

A Regression-Based Predictive Analysis of National Happiness Index: Evidence from the World Happiness Report (2006–2023)

Chuwei Wang *

School of Data Science, Xiamen University Malaysia, Sepang, Selangor, Malaysia

* Corresponding Author Email: dsc2309260@xmu.edu.my

Abstract. This study investigates the relationship between economic development and national happiness using panel data from the World Happiness Report covering the period from 2006 to 2023. Previous studies have generally suggested that income is positively associated with subjective well-being, but questions remain regarding the stability of this relationship across countries and over time. To address these issues, this study employs a regression-based empirical framework that combines correlation analysis, baseline linear regression, income-group comparison using a Welch two-sample t-test, and a time-interaction regression model. In addition, a prediction framework is developed to estimate happiness outcomes under different economic conditions. The empirical results reveal a strong positive association between log GDP per capita and national happiness levels. Significant differences in happiness are also observed between high-income and low-income countries. However, the interaction analysis suggests that the economic effect on happiness remains relatively stable over the sample period. The prediction model further demonstrates that economic indicators possess substantial explanatory and predictive power for national happiness. Overall, the findings indicate that economic development plays an important but incomplete role in shaping subjective well-being across countries and that non-economic factors may also contribute to differences in happiness levels.

Keywords: Happiness index; GDP per capita; regression analysis.

1. Introduction

In recent years, when discussing social development, researchers have increasingly focused on the concept of subjective well-being. Traditional economic indicators, especially the Gross Domestic Product (GDP), are usually used to measure a country's economic performance. However, these indicators do not always accurately reflect the true level of satisfaction that people have with their lives [1, 2].

Many global studies have shown that there is a certain relationship between income levels and happiness. To be general, the residents in the countries with higher incomes tend to report higher levels of subjective happiness, especially when comparing countries at different stages of development [3-5]. At the same time, relying solely on income cannot fully explain the differences in happiness levels among countries [6].

Some studies have indicated that the impact of income on happiness may vary over time and with different stages of development. Therefore, it is necessary to conduct further research to determine whether the relationship between economic development and happiness is stable, and whether this relationship can be used to predict levels of happiness.

This study analyzed the data from the "World Happiness Report" from 2006 to 2023. It employed a variety of empirical methods, including correlation analysis, regression models, income group comparisons, and time interaction analysis, to examine the role of economic factors in explaining and predicting the happiness levels of different countries.

The structure of this article is as follows: The second part introduces the data sources and main variables; the third part explains the empirical research methods; the fourth part presents the research results and discusses the relationship between economic development and happiness.

2. Data and Variables

2.1. Data source: World Happiness Report (2006–2023)

Data in this study are extracted from World Happiness Report from the year 2006 to 2023. World Happiness Report constructs cross-country data sets from mostly sources like Gallup World Poll which offers good international comparability and wide coverage for countries and regions. In this study, we use country aggregates panel data to explore the statistical relationship between economic development and national happiness, and in the following sections we will carry out income-group comparison, time-stability test and prediction analysis.

2.2. Variable Definitions: Life Ladder, Log GDP Per Capita, and Income Group Classification

Dependent variable: National happiness (Life Ladder score).

Life Ladder reflects the individual's self-evaluation of his/her overall life satisfaction. It is a continuous measure ranging from 0 to 10, and the higher the score, the happier the citizen.

The main explanatory variable: GDP per capita (national economic development). Due to large income differences across countries and possible skewness of the variable, GDP per capita is converted into its natural logarithm named Log GDP per capita. This way, extreme values are attenuated and the linear model fits better.

Income group classification variable. Countries are grouped into high income, middle income or low income groups. This grouping is further used for comparative statistics and group-based regression analysis to test whether economic effect on happiness is the same across income groups.

Table 1. Variable description and data types

Variable Name	Description	Data Type
Country name	The name of the country or territory	Character
Year	The calendar year of the survey response (2006–2023)	Integer
Life Ladder	The national average response to the Cantril Ladder question (0–10), serving as the primary measure of subjective well-being (happiness score)	Numeric (Float)
Log GDP per capita	The natural log of GDP per capita adjusted by purchasing power parity (PPP)	Numeric (Float)
Social support	A measure showing whether people have friends or relatives they can rely on when facing difficulties	Numeric (Float)
Healthy life expectancy at birth	The expected number of years a person can live in good health from birth	Numeric (Float)
Freedom to make life choices	A measure of how satisfied people are with their freedom to make their own life decisions	Numeric (Float)
Generosity	A measure based on how much people donate to charity compared with their income level	Numeric (Float)
Perceptions of corruption	A measure of how people view corruption in government and business	Numeric (Float)
Positive affect	The average level of positive feelings such as happiness, enjoyment, or laughter	Numeric (Float)
Negative affect	The average of previous-day affects measures for worry, sadness, and anger	Numeric (Float)

Although the dataset contains multiple covariates, this paper focuses on Life Ladder as the dependent variable and Log GDP per capita as the primary explanatory variable. The remaining variables are reported for completeness and are not explicitly modeled in the subsequent analysis.

Table 1 summarizes the key variables and their data types used in this study. Life Ladder is treated as the dependent variable, while Log GDP per capita serves as the primary explanatory variable.

3. Methodology

3.1. Correlation Analysis and Baseline Regression Model

The empirical analysis starts with a correlation analysis between national happiness (Life Ladder) and economic development (log GDP per capita). This allows us to obtain some initial ideas regarding the direction and strength of the association between happiness and economic development, as well as to check whether there is any linear relationship between the two variables in the data.

On this basis, we estimate a baseline linear regression model, in which Life Ladder is taken as the dependent variable, and log GDP per capita is taken as the main explanatory variable. The baseline model captures the overall effect of wealth on happiness and contains the key parameter estimates used in the analysis below.

3.2. Income-Group Comparison and Hypothesis Testing

To examine whether happiness differs across stages of economic development, countries are classified into income groups. The average happiness levels of high-income and low-income countries are then compared.

Because the sample groups may differ in variance and sample size, Welch's two-sample t-test is employed to test whether the difference in mean happiness between the two groups is statistically significant. This approach allows for more robust inference under unequal variance conditions.

3.3. Time-Interaction Regression Model

To examine whether the economic happiness indicator is affected by the economic development in different ways at different times, we construct a time-interaction regression model with an interaction term between log GDP per capita and the year variable.

The interaction term reflects the possible changes in the strength of the economic-happiness relationship over time. If the interaction coefficient is statistically significant, it means that the effect of economic development on happiness has changed over the sample period.

Regression analysis has been widely used in empirical studies of happiness economics to estimate the relationship between economic development and subjective well-being [7]. Previous cross-country studies on life satisfaction have also applied similar regression approaches [8].

3.4. Happiness Prediction Framework

In addition to analyzing the relationship between economic development and happiness, this study also established a simple prediction model. By substituting different log GDP per capita values into the regression equation, corresponding predicted values of happiness levels can be obtained.

Meanwhile, this study also calculated the confidence intervals of the predicted values to reflect the uncertainty of the prediction results. This method can be used to evaluate the effectiveness of economic indicators in predicting the level of national happiness.

4. Economic Factors and Happiness

4.1. Overall Relationship between Economic Development and Happiness

However, in order to get an overview of the relationship between economic development and happiness, in this section we first look at the relationship between log GDP per capita and national happiness through graphical and regression analysis.

Table 2. R output for the simple regression of happiness on Log GDP per capita

Variable	Coefficient	Std. Error	T-value	P-value
Intercept	-1.661	0.118	-14.04	< 0.001
Log GDP per capita	0.761	0.012	60.91	< 0.001
Statistic			Value	
Observations			2335	
R-squared			0.614	
Adjusted R-squared			0.614	
Residual Std. Error			0.695	
F-statistic			3710 (p < 0.001)	

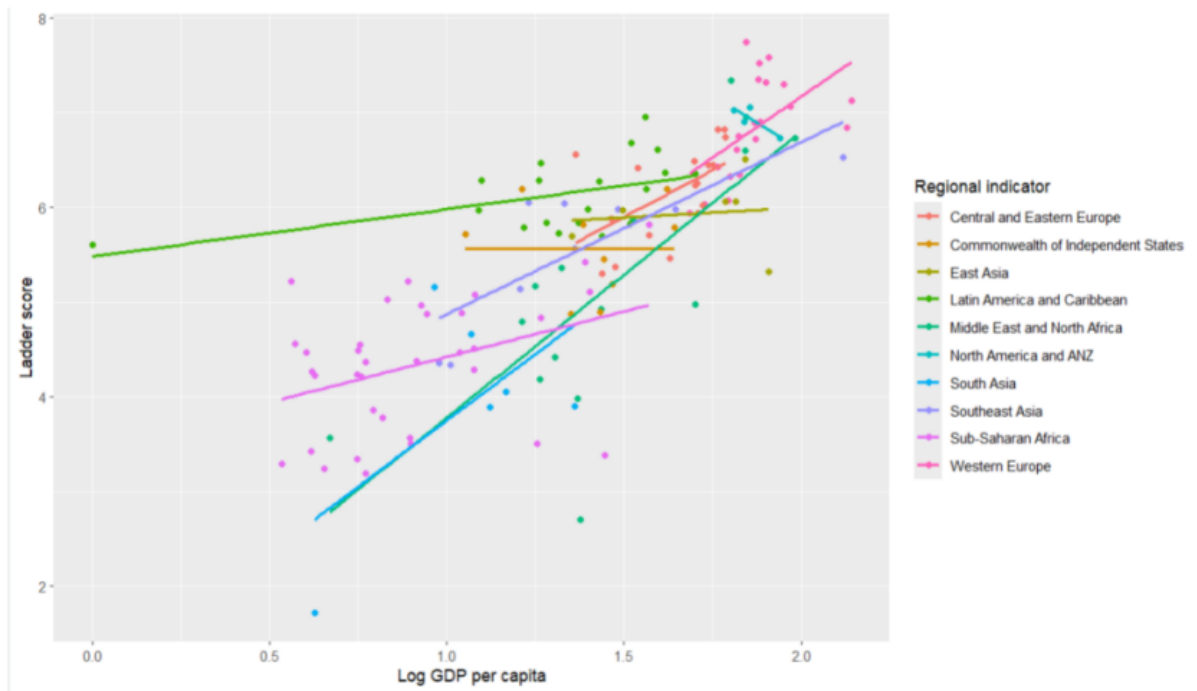
**Fig. 1** Scatter plot of Log GDP per capita vs. happiness score (with regression line).

Figure 1 illustrates the positive relationship between log GDP per capita and national happiness. As reported in Table 2, the regression results show a statistically significant positive effect of income on happiness ($p < 0.001$), with an R^2 of 0.614.

4.2. Happiness Differences between High-Income and Low-Income Countries

Building on the overall relationship, this study compares happiness levels across income groups. Countries are classified into high-income and low-income groups based on income classification standards.

Figure 2 shows that high-income countries consistently exhibit higher happiness scores. The statistical comparison results, reported in Table 3, indicate a highly significant difference in mean happiness levels between the two groups ($p < 0.001$).

This suggests that economic development plays a crucial role in shaping national happiness levels.

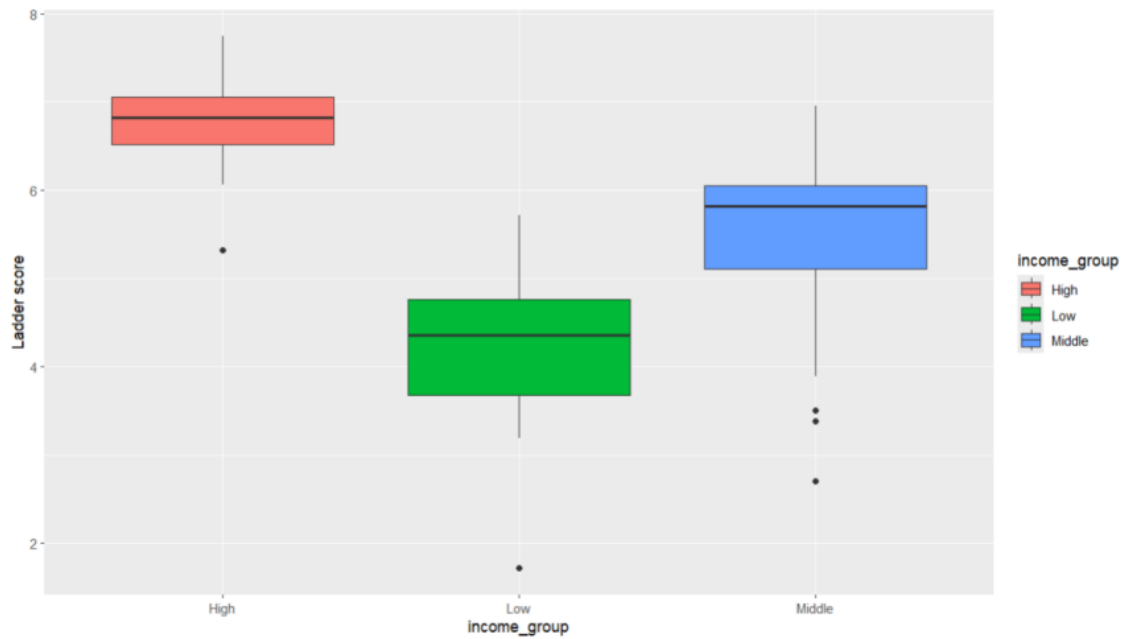


Fig. 2 Happiness score distribution by income group.

Table 3. R output for the Welch two-sample t-test between high- and low-income countries

Statistic	Value
Mean happiness (high-income)	6.785
Mean happiness (low-income)	4.257
Mean difference	2.528
t-statistic	15.856
Degrees of freedom	56.423
p-value	< 0.001
95% Confidence interval	[2.209, 2.847]

This paper further examines regional differences by computing regional averages of Life Ladder scores and Log GDP per capita. The results indicate clear regional patterns. Regions such as North America and ANZ as well as Western Europe exhibit the highest average happiness levels, while Sub-Saharan Africa and South Asia record the lowest. Overall, regions with higher levels of economic development tend to display higher happiness scores. The detailed results are summarized in Table 4.

Table 4. Regional mean happiness scores and Log GDP per capita

Region	Mean Life Ladder	Mean Log GDP
Central and Eastern Europe	5.65	10.00
Commonwealth of Independent States	5.21	9.22
East Asia	5.65	10.20
Latin America and Caribbean	6.09	9.41
Middle East and North Africa	5.37	9.71
North America and ANZ	7.21	10.80
South Asia	4.50	8.41
Sub-Saharan Africa	4.35	8.03
Southeast Asia	5.37	9.25
Western Europe	6.86	10.80

4.3. Temporal Stability of the Economic–Happiness Relationship

To examine whether the impact of economic development on happiness changes over time, a regression model including an interaction term between log GDP per capita and year is estimated.

As shown in Table 5-6 (time-interaction regression results), the interaction term is statistically insignificant, indicating that the relationship between economic development and happiness remains stable over the period from 2006 to 2023.

Table 5. Baseline regression results for the time-interaction model

Variable	Coefficient	Std. Error	t-value	p-value
Intercept	-1.669	0.119	-14.06	< 0.001
Log GDP per capita	0.762	0.013	60.77	< 0.001
Year (centered)	-0.0009	0.024	-0.04	0.969
Log GDP $\times$ Year	-0.00017	0.0025	-0.07	0.944

Table 6. R output for the time-interaction model

Statistic	Value
Observations	2335
R-squared	0.614
Adjusted R-squared	0.614
Residual Std. Error	0.695
F-statistic	1236 (p < 0.001)

To visualize group-level trends over time, Figure 3 presents average Life Ladder scores for high- and low-income countries from 2006 to 2023.

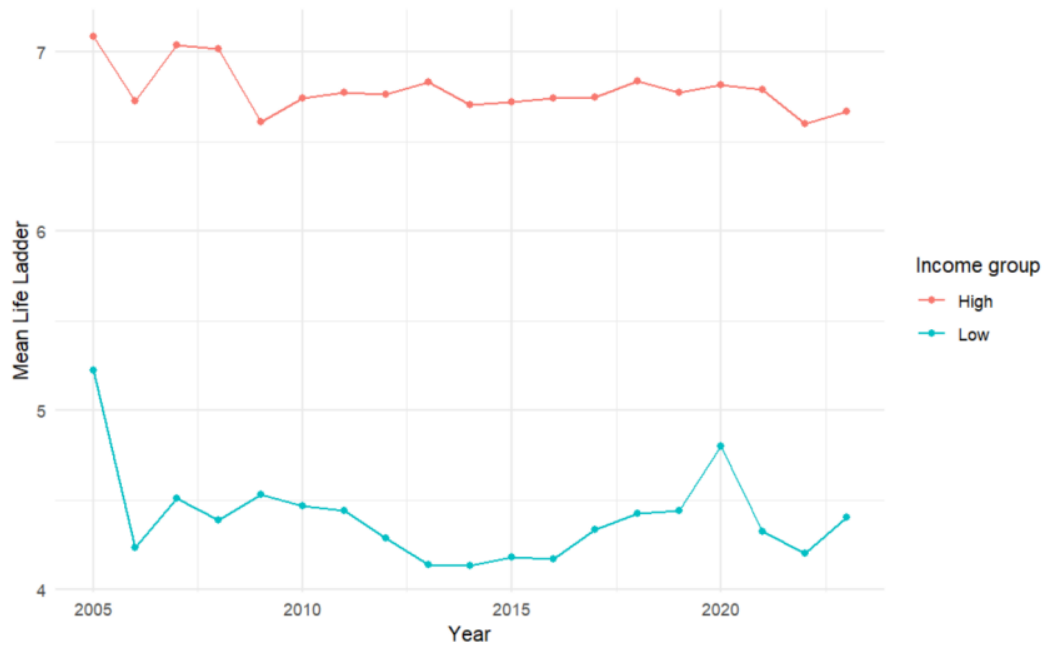


Fig. 3 Happiness trends over time for high- and low-income countries

5. Predicting Happiness Using Economic Indicators

5.1. Construction of a GDP-based Happiness Prediction Model

Based on the empirical results, a happiness prediction model is constructed using economic indicators. Predicted happiness scores are generated under different levels of GDP per capita.

The estimation results of the prediction model are reported in Table 7 (prediction regression output), showing that GDP per capita has strong predictive power for national happiness.

Table 7. Three Scheme comparing.

Variable	Coefficient	Std. Error	t-value	p-value
Intercept	-1.661	0.118	-14.04	< 0.001
Log GDP per capita	0.761	0.012	60.91	< 0.001

5.2. Predicted Happiness under Different Economic Levels

Based on the prediction model, predicted happiness levels under different economic conditions are reported in Figure 4.

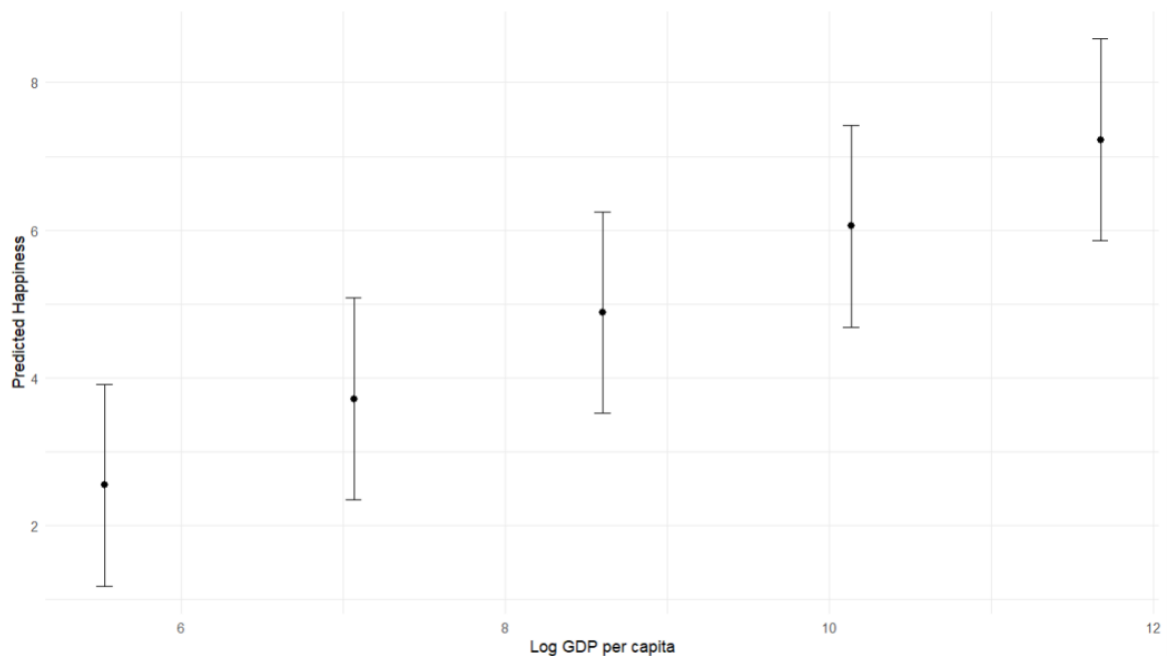


Fig. 4 Predicted happiness scores across economic levels

The results indicate a clear upward trend in predicted happiness as GDP per capita increases, while the confidence intervals reflect varying levels of uncertainty across economic stages.

6. Discussion

Although the results reveal that there is a positive correlation between economic development and national happiness, there are several drawbacks to be aware of when analyzing the results.

This paper focuses only on GDP per capita as the main explanatory variable. However, while income is an important influence on happiness it cannot be taken to represent all facets of well-being. There may be other significant elements (social support, public health system, quality of education, government effectiveness etc.) affecting the way individuals assess their lives, and such elements may differ across countries. However, as these variables are not included in the present regression specification, some of the differences in happiness between countries are still unaccounted for [8].

Second, the analysis is based on cross-country observational data. Regression analysis cannot tell us if two variables are causally related to each other, even if we know that they are statistically related. For instance, it is plausible to believe that income leads to happiness, but it is also reasonable to think that happy societies are the conditions for strong economic performance. What makes it hard to have a clear causal interpretation is that this potential two-way relation is rather prominent [9].

Third, cultural values and social norms may also play a role in how people report happiness. People in different countries may give different answers to the same survey questions, which could lead to measurement bias even in cross-country comparisons. These challenges could be tackled in

future studies in various manners. One avenue is to add more social and institutional factors, for example access to health care, education systems and governance indicators to better capture the determinants of happiness. A further avenue is the exploitation of more sophisticated panel econometric methods that enable a more thorough investigation of the dynamic interactions between economic development and well-being.

In addition, future studies may extend the dataset by including more countries or longer time periods. Combining economic indicators with social indicators may also help build more accurate models for predicting national happiness [1,10].

7. Conclusion

Using panel data of the World Happiness Report over the period of 2006~2023, this paper investigates the connection between human well-being and economic growth. Several empirical methods are used in the analysis, including correlation analysis, regression models, income-group comparison, and time-interaction analysis. They enable a better understanding of how economic instruments are related to national happiness.

The findings reveal a positive relationship between GDP per capita and happiness. Higher Life Ladder scores tend to be reported by richer countries. The differences between high and low-income countries are also marked on happiness. High income countries generally have substantially higher happiness averages than low income countries. This indicates that economic growth is a significant source of enhancement of national well-being.

The time-interaction results also confirm that the nexus between economic development and well-being is homogenous throughout the period of investigation. That is to say, the association between GDP per capita and happiness is not very volatile over time.

Besides, based on GDP per capita, we establish a rudimentary forecast model of national happiness in this paper. The output of prediction indicates that among all studied happiness determinants, economic indices can explain most of the happiness differences between countries. However, it's also true that economic variables can't fully predict happy-country results.

In sum, this study demonstrates that economic growth is an important predictor of national happiness but is by no means the only one. Social systems, health care, and other non-material factors can affect well-being as well. Future research can include more social and institutional variables to better understand the determinants of happiness.

References

- [1] J. F. Helliwell et al., World Happiness Report 2024, Sustainable Development Solutions Network, 2024.
- [2] M. A. Killingsworth, "Experienced well-being rises with income, even above \$75,000 per year," Proceedings of the National Academy of Sciences, 2021.
- [3] M. A. Killingsworth, D. Kahneman, and B. A. Mellers, "Income and emotional well-being: A conflict resolved," Proceedings of the National Academy of Sciences, 2023.
- [4] J. F. Helliwell, R. Layard, J. Sachs, and J. De Neve, World Happiness Report 2023, Sustainable Development Solutions Network, 2023.
- [5] J. F. Helliwell et al., World Happiness Report 2022, Sustainable Development Solutions Network, 2022.
- [6] C. De Neve and A. Ward, "Happiness and economic growth," Annual Review of Economics, 2022.
- [7] A. Clark, S. Fleche, R. Layard, N. Powdthavee, and G. Ward, The Origins of Happiness: The Science of Well-Being over the Life Course, Princeton University Press, 2018.
- [8] R. Layard, "Can we be happier? Evidence and ethics," Journal of Economic Behavior & Organization, 2020.
- [9] D. Asteriou and S. G. Hall, Applied Econometrics, 3rd ed., Palgrave Macmillan, 2015.
- [10] J. Sachs, L. Akinin, S. Wang, and E. Woo, "The determinants of happiness and misery," World Happiness Report 2021.