. `https://data.worldbank.org/indicator/SP.POP.TOTL?locations=1W`. 2024.
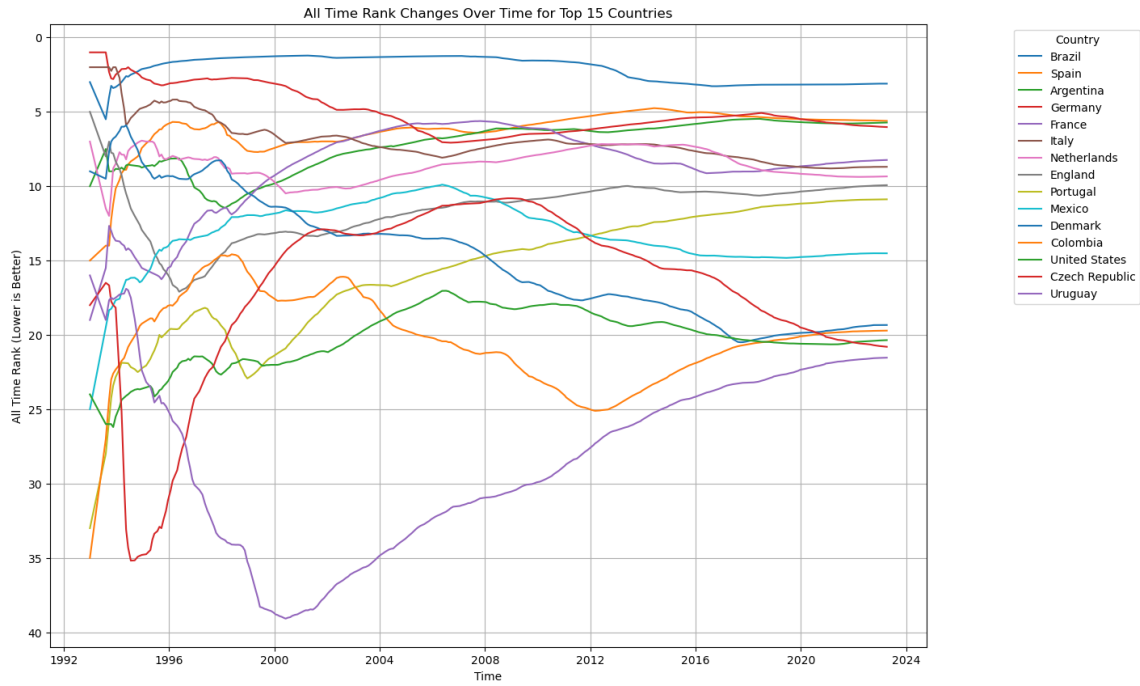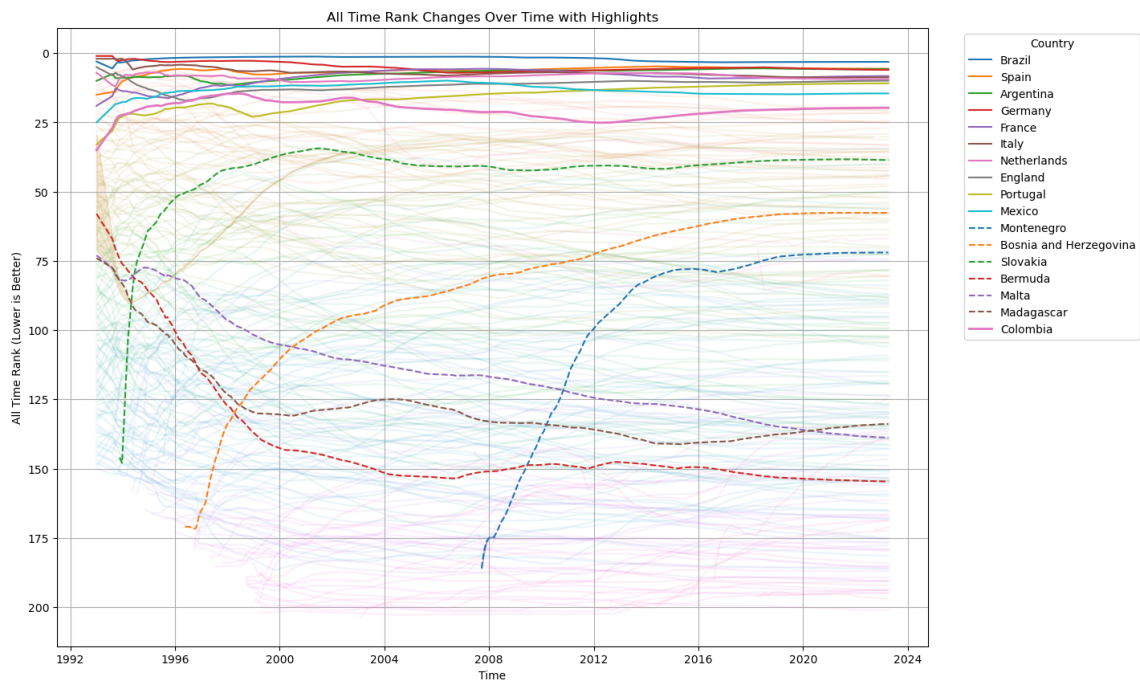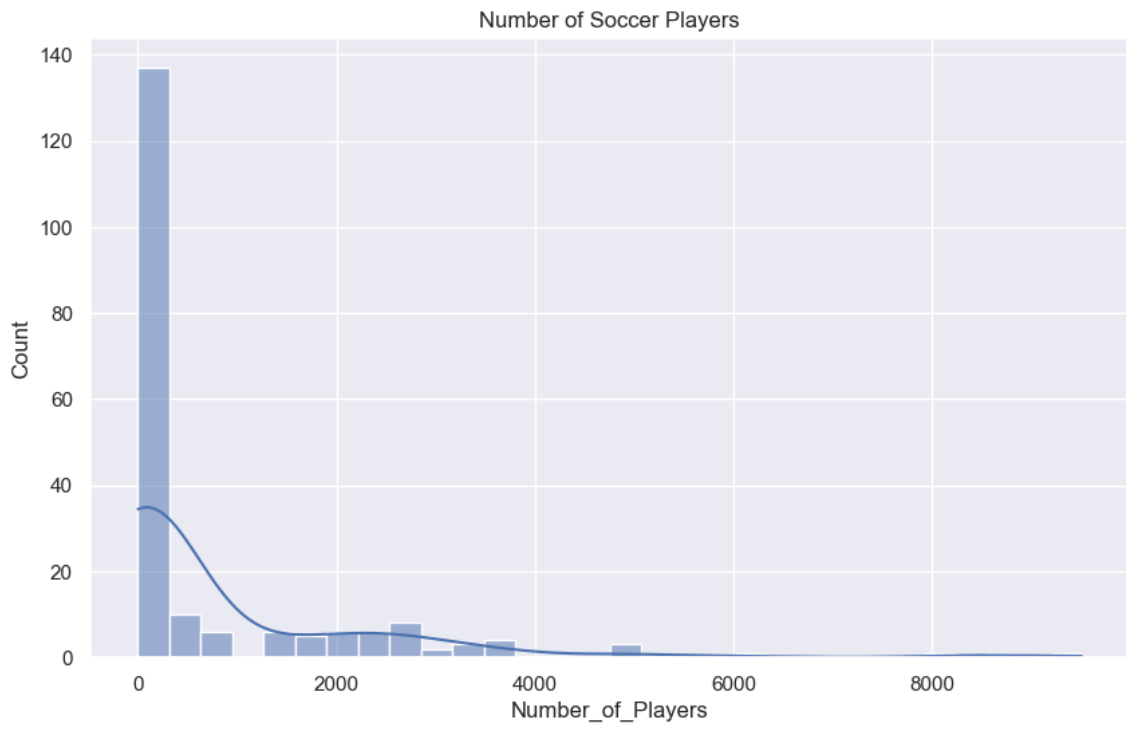
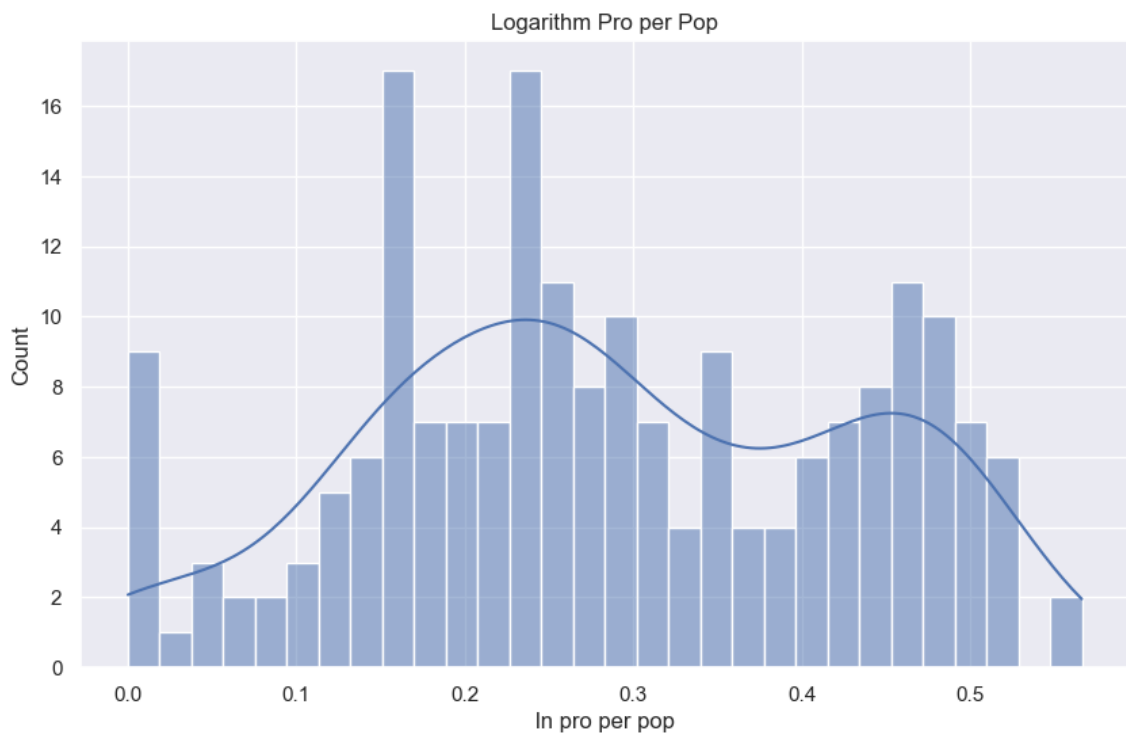## Acknowledgments

7.3 Tables and Figures

Figure 1

Figure 2

Figure 3

Figure 4
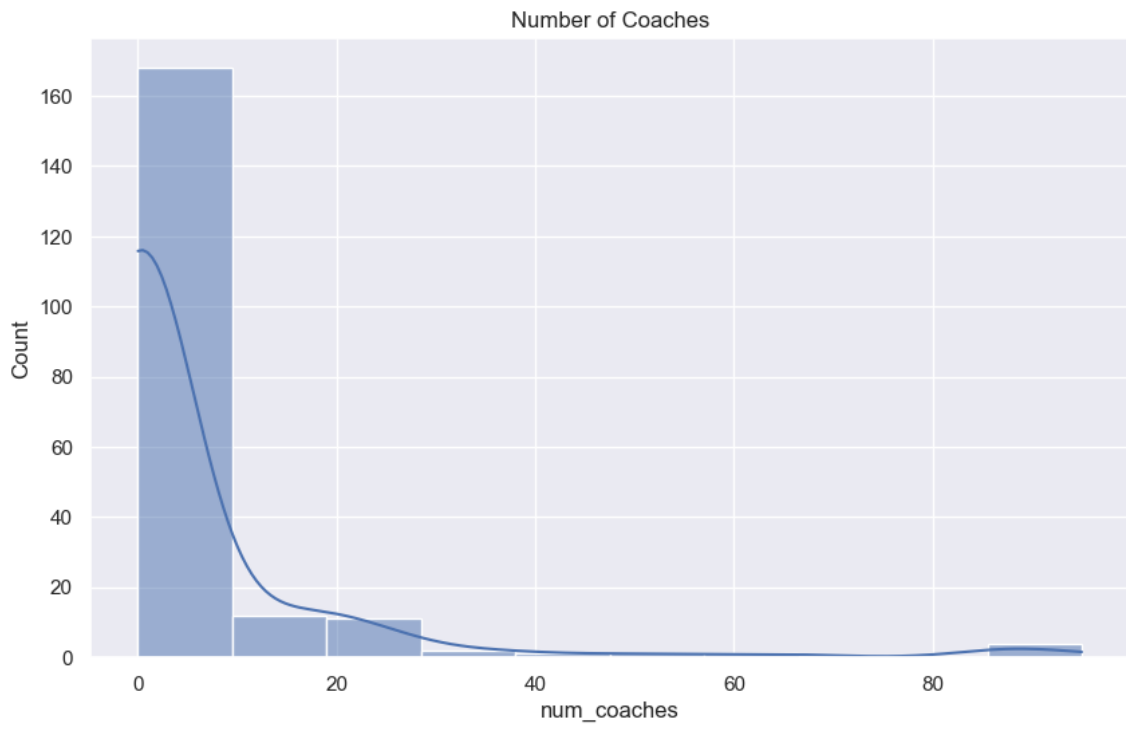
Figure 5
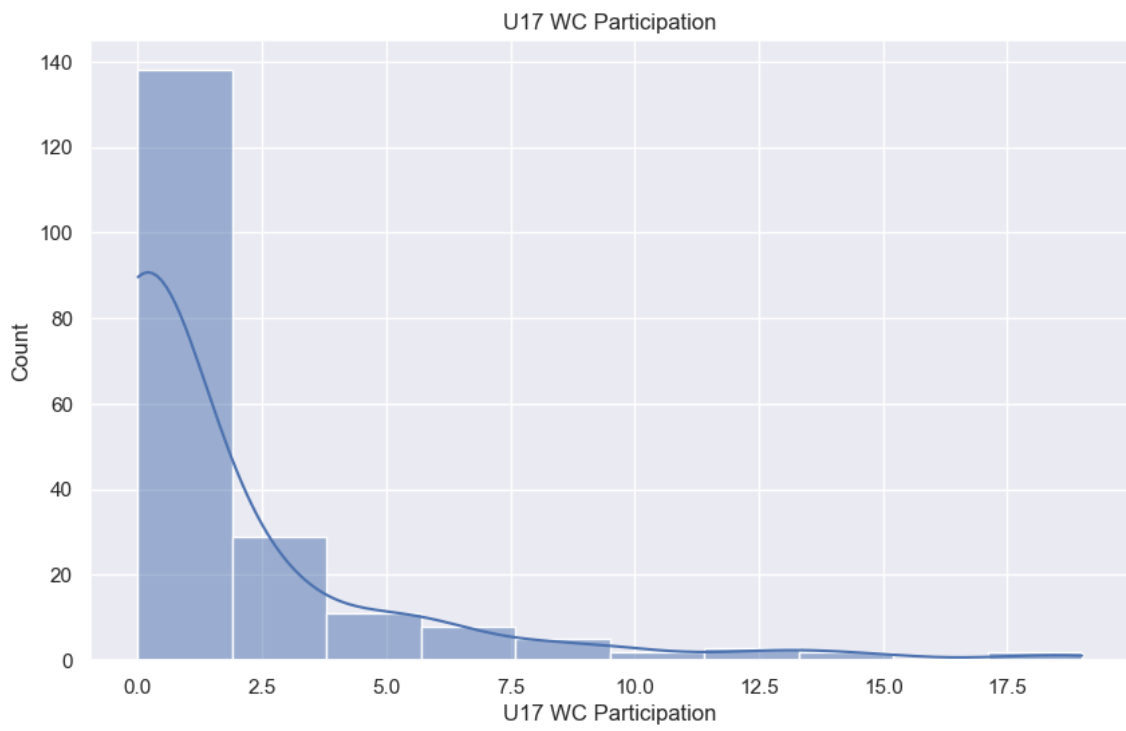
Figure 6
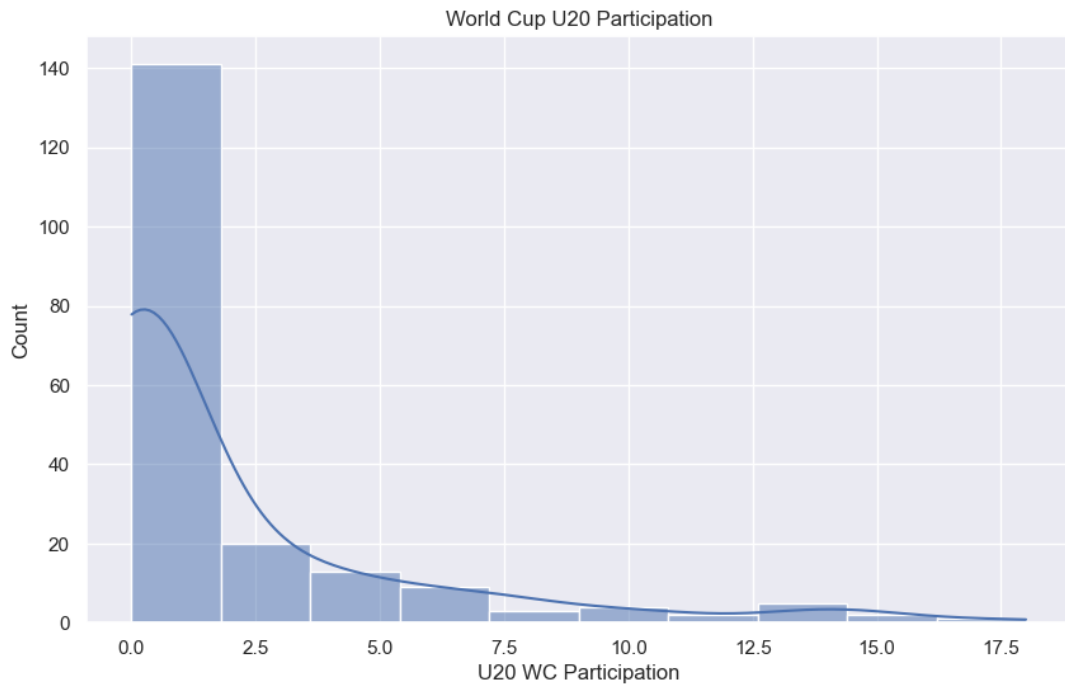
Figure 7
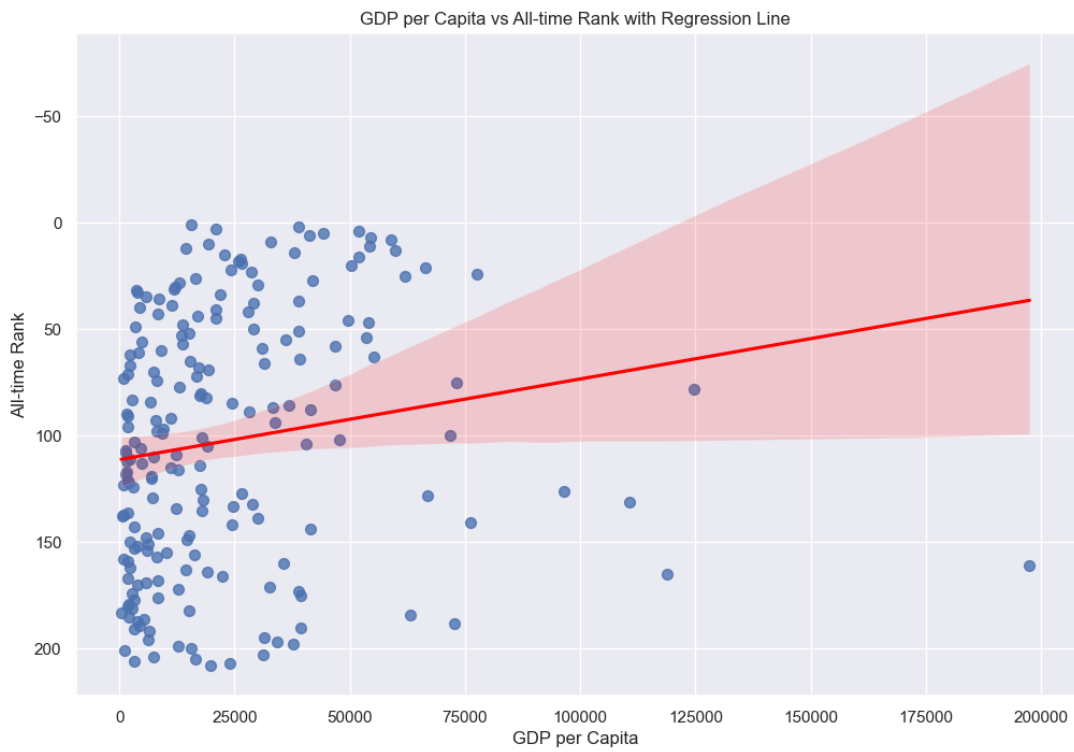
Figure 8

Table 1: OLS Regression Institutions

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.423	
Method:	Least Squares	Adj. R-squared:	0.421	
		F-statistic:	145.4	
Prob (F-statistic):			1.81e-25	
Log-Likelihood:			-1028.1	
No. Observations:			200	
Df Residuals:			198	
Df Model:			1	
Covariance Type:	nonrobust			

	coef	std err	t	P> |t|
const	120.3907	3.355	35.889	0.000
Institutional Index	-2.1884	0.181	-12.060	0.000

Table 2: OLS Regression Institutions and GDP per Capita

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.426	
Method:	Least Squares	Adj. R-squared:	0.421	
		F-statistic:	73.24	
Prob (F-statistic):			1.65e-24	
Log-Likelihood:			-1027.6	
No. Observations:			200	
Df Residuals:			197	
Df Model:			2	
Covariance Type:	nonrobust			

	coef	std err	t	P> |t|
const	122.8139	4.118	29.824	0.000
Institutional Index	-2.1573	0.184	-11.723	0.000
GDP_per_capita	-0.0001	0.000	-1.014	0.312

Table 3: OLS Regression Cultural

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.288	
Method:	Least Squares	Adj. R-squared:	0.285	
		F-statistic:	80.24	
Prob (F-statistic):			2.45e-16	
Log-Likelihood:			-1049.1	
No. Observations:			200	
Df Residuals:			198	
Df Model:			1	
Covariance Type:	nonrobust			
	coef	std err	t	P > |t|
const	161.6183	7.529	21.467	0.000
Cultural Index	-1.4640	0.163	-8.958	0.000

Table 4: OLS Regression Cultural and GDP per Capita

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.320	
Method:	Least Squares	Adj. R-squared:	0.313	
		F-statistic:	46.29	
Prob (F-statistic):			3.31e-17	
Log-Likelihood:			-1044.6	
No. Observations:			200	
Df Residuals:			197	
Df Model:			2	
Covariance Type:	nonrobust			
	coef	std err	t	P > |t|
const	170.7793	7.982	21.396	0.000
Cultural Index	-1.4771	0.160	-9.217	0.000
GDP_per_capita	-0.0004	0.000	-3.012	0.003

Table 5: OLS Regression Cultural and Institutional

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.576	
Method:	Least Squares	Adj. R-squared:	0.572	
		F-statistic:	134.0	
Prob (F-statistic):			1.85e-37	
Log-Likelihood:			-997.28	
No. Observations:			200	
Df Residuals:			197	
Df Model:			2	
Covariance Type:	nonrobust			
	coef	std err	t	P > |t|
const	163.0601	5.826	27.990	0.000
Cultural Index	-1.0985	0.130	-8.429	0.000
Institutional Index	-1.8599	0.161	-11.570	0.000

Table 6: OLS Regression Cultural + Institutional + GDP

	Dep. Variable:	All Time Rank		
Model:	OLS	R-squared:	0.583	
Method:	Least Squares	Adj. R-squared:	0.577	
		F-statistic:	91.38	
Prob (F-statistic):			4.98e-37	
Log-Likelihood:			-995.66	
No. Observations:			200	
Df Residuals:			196	
Df Model:			3	
Covariance Type:	nonrobust			

	coef	std err	t	P> |t|
const	167.3613	6.272	26.684	0.000
Cultural Index	-1.1149	0.130	-8.581	0.000
Institutional Index	-1.8080	0.162	-11.128	0.000
GDP_per_capita	-0.0002	9.91e-05	-1.790	0.075