Sanghoon, L. (2020). Income and international soccer performance. *Journal of Sports Economics & Management, 10*(3), 139-154.

INCOME AND INTERNATIONAL SOCCER PERFORMANCE

Ingresos y rendimiento del fútbol internacional

Sanghoon Lee

Department of Economics, Hannam University, Republic of Korea

ABSTRACT: This study empirically examines how income affects a country's performance in international soccer games. Previous studies suggest a positive link between income and soccer performance as well as an inverse U-shaped relationship. The paper contributes to the literature by using various estimation techniques such as a dynamic GMM panel data approach and explaining the mixed results of the existing empirical studies. The empirical results show that the cross-sectional analysis confirms the positive effect of income on soccer performance, while the panel data analysis supports the inverse U-shaped relationship. This evidence implies that high income countries perform well in international soccer, while low income countries perform poorly. On the other hand, as a country's income rises, soccer performance improves at a decreasing rate, and then gets worse beyond some level of income. The inverse U-shaped relationship is also supported by the split sample regression result that the positive effect is found in low income countries only.

KEYWORDS: Soccer, FIFA World Ranking, Income

RESUMEN: *Este estudio examina empíricamente cómo la renta afecta al rendimiento de un país en los partidos de fútbol internacionales. Los estudios anteriores sugieren una relación positiva entre la renta y el rendimiento futbolístico, así como una relación en forma de U inversa. El artículo contribuye a la literatura utilizando varias técnicas de estimación, como un enfoque dinámico de datos de panel GMM, y explicando los resultados mixtos de los estudios empíricos existentes. Los resultados empíricos muestran que el análisis transversal confirma el efecto positivo de la renta en el rendimiento futbolístico, mientras que el análisis de datos de panel apoya la relación en forma de U inversa. Esta evidencia implica que los países de renta alta obtienen buenos resultados en el fútbol internacional, mientras que los países de renta baja obtienen malos resultados. Por otra parte, a medida que aumenta la renta de un país, el rendimiento futbolístico mejora a un ritmo decreciente, y luego empeora a partir de cierto nivel de renta. La relación en forma de U inversa también se ve respaldada por el resultado de la regresión de muestra dividida, según el cual el efecto positivo se encuentra únicamente en los países de renta baja.*

PALABRAS CLAVE: Fútbol, clasificación mundial de la FIFA, ingresos

Contact information:

Corresponding author

Sanghoon Lee
sanghoonlee@hnu.kr
Hannam University
70 Hannam-ro, Daedeok-gu,
Daejeon 306-791, South Korea

1. Background

Soccer has grown to a billion-dollar global industry, which generates huge financial revenues. The success of national soccer teams brings financial and socio-economic rewards to the country's economy (Hoffmann et al., 2002b, pp.254-255). Thus, it is important to identify the determinants of success in international soccer. Previous empirical studies show that economic factors are important in explaining international soccer performance. This paper examines the relationship between international soccer performance and national income.

Some argue that income of a country may be neutral to soccer performance since soccer requires little equipment. Others hold that high-income countries are likely to spend more resources on leisure activities, and thus such countries are expected to perform better in international soccer competition.

Previous studies tend to confirm the positive relationship between income and international soccer performance. Houston and Wilson (2002) use data for 179 countries to estimate international soccer rankings and report that leisure proficiency is positively related to the level of income. Gelade and Dobson (2007) investigate the data of 201 national teams and provide evidence of a positive effect of national wealth. Based on data for 178 countries, Leeds and Leeds (2009) show that countries with higher GDP per capita outperform poorer countries. The same result is obtained for a sample of African countries (Luiz and Fadal, 2011) and for women's international soccer (Hoffmann et al., 2006). Yamamura (2009, 2012) also finds the same effect. A recent paper by Batarfi and Reade (2020) shows that higher level of GDP per capita leads a country to win more often.

In contrast, in a seminal paper, Hoffmann et al. (2002b) report regression results that there is an inverse U-shaped relationship between the FIFA ranking and per capita wealth. While higher income is expected to lead to better soccer performance, after a certain level, higher income is no longer associated with better performance. The authors claim that soccer is an inexpensive sport that can be played by poorer children with little equipment and, as a country becomes wealthier, people will be more likely to engage in expensive activities. Macmillan and Smith (2007) extend the data set of Hoffmann et al. (2002b) by adding 100 more countries to the sample and confirm the inverse U-shaped pattern.

As discussed above, previous studies have produced mixed results. The findings of some studies support a positive effect of income on soccer performance, while others suggest an inverse U-shaped relationship, which are summarized in Table 1. In addition, a recent study by Wan et al. (2020) finds an insignificant relationship between FIFA score and GDP per capita.

These mixed findings might be due to methodological limitations. The existing studies aim at identifying the determinants of international soccer performance and usually employ regression analysis with cross-sectional data. However, cross-sectional studies do not capture a causal relationship. Social, cultural and political factors may affect both

economic development and sports performance. If this is the case, the OLS regression of income on soccer performance may produce a spurious relationship. In addition, since it is almost impossible to control all potential factors relevant to determining soccer performance, omitted variable bias may be a serious concern.

Table 1. Previous empirical studies of the FIFA ranking

Literature	sample	explanatory variables	findings
Hoffmann et al. (2002b)	76	GNP per capita	∩
		temperature	∩
Houston and Wilson (2002)	179	GDP per capita	+
		population	+
Hoffmann et al. (2006)	88	GDP per capita	+
		temperature	∩
		population	+
Gelade and Dobson (2007)	201	GDP per capita	+
		medium-pressure	+
		regular players	+
Macmillan and Smith (2007)	176	GNP per capita	∩
		temperature	∩
		population	+
Leeds and Leeds (2009)	178	GDP per capita	+
		population	+
Yamamura (2009)	unknown	GDP per capita	+
		population	+
Luiz and Fadal (2011)	152	GDP per capita	+
		population	o
		mild climate	+
Yamamura (2012)	unknown	GDP per capita	+
		population	+

Notes: The table shows previous empirical studies of the determinants of international soccer performance. +, -, o, U, and ∩ refer to a positive, negative, insignificant, U-shaped, and inverse U-shaped relationship between variables, respectively.

2. Methods

This study conducts regression analysis to test the relationship between income and international soccer performance. The current analysis contributes to the literature on international soccer performance by using various estimation techniques. We employ panel data and a dynamic GMM model to address some of the methodological concerns. The details of the methods and results obtained are presented and discussed below.

2.1. A cross-section analysis

Following most previous research, we begin with cross-sectional data so that we can compare our results with the previous results. The ordinary least squares (OLS) model used in the study is represented by the following regression equation:

$$FIFA\ rank_{i,t} = \beta_0 + \beta_1 log(gdp/n)_{i,t-1} + \beta_2 log(gdp/n)^2_{i,t-1} + \beta_3 (temp-14)^2_{i,t} + \beta_4 log(players)\}\_\{i,t\} + \varepsilon_{i,t} \quad (1)$$

where the subscript i refers to the country, t to time period (t=2014), and ε to the error

term.

For a dependent variable of interest, we employ the FIFA World ranking (*FIFA rank*) as a proxy for international soccer performance. The FIFA ranking is based on the average points gained from international matches and the data are obtained from the FIFA website (http://www.fifa.com). The FIFA ranking was introduced in 1993. Among the 209 FIFA member countries, we exclude countries with missing data from the data set (*n=188*). We add a negative sign to the ranking, so that higher value represents better performance.

As is usual in the literature, the log of GDP per capita (*log(gdp/n)*) is used as a proxy for income. The rationale behind the use of income as an explanatory variable is that high income leads to more leisure time for sports activity and improved performance. In addition, income represents public/private funding for sport and access to infrastructure, which are important for the development of soccer talent. Thus, the positive effect of income on soccer performance is expected. However, since soccer is a capital-unintensive sport, the positive effect is expected in low income countries but not in high income countries (Hoffmann et al., 2002b; Houston and Wilson, 2002). This can lead to diminishing returns or an inverse U-shaped relationship. To identify this effect, the quadratic variable of per capita GDP is included in the model. The GDP data are obtained from the World Bank's World Development Indicators.

In addition to the explanatory variable of income, two control variables are employed in the cross-sectional analysis: temperature ($(temp-14)^2$) and the log of the number of soccer players (*log(players)*). It is natural to expect that countries with mild climates are successful at sports. Climates with extremes in temperatures are obstacles to outdoor activity, which will compromise the development of talent in sport. Hoffmann et al. (2002a) show that the average annual temperature of 14 degrees Celsius is optimal for sports performance. Thus, following the work of Hoffmann et al. (2002b), we include the squared deviation of temperature from the 14 degree Celsius in the regression equation. The temperature data are collected from the Weatherbase website (http://www.weatherbase.com).

Countries with a large talent pool are likely to be successful at the sport. Gelade and Dobson (2007) hold that the strength of the national team will be a function of the pool size from which the team is chosen. This study employs the log of the number of both registered and unregistered soccer players to control the size of the talent pool. The players data are obtained from the 2006 Big Count of the FIFA website (http://www.fifa.com).

2.2. A panel analysis

International soccer performance is affected by various factors. For example, Foer (2004, 2006) argues that a country's soccer performance at the international level is dependent upon its social, cultural, and political climate. These factors are difficult to classify and measure in empirical studies. A cross-sectional approach does not account for the unobserved heterogeneity across countries, and thus may cause a serious omitted variable bias. Panel data analysis allows us to overcome the limitation inherent in cross-sectional studies (Baltagi, 2005, pp.4-7). Thus, we investigate country-level panel data over the

period 1994-2014 to examine the relationship between soccer performance and income.

This paper uses the panel data regression model given below:

$$FIFA\ rank_{i,t} = \beta_0 + \beta_1 log(gdp/n)_{i,t-1} + \beta_2 log(gdp/n)^2_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

where the subscript *i* refers to the country, *t* to time period, and ε to the error term. The two control variables used in the cross-sectional estimation are not included due to lack of panel data.

In studies of panel data, it is common to use two econometric techniques--fixed effects model and random effects model. We conduct a Hausman test to determine which model is more appropriate to our data. The current panel study uses fixed effects model, which is shown to be appropriate to the sample data by the Hausman test. Year dummies are included in all regressions to control for the time effect. The results of the Hausman test and the year dummies are not reported for brevity.

Similarly to the above method, Yamamura (2009, 2012) applies static panel data regression models to examine the FIFA ranking. We add to the studies by using a dynamic panel data model represented by the following equation:

$$FIFA\ rank_{i,t} = \beta_1 log(gdp/n)_{i,t-1} + \beta_2 log(gdp/n)^2_{i,t-1} + \beta_3 FIFA\ rank_{i,t-1} + \varepsilon_{i,t} \quad (3)$$

where the dependent variable (*FIFA rank*) is included as an independent variable as well. The static OLS or panel models assume fixed exogenous variables, but this imposes unrealistic restrictions on dynamic behavior and thus leads to inconsistent estimators when the fixed regressor assumption is relaxed (Wooldridge, 2002, p.10). The dynamic model used in the study allows us to overcome the shortcoming of the static models and study the dynamic relationships between variables over time.

The dynamic model uses the *t-1* lagged variable of the FIFA ranking as an independent variable, which plays the role of a control variable to capture unobserved characteristics of a country. Thus, the dynamic model addresses the problem of the lack of control variables in the static panel model. Since the independent variable is likely to be endogenous in a standard dynamic model, the OLS estimates will be biased and inconsistent. Thus, we estimate Equation 3 by the difference GMM method (Arellano and Bond, 1991) to obtain consistent and efficient estimates. We use *t-2* through *t-4* lagged values of the FIFA ranking as GMM instruments since very remote lags are not informative in practice (Bond and Meghir, 1994). The Sargan test of overidentifying restrictions and the second-order autocorrelation test are conducted to evaluate the validity of the instruments and the specification of the model.

2.3. A split sample analysis

This study uses a split sample approach. We divide countries into groups by characteristics and perform the regressions to examine the effect of income on soccer performance in each of the groups. If the effects are significantly different between groups, the observed difference should indicate the pure effect of income on soccer performance. Even though individual estimates of the income coefficients may be biased, the estimated difference in the coefficients between groups is an unbiased estimate of the true difference.

In the study, we split the sample countries into two groups i) based on the level of income and ii) based on the number of (both registered and unregistered) soccer players. First, we separate countries into two equal-sized groups according to the level of income as measured by the 2013 GDP per capita: low income countries and high-income countries. We expect that, if, as argued by Hoffmann et al. (2002b), the inverse U-shaped relationship exists, the positive effect of income on soccer performance is pronounced for low income countries.

Second, the sample countries are divided into two equal-sized groups by the number of soccer players: countries with a low number of players and countries with a high number of players. We would expect that the positive effect of income, if any, on soccer performance would be much more noticeable in countries with a high popularity of soccer. It is reasonable to suppose that, in countries in which soccer is not popular, high income may provide resources for sports other than soccer.

Table 2. Summary statistics

		Cross-section			Panel		
		median	***mean***	***s.d.***	***median***	***mean***	***s.d.***
all	*FIFA rank*	99.50	99.92	58.65	94.00	95.71	56.34
	log(gdp/n)	8.75	8.69	1.51	8.01	8.03	1.65
	temp	21.20	18.91	7.58			
	log(players)	12.84	12.51	2.00			
low income	*FIFA rank*	114.00	114.50	54.91	111.00	111.30	51.19
	log(gdp/n)	7.46	7.42	0.86	6.69	6.73	0.95
	temp	23.70	21.56	5.83			
	log(players)	12.97	12.74	1.61			
high income	*FIFA rank*	79.00	85.66	58.96	73.50	81.04	57.01
	log(gdp/n)	9.82	9.94	0.81	9.40	9.32	1.08
	temp	15.75	16.29	8.19			
	log(players)	12.59	12.29	2.31			
low players	*FIFA rank*	114.00	116.40	49.94	108.00	110.20	50.20
	log(gdp/n)	7.73	8.08	1.56	6.91	7.34	1.63
	temp	23.70	20.89	6.80			
	log(players)	12.80	12.60	1.80			
high players	*FIFA rank*	69.00	83.40	62.22	70.00	81.90	58.36
	log(gdp/n)	9.23	9.31	1.18	8.59	8.70	1.36
	temp	17.35	17.04	7.83			
	log(players)	12.90	12.42	2.19			

Notes: The table shows the summary statistics for the variables used in the study. *s.d.* refers to standard deviation. This study uses data of 188 countries.

Summary statistics are listed in Table 2. The table shows that the FIFA rank tends to be higher for the high-income countries than for the low-income countries, and to be higher for the high players countries than for the low players countries, as usually expected.

2.4 A Robustness Method

In addition to the main analysis explained above, we collect additional data for the period 2015 to 2021 and conduct the panel analysis using the data 1994-2021. In this extended sample, the number of countries are 209, and the total number of observations is 5138. The data are composed of FIFA ranking and GDP per capita. For the natural log of GDP per capita, the median is 8.351, the mean is 8.357, and the standard deviation is 1.647. We conducted the F test, LM test, and Hausman test (Hausman, 1978) to statistically

determine whether we use fixed effects or random effects, and the test results indicated that the fixed effects model was appropriate for the data used in this analysis.

3. Findings

3.1. Full Sample

In this section we report the empirical findings. As a preliminary step, we present scatter plots of log(gdp/n) and -FIFA rank in Figure 1 (cross-sectional data) and Figure 2 (panel data). From the figures we observe that the relationship between national income and soccer performance is positive.

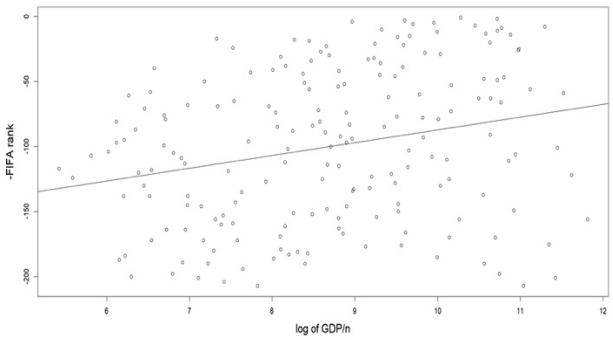

Figure 1. log(gdp/n) and -FIFA rank Cross-sectional data)

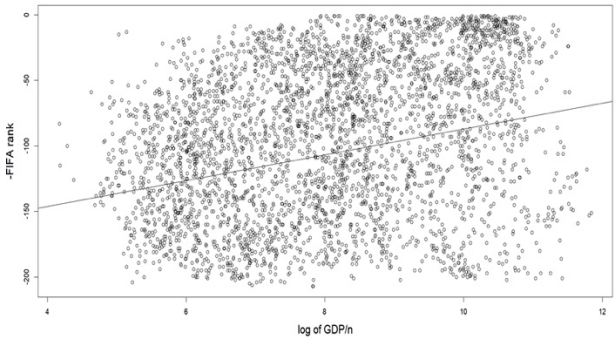

Figure 2. log(gdp/n) and -FIFA rank (Panel data)

First of all, the results of the regression analysis for the full sample are discussed. The regression results for the full data using the equations 1, 2, and 3 are presented in Table 3. In the OLS results, while the linear model produces a positive and significant estimate of income, the quadratic model does not present a significant effect. Thus, the cross-sectional results support the hypothesis of a positive influence of income on soccer performance, but not the inverse U-shaped pattern. The coefficient estimates for the control variables have signs consistent with expectations. The inverted U-shaped effect of temperature and the positive effect of the number of players on soccer performance are confirmed by the OLS results.

In contrast to the OLS results, the fixed effects results confirm the inverse U-shaped relationship as well as the positive relationship between income and soccer performance. For the quadratic regression, the linear term is positive and the quadratic term is negative, all of which are statistically significant. However, the result does not necessarily guarantee an inverse U-shaped relationship. "To do so would require the demonstration of an inflection point beyond which the curve becomes downward sloping, as opposed to just asymptotic, and a demonstration that this point is not just a statistical abstraction, but that it is within the range of acceptable or realistic values of the independent variable" (Herold et al., 2006, p.384).

Table 3. Regression results. Full sample.

	OLS		Fixed effects		Panel GMM	
$log(gdp/n)_{t-1}$	12.8453***	29.6178	4.4787**	17.6879***	1.2787	31.4115**
	(5.9480)	(1.3090)	(3.1446)	(4.0778)	(0.4048)	(2.7573)
$log(gdp/n)^2_{t-1}$		-0.9698		-0.8881**		-1.9801**
		(-0.7440)		(-3.2235)		(-2.6616)
$(temp-14)^2_t$	-0.1226*	-0.1225*				
	(-2.2780)	(-2.2720)				
$log(players)_t$	17.3366***	17.3282***				
	(10.5360)	(10.5170)				
$FIFA\ rank_{t-1}$					0.6705***	0.5954***
					(17.2028)	(11.8445)
R^2	0.4897	0.4913	0.0054	0.0084		
Sargan					0.0605	0.0626
AR(2)					0.0631	0.0547

Notes: The table shows the full sample regression results. Figures are regression coefficient estimates, and *t* values are shown in parentheses below coefficient estimates. ***, **, and *, respectively, indicate significance levels at 0.1%, 1%, and 5% levels. *Sargan* and *AR(2)* refers to p values for the Sargan test and the autocorrelation test for AR(2) process, respectively.

Thus, we determine whether the requirement is satisfied or not by plotting the relationship between income and soccer performance using the fixed effects estimates.

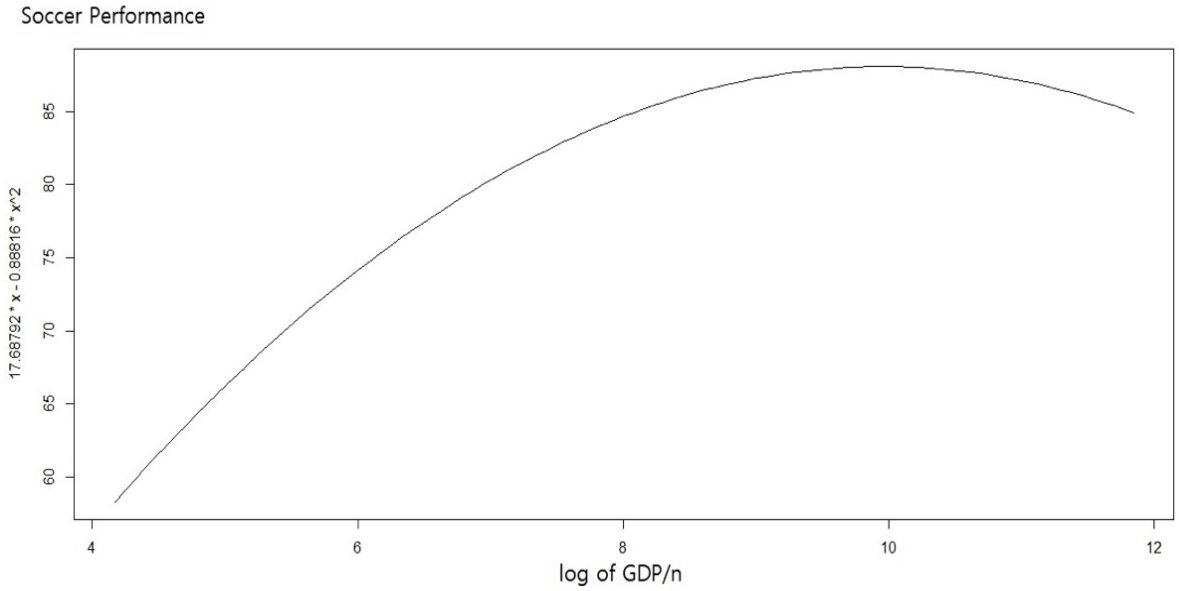

Figure 3. Fixed effects regression with the full sample

Figure 3 shows the relationship of the fixed effects regression, in which the horizontal axis represents the level of income and the vertical axis represents the level of international soccer performance. Since the figure indicates that the inflexion point is within the acceptable range, the inverted U-shaped relationship is confirmed in the fixed effects model.

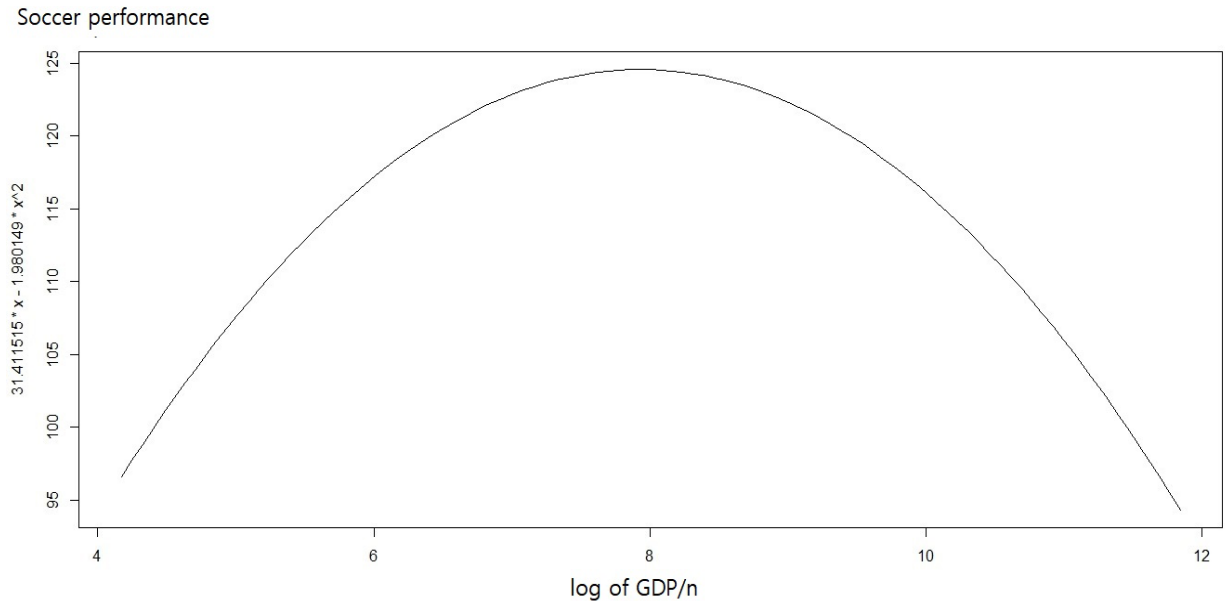

Figure 4. Dynamic GMM regression with the full sample

The dynamic panel GMM estimation corroborates the inverted-U relationship. In the dynamic model, while the linear regression does not produce a significant coefficient, the quadratic regression shows significant coefficients. The results appear to support the inverse U-shaped pattern, but not the positive effect. The requirement for an inverse U-shaped relationship is satisfied, which is illustrated in Figure 4. When comparing the fixed effects result and the panel GMM result, the inverted-U shape is more pronounced in the dynamic regression than in the fixed effects regression. For the dynamic panel regressions, neither the Sargan test nor the second-order autocorrelation test rejects the null hypothesis at the 5 percent level. Since the null hypothesis is that the specification of the model is appropriate and the instruments are valid (Hansen, 1982; Sargan, 1958), there is no problem with the model specification and the instruments.

3.2. Split Sample

Here the results of the split sample regressions are discussed. First, the regression results based on the division between high income and low-income countries are presented in Table 4.

For the low-income countries, the OLS and fixed effects regressions support the positive effect of income on soccer performance as expected, while the panel GMM does not provide significant results. In contrast, for the high-income countries, all the three regressions do not show a significant positive effect.

Table 4. Regression results: split sample I

OLS	**low income**		**high income**	
$log(gdp/n)_{t-1}$	12.5073*	-166.9297	7.4829	257.5000*
	(2.1000)	(-1.5780)	(1.5890)	(2.1860)
$log(gdp/n)^2_{t-1}$		12.2903		-12.4300*
		(1.6980)		(-2.1240)
$(temp-14)^2_t$	-0.1460	-0.1170	-0.0788	-0.0895
	(-1,7840)	(-1.4140)	(-1.1130)	(-1.2860)
$log(players)_t$	15.3106***	15.4948***	18.7305***	18.2000***
	(4.7890)	(4.8970)	(10.3010)	(10.1070)
R^2	0.2682	0.2922	0.6328	0.6511
Fixed effects	low income		high income	
$log(gdp/n)_{t-1}$	5.8582**	12.3133	3.3581	20.3913*
	(2.7383)	(1.2860)	(1.7347)	(2.1779)
$log(gdp/n)^2_{t-1}$		-0.4795		-1.0675
		(0.6917)		(-1.8593)
R^2	0.0103	0.0106	0.0143	0.0163
Panel GMM	low income		high income	
$log(gdp/n)_{t-1}$	6.7526	-35.3390	3.1450	-9.6902
	(0.7638)	(-0.5651)	(0.5633)	(-0.2879)
$log(gdp/n)^2_{t-1}$		2.8269		0.7248
		(0.5917)		(0.3642)
$FIFA\ rank_{t-1}$	0.6912***	0.6304***	0.5636***	0.5602***
	(8.9826)	(7.2838)	(8.1779)	(8.2556)
Sargan	0.9415	0.9689	0.9325	0.9340
AR(2)	0.1141	0.1335	0.2055	0.2033

Notes: The table shows the split sample regression results. Figures are regression coefficient estimates, and *t* values are shown in parentheses below coefficient estimates. ***, **, and *, respectively, indicate significance levels at 0.1%, 1%, and 5% levels. *Sargan* and *AR(2)* refers to p values for the Sargan test and the autocorrelation test for AR(2) process, respectively.

On the other hand, the OLS quadratic regression provides significant results, which is shown in Figure 5. The figure seems to confirm the diminishing trend rather than the inverse U-shaped pattern. Overall, the first split sample analysis supports that international soccer performance improves as income increases when income levels are in the lower stages, but when the income level is at a high stage, little relationship between soccer performance and income is observed.

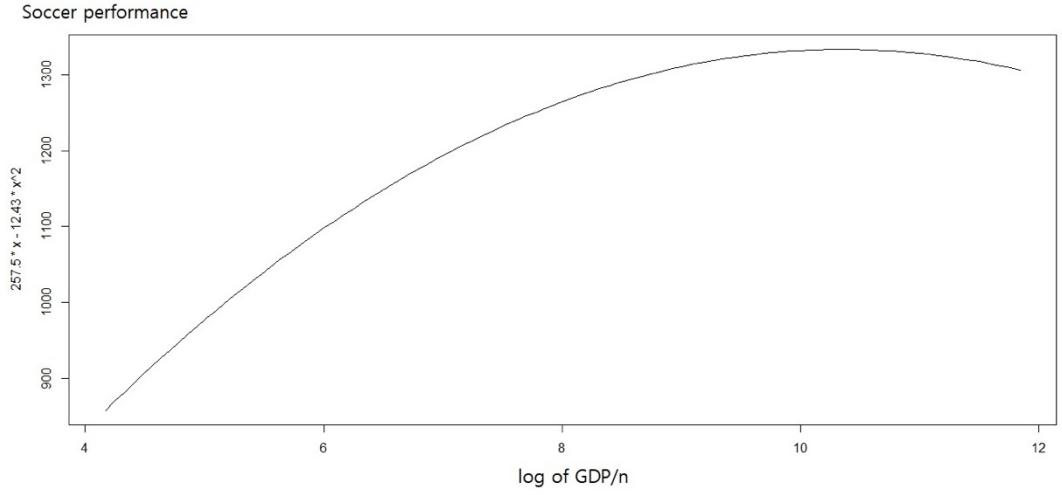

Figure 5. OLS regression with the high-income sample

The results of the split sample regression based on the number of players is presented in Table 5.

Table 5. Regression results: split sample II

OLS	**low players**		**high players**	
$log(gdp/n)_{t-1}$	10.0703**	-40.3391	8.2273*	91.7143*
	(3.1870)	(-1.1610)	(2.5230)	(2.3580)
$log(gdp/n)^2_{t-1}$		3.0207		-4.5016*
		(1.4560)		(-2.1530)
$(temp-14)^2_t$	-0.0595	-0.0611	-0.1320	-0.1293
	(-0.7990)	(-0.8260)	(-1.8020)	(-1.8000)
$log(players)_t$	11.2154***	11.6195***	21.1448***	21.1363***
	(4.0720)	(4.2260)	(11.1760)	(11.3950)
R^2	0.2283	0.2478	0.6722	0.6885
Fixed effects	low players		high players	
$log(gdp/n)_{t-1}$	3.6697	19.6154**	6.1742**	25.2426***
	(1.8960)	(3.2583)	(2.7820)	(3.3245)
$log(gdp/n)^2_{t-1}$		-1.0695**		-1.2502**
		(-2.7966)		(-2.6256)
R^2	0.0123	0.0170	0.0131	0.0171
Panel GMM	low players		high players	
$log(gdp/n)_{t-1}$	2.7782	31.9464	2.2710	27.2981
	(0.4361)	(1.2198)	(0.4662)	(0.8903)
$log(gdp/n)^2_{t-1}$		-1.9637		-1.7069
		(-1.0909)		(-0.9087)
$FIFA\ rank_{t-1}$	0.6636***	0.6219***	0.5342***	0.5128***
	(11.6339)	(8.8369)	(7.6789)	(7.1194)
Sargan	0.9125	0.9166	0.9739	0.9657
AR(2)	0.1819	0.2754	0.5779	0.5868

Notes: The table shows the split sample regression results. Figures are regression coefficient estimates, and *t* values are shown in parentheses below coefficient estimates. ***, **, and *, respectively, indicate significance levels at 0.1%, 1%, and 5% levels. *Sargan* and *AR(2)* refers to p values for the Sargan test and the autocorrelation test for AR(2) process, respectively.

According to the OLS results, the positive effect of income on soccer performance is confirmed for both 'low players' and 'high players' countries. In addition, for the 'high players' countries, the fixed effects regression also confirms the positive relationship.

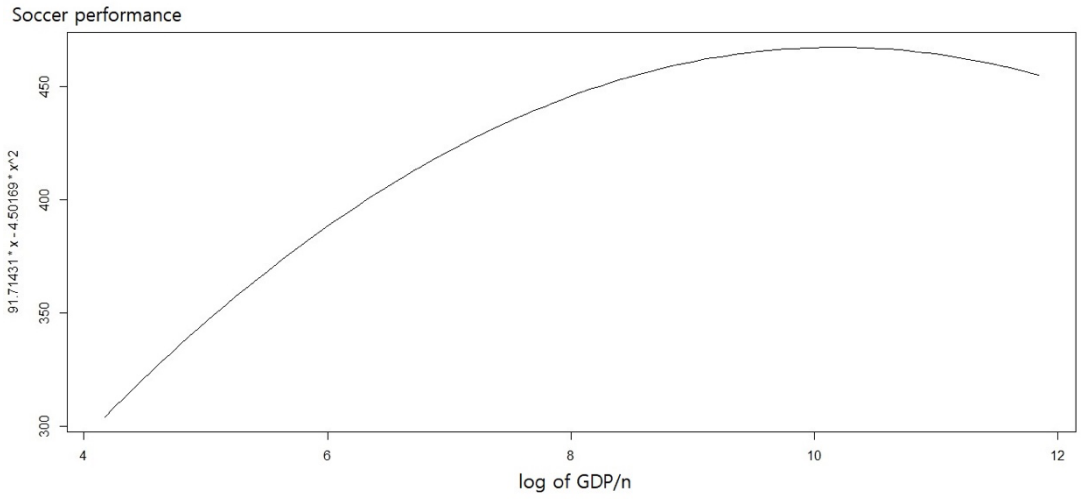

Figure 6. OLS regression with the high-player sample

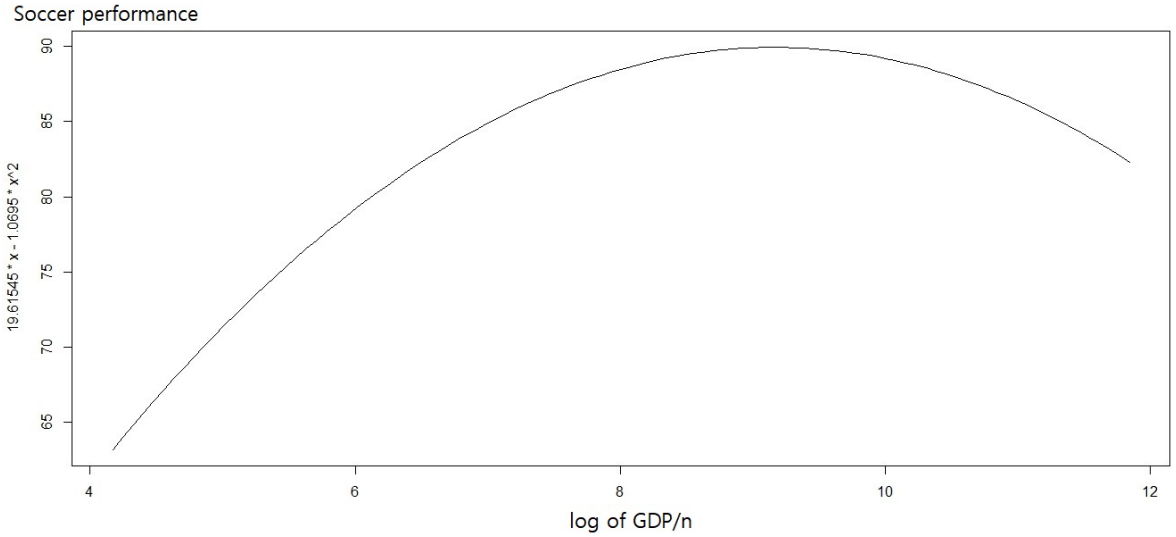

Figure 7. Fixed effects regression with the low-player sample

The quadratic regressions of the cross-sectional and static panel models show statistically significant results, except for the OLS for 'low players' countries. The quadratic regression results are illustrated in Figure 6, Figure 7, and Figure 8. The figures show a pattern of the inverse U-shaped relationship or a diminishing return trend. In contrast to the OLS and fixed effects regressions, the dynamic panel regressions show insignificant results. In sum, the relationship between soccer performance and income is not significantly different between the two groups.

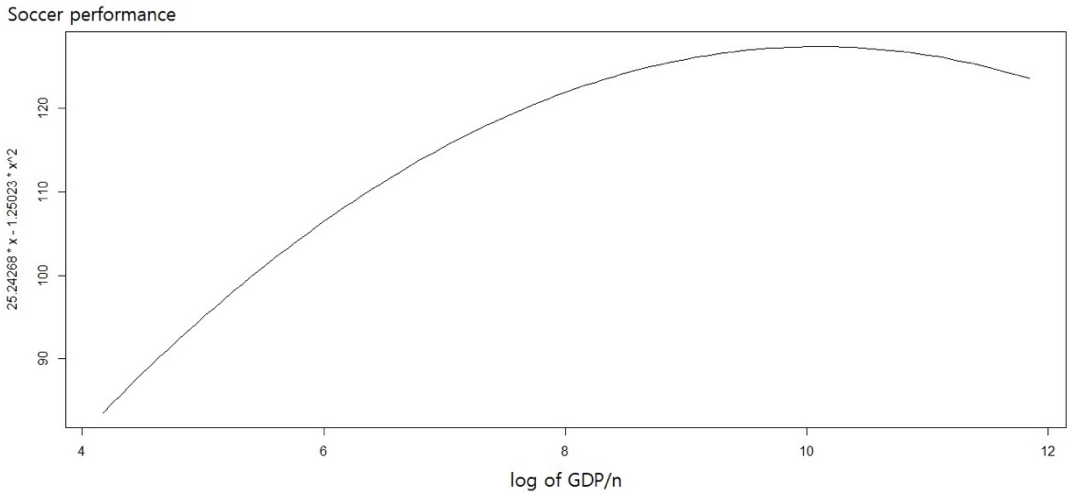

Figure 8. Fixed effects regression with the high-player sample

3.3 Robustness Checks

The panel analysis using the extended data are summarized by Table 6. First, the pooled OLS regression finds the positive relationship between the FIFA ranking and GDP per capita. This confirms the the cross-sectional result mentioned above, which supports the hypothesis of a positive influence of income on soccer performance. Second, the fixed

effects regression reports an insignificant relationship between the FIFA ranking and GDP per capita. This result is consistent with the recent study by Wan et al. (2020), and can be related to a nonlinear relationship. Thus, we conduct a quadratic regression, and it shows an inverse U-shaped relationship between the FIFA ranking and GDP per capita, which is statistically significant.

Table 6. Regression results: extended sample

	OLS	**Fixed effects**		**Panel GMM**
$log(gdp/n)_{t-1}$	7.0017***	1.0736	19.1666***	-2.5393***
	(14.29)	(0.9143)	(5.2959)	(-7.8617)
$log(gdp/n)^2_{t-1}$			-1.2437	
			(-5.2835)	
$FIFA\ rank_{t-1}$				-0.7626***
				-28.0408
R^2	0.0384	0.0033	0.0089	
Sargan				0.1118
AR(2)				0.0298

Notes: The table shows the full sample regression results. Figures are regression coefficient estimates, and *t* values are shown in parentheses below coefficient estimates. ***, **, and *, respectively, indicate significance levels at 0.1%, 1%, and 5% levels. *Sargan* and *AR(2)* refers to p values for the Sargan test and the autocorrelation test for AR(2) process, respectively.

Figure 9 shows the relationship of the fixed effects regression, in which the horizontal axis represents the GDP per capita and the vertical axis represents the level of international soccer performance. The figure confirms that the inverse U-shaped relationship since the inflexion point is within the acceptable range. The quadratic regression result is consistent with the results of the main analysis mentioned above.

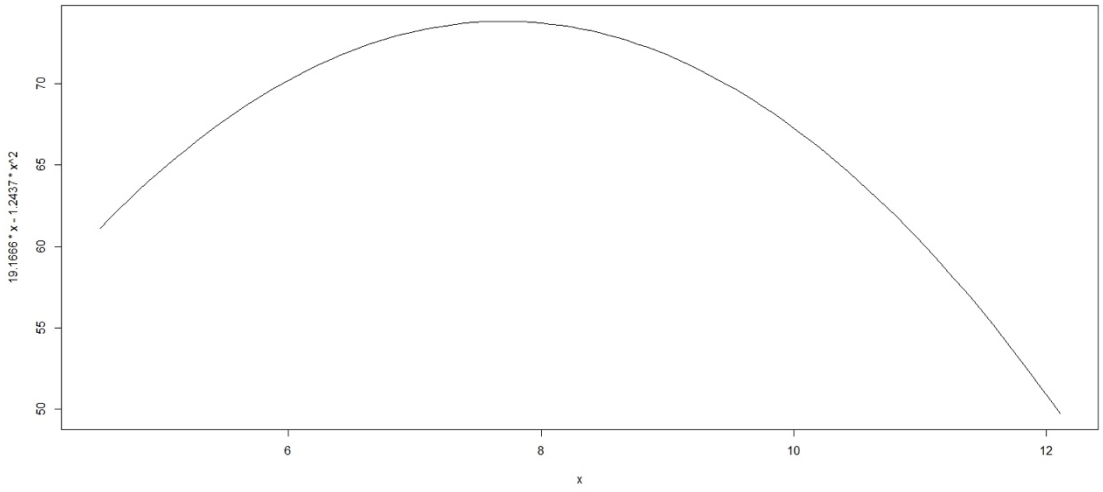

Figure 9. Fixed effects regression with the extended sample

Third, the dynamic GMM regression reports a negative effect of GDP per capita on the FIFA ranking. This is an unexpected result and needs an explanation. It might be due to some autocorrelation problem since the lagged value of the FIFA ranking also shows a negative coefficient and the p-value of AR(2) is 0.02 (less than 0.05). A robust explanation for this negative relationship is left to future work.

4. Conclusion

The effect of economic factors on sports performance is a recent topic of interest. This study investigates how income affects a country's performance in international soccer games. Previous studies suggest a positive link between income and soccer performance as well as an inverse U-shaped relationship between them. These two relationships are empirically examined in the current study. This paper contributes to the literature by using various estimation techniques such as a dynamic GMM panel data approach and explaining the mixed results of the previous empirical studies.

The empirical results for the full sample show that the OLS and fixed effects regressions confirm the positive effect, while the fixed effects and the dynamic GMM regressions confirm the inverse U-shaped pattern. Overall, the cross-sectional analysis shows the positive effect of income on soccer performance, while the panel data analysis supports the inverse U-shaped relationship. This evidence implies that high income countries perform well in international soccer, while low income countries perform poorly. On the other hand, as a country's income rises, soccer performance improves at a decreasing rate, then gets worse beyond some level of income. The inverse U-shaped relationship is also supported by the split sample regression result that the positive relationship is found in low income countries only. The empirical findings are consistent with previous mixed results and reconcile the previous results.

References

Arellano, M. & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies* 58 (2), 277-297.

Baltagi, B. H. (2005). *Econometric analysis of panel data*. West Sussex, England: John Wiley & Sons.

Batarfi, M. & Reade, J. (2020). Why are we so good at football, and they so bad? Institutions and National Footballing Performance. *De Economist* 169, 63-80.

Bond, S. & Meghir, C. (1994). Dynamic investment models and the firm's financial policy. *Review of Economic Studies* 61 (2), 197-222.

Foer, F. (2004). *How Soccer Explains the World: An Unlikely Theory of Globalization*. New York: HarperCollins.

Foer, F. (2006). How governments nurture soccer. http://www.newrepublic.com/article/how-governments-nurture-soccer. Accessed: 2014-11-06.

Gelade, G. A. & Dobson, P. (2007). Predicting the comparative strengths of national football teams. *Social Science Quarterly* 88 (1), 244-258.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50 (4), 1029-1054.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica* 46(6), 1251-1271.

Herold, D. M., Jayaraman, N., & Narayanaswamy, C. R. (2006). What is the relationship between organizational slack and innovation? *Journal of Managerial Issues* 18 (3), 372-392.

Hoffmann, R., Ging, L. C., Matheson, V., & Ramasamy, B. (2006). International women's football and gender inequality. *Applied Economics Letters* 13 (15), 999-1001.

Hoffmann, R., Ging, L. C., & Ramasamy, B. (2002a). Public policy and Olympic success. *Applied Economics Letters* 9 (8), 545-548.

Hoffmann, R., Ging, L. C., & Ramasamy, B. (2002b). The socio-economic determinants of international soccer performance. *Journal of Applied Economics* 5 (2), 253-272.

Houston, R. G. & Wilson, D. P. (2002). Income, leisure and proficiency: An economic study of football performance. *Applied Economics Letters* 9 (14), 939-943.

Leeds, M. A. & Leeds, E. M. (2009). International soccer success and national institutions. *Journal of Sports Economics* 10 (4), 369-390.

Luiz, J. M. & Fadal, R. (2011). An economic analysis of sports performance in Africa. *International Journal of Social Economics* 38 (10), 869-883.

Macmillan, P. & Smith, I. (2007). Explaining international soccer rankings. *Journal of Sports Economics* 8 (2), 202-213.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica* 26 (3), 393-415.

Wan, K-M., Ng, K-U., & Lin, T-H. (2020). The political economy of football: Democracy, income inequality, and men's national football performance. *Social Indicators Research* 151, 981-1013.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.

Yamamura, E. (2009). Technology transfer and convergence of performance: An economic study of FIFA football ranking. *Applied Economics Letters* 16 (3), 261-266.

Sanghoon, L. (2020). Income and international soccer performance. *Journal of Sports Economics & Management, 10*(3), 139-154.

Yamamura, E. (2012). Effect of linguistic heterogeneity on technology transfer: An economic study of FIFA football rankings. *Atlantic Economic Journal* 40 (1), 85-99.