My title*

Carl Fernandes

April 19, 2024

This study investigates the connection between income inequality and national hapiness, questioning the impact of the Gini index. By examining cross-sectional data we find that income inequality has a limited direct effect on happiness. It is evident that factors like the quality of institutions and levels of development as measured by the Democracy Index and the Inequality adjusted Human Development Index (IHDI) have a more significant impact on shaping happiness. Our results indicate that focusing solely on income inequality doesn't provide a complete understanding. Instead a wider range of indicators are. This research contributes to the conversation on disparity and satisfaction by emphasizing the importance of using multifaceted approaches in happiness research and policy development.

1 Introduction

1.1 Background

Global discourse on well-being has increasingly recognized happiness as a critical indicator of societal progress. This paper embarks on an empirical inquiry into the determinants of happiness, with a particular focus on income distribution. Traditional economic theories suggest a correlation between economic prosperity and happiness, but emerging research prompts a re-evaluation of this association. In this context, income inequality, typically measured by the Gini index, emerges as a potential factor influencing happiness. Yet, the depth of its impact remains under-examined and is often overshadowed by more visible economic indicators.

---

*Code and data are available at: https://github.com/Carl-1506/Happiness-and-inequality.git

1.2 Research Question and Objectives

The primary aim of this study is to interrogate the extent to which income inequality affects national happiness levels when considered in isolation. Secondary to this is an evaluation of the interdependent effects of democratic governance and human development. The research questions guiding this analysis are: Does income inequality significantly predict happiness across nations, and if so, to what degree? Furthermore, how do variables such as the state of democracy and human development refine our understanding of this relationship? The objectives are twofold: to quantify the influence of income inequality on happiness and to discern the moderating roles of democratic governance and human development.

1.3 Theoretical Framework

Our analysis is based on a range of literature that explores the connection, between gaps and personal well being. A key focus of our framework is the research by Oishi, Kesebir and Diener (2011) who argue for a relationship between income inequality and individual happiness. They suggest that perceptions of fairness and levels of trust play a role in shaping this connection. Particularly they highlight that individuals with incomes tend to be less happy when faced with inequality not necessarily because of their own earnings but due to reduced trust and fairness perceptions. Their results indicate that how people subjectively experience inequality can be just as significant as the income gaps.

On the hand Meyer and Sullivan (2013) present a perspective on inequality compared to the mainstream discussions. Their study during the Recession reveals a contrast; while income inequality increased consumption inequality showed a different trend. This contrast prompts a look at measures of inequality stressing the importance of exploring other aspects of economic experiences like spending habits when assessing their impact, on happiness. Our research focuses on exploring if income inequality, by itself can predict happiness or if its impact depends on socio factors, like trust, fairness, consumption levels and the quality of democratic governance. The model we use aims to separate these interconnected influences emphasizing that a complete view of happiness goes beyond measures.

1.4 Paper Structure

The paper proceeds as follows: Section 2 presents a review of the pertinent literature, laying the groundwork for our empirical strategy. In Section 3, we detail the data sources and describe the variables under study. Section 4 outlines our methodology, articulating the econometric approach and data handling procedures. Section 5, the core of our empirical analysis, presents the results and explores their robustness. Section 6 discusses the implications of our findings within the broader socio-economic context, and Section 7 concludes by synthesizing the study's insights and suggesting avenues for future research.

The remainder of this paper is structured as follows. Section 3....

2 Literature Reivew

2.1 Previous Research on Happiness and Economic Inequality

The discussion about what makes people happy has grown, with researchers focusing on how economic inequality affects peoples well being. Oishi et al. (2011) shed light on this topic by showing a connection between a country's income gap and individual happiness levels in their in depth study using data from the General Social Survey. They found that feelings of fairness and trust decrease when inequality rises especially affecting those with incomes. These findings suggest that how people perceive conditions can significantly impact societal happiness.

Meyer and Sullivan (2013) add to this conversation by analyzing how consumption and income inequality change during periods of instability. Their study presents a different viewpoint; even though income inequality rose during the challenges of the early 21st century consumption inequality did not follow the same trend. This challenges the idea that only income differences influence happiness, prompting exploration into how people navigate difficulties.

Overall existing research indicates that the link between inequality and happiness is complex shaped by both measures and subjective social views. It also suggests that elements like trust in systems and perceptions of fairness are crucial, in shaping this relationship. Exploring the dynamics involved requires an investigation, into the factors that influence how economic inequality impacts our sense of happiness.

2.2 Gaps in Literature

While the current body of research provides a foundational understanding of economic inequality's impact on happiness, there remains a lacuna in quantifying the relative influence of associated socio-economic factors. Notably, the existing literature often examines income disparity without fully integrating other potential variables that may exert a more pronounced effect on subjective well-being.

Moreover, the dynamic between perceived economic fairness and societal trust as intermediaries in the inequality-happiness paradigm warrants further exploration. Studies have largely concentrated on aggregate data and broad metrics, which may overlook the micro-level nuances and the heterogeneity of individual experiences within various economic strata. Additionally, the diverse methodologies and scopes of previous research present a fragmented view, indicating a need for more harmonized and cross-culturally applicable analyses.

This study aims to fill these gaps by offering a more holistic approach, incorporating a multi-variable framework that considers not only the direct impact of income inequality but also

the roles of democratic governance and human development. Such an approach is designed to provide a more granular understanding of how income inequality interplays with other key factors to influence happiness across different national contexts.

3 Data

3.1 Data Sources

To conduct an examination of the factors we're interested, this study utilizes data from various reliable sources each contributing important metrics for our analysis.

3.1.1 Happiness Score

The happiness scores come from the World Happiness Report, a survey that assesses subjective well being. This dataset offers happiness ratings for countries giving us insights into how individuals perceive their own life satisfaction. These scores reflect a country's happiness. Are crucial for understanding the psychological impact of economic circumstances. The data range is 0-10, with 10 being the most happy.

3.1.2 Gini Index

The Gini index indicates income distribution within countries and is sourced from the database of the World Bank. This metric is significant in our analysis as it serves as a measure of income inequality with values indicating greater disparities.The index is measured as a percentage, with lower scores indicating more income equality.

3.1.3 Democracy Index Data Source

The Democracy Index, which evaluates governance and democratic freedom across countries worldwide is sourced from Our World in Datas database. It offers a perspective on systems, essential, for interpreting the broader socio political landscape where economic and well being indicators intersect. It ranges from 0-10 with 10 being the most democratic.

3.1.4 IHDI

The Inequality adjusted Human Development Index (IHDI) data is sourced from the United Nations Development Programme (UNDP). Unlike the Human Development Index (HDI) IHDI considers inequality, in distribution among the population offering a more comprehensive assessment of human development gaps. It ranges between 0 and 1 with 1 being most developed.

3.2 Sample Description

The study utilizes a dataset comprising observations from over 150 countries for the year 2022. This snapshot approach allows for a detailed examination of contemporary socio-economic conditions and their impact on happiness across diverse national contexts. Focusing on a single year enhances the consistency of data interpretation and avoids the complexities associated with longitudinal trend analysis.

This dataset's geographic and temporal scope provides a robust basis for evaluating the relationships between happiness, economic disparity, democratic governance, and human development, offering insights into how these dynamics manifest in the modern world.

3.3 Descriptive Statistics

3.3.1 Happiness Score Distribution

In Figure 1 we see how happiness scores are spread out in the sample. The graph reveals that there are peaks indicating common levels of reported happiness. Most countries fall in the mid-range of scores suggesting that a moderate level of self reported well being is most common. Notably few countries report high or low happiness levels showing that globally such extreme ratings are rare.

3.3.2 Gini Index Distribution

The distribution of Gini index values shown in Figure 2 displays a pattern indicating that most countries have lower levels of inequality and fewer have high inequality. This skewness suggests that although income inequality exists extreme cases are not as prevalent.

3.3.3 Relationship Between Happiness and Gini Index

In Figure 3 there's a scatter plot comparing happiness scores to Gini index values. The plot doesn't show a persistent pattern or strong linear relationship implying that factors other than income inequality may also influence happiness.

3.3.4 Relationship Between Happiness and IHDI

Lastly Figure 4 illustrates the connection between happiness scores and IHDI. A positive trend is visible with higher IHDI values corresponding to happiness scores. This suggests a link where countries with equitable human development tend to have happier citizens.

3.3.5 Discussion

The thorough analysis lays the foundation for an investigation through methods. From the visuals it appears that although income inequality, measured by the Gini index is evident in the data, its impact on happiness is not easily defined and probably intertwined with other socio-economic aspects. This theory will be examined further in the multivariate analysis.

4 Model

This research utilizes two approaches to understand the connection between happiness and various socio-economic factors. The main objective is to examine how income inequality, represented by the Gini index along with democratic governance and human development, relate to self reported happiness levels. These models aim to offer a perspective considering both indirect impacts of these factors on a countrys overall well being.

The initial model is a linear regression that suggests happiness scores are directly linked to the Gini index in a linear manner. This model acts as a foundation for understanding the association between income inequality and happiness without considering variables.

The second model is a regression that builds upon the first by introducing additional predictors; Democracy index and IHDI. This model enables us to investigate how income distribution, democratic governance and human development collectively influence happiness presenting a examination of the socio-economic dynamics involved.

4.1 Model set-up

4.1.1 Simple Linear Model

The initial model considers the Gini index as the sole predictor. This setup allows for an assessment of how variations in income inequality might align with changes in happiness scores. It is represented by the equation below where α is the intercept, β is the slope coefficient for the Gini index, and ϵ_i is the error term.

$$HappinessScore_i = \alpha + \beta \times GiniIndex_i + \epsilon_i$$
$$\epsilon_i \sim N(0, \sigma^2)$$

4.1.2 Multi-Linear Model

Expanding on the initial model, the second regression introduces additional explanatory variables and is represented below. In this model, α represents the intercept, β_1, β_2, and β_3 are the coefficients for Democracy index, Gini index, and IHDI, respectively, while ϵ denotes the error term.

$$\text{Happiness\_score}_i = \alpha + \beta_1 \times \text{Democracy\_index}_i + \beta_2 \times \text{Gini\_index}_i + \beta_3 \times \text{IHDI}_i + \varepsilon_i \quad (1)$$
$$\varepsilon \sim \text{Normal}(0, \sigma^2) \quad (2)$$

Both models are estimated using ordinary least squares (OLS), which minimizes the sum of squared differences between the observed and predicted values. This method was chosen for its robustness and simplicity, facilitating the interpretation of coefficient estimates.

The data are prepared for analysis by ensuring no multicollinearity among predictors, normality of residuals, homoscedasticity, and independence of observations—key assumptions for the validity of OLS regression. Violations of these assumptions are addressed through appropriate transformations and robust standard error calculations where necessary. For more information on how this is done, please refer to Section B.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

4.2 Model justification

The selection of linear regression as the analytical method in this study is justified by its suitability for quantifying relationships between a continuous dependent variable and one or more independent variables. The first model, examining the influence of income inequality on happiness, utilizes a single predictor, the Gini index, to isolate its effect. This model serves as a point of reference for understanding the baseline impact of inequality on national happiness levels without additional factors.

In contrast, the second model's inclusion of the Democracy index and IHDI variables alongside the Gini index allows for a deeper investigation into how happiness correlates with a broader set of economic and socio-political variables. This approach is informed by the theoretical foundation suggesting that happiness is a multifaceted construct influenced by various aspects of societal development and governance.

Both models are grounded in empirical evidence from prior research indicating potential links between these variables and happiness. By employing OLS regression, we aim to build on these findings, providing clear, interpretable coefficients that measure the strength and direction of these relationships. The simplicity and transparency of OLS facilitate the communication of

results to a broad audience, which aligns with the study's objective of contributing to the policy discourse on enhancing societal well-being.

5 Results

5.1 Baseline Model

The baseline results are summaried in Table 1. Findings show that as income inequality increases happiness scores tend to decrease with a negative coefficient ($\beta = -0.043$) and a p value of 0.001 supporting this trend. This means a 1% increase in inequality is associated with a 0.043 drop in the happiness score of a country. The models intercept is positive ($\alpha = 7.263$), indicating the happiness score in the absence of income inequality. However with an R squared value of 0.103, 10.3% of the variations in happiness scores can be explained by the Gini index alone, hinting at other contributing factors beyond this models scope. The models credibility is bolstered by a F statistic (F = 11.489) affirming the significance of the regression equation. Additionally an RMSE of 1.00 suggests how much happiness scores deviate from predicted values by the model on average.

While it aligns with expectations that higher income inequality corresponds to lower reported happiness levels, the limited explanatory power emphasizes the necessity for a more comprehensive model incorporating additional variables. The upcoming parts will expand on this analysis by incorporating socio-economic factors into the regression model.

5.2 Extended Model

The extended model combines the Democracy index and IHDI with the Gini index to explore how they collectively influence happiness levels. The findings reveal a connection between IHDI and happiness scores ($\beta_3 = 4.112$, p < 0.001) indicating that nations with higher human development levels, adjusted for inequality show higher happiness among their populace.

It's worth noting that the Democracy index also demonstrates a positive impact ($\beta_1 = 0.096$, p = 0.019) suggesting that democratic governance is linked to increased happiness scores. This aligns with the expected theory that democratic societies offering freedoms and citizen engagement can enhance well being.

In this model the Gini index coefficient is positive ($\beta_2 = 0.013$) but not statistically significant (p = 0.124) differing from the base model results. This change emphasizes the importance of considering socio-economic factors when assessing national welfare.

The models explanatory capability has been notably improved, with an R squared value of 0.735 showing that 73.5% of happiness score variations are accounted for by the model. The adjusted R squared value of 0.727 considers predictor variables. Confirms the models reliability.

The decrease in RMSE to 0.53 when compared to the model indicates a predictive capability of this expanded analysis.

5.3 Diagnostic checks

The Q-Q plot for the residuals of the second model in Figure 5 exhibits a slight deviation from normality, particularly in the tails, suggesting the presence of outliers or heavy tails in the data distribution. This deviation from the expected straight line in the Q-Q plot could imply that the residuals do not follow a perfect normal distribution, which is an assumption for the OLS regression. However, the majority of the points do align well with the theoretical quantiles, indicating that any deviations from normality are not extreme.

Moreover, this departure from normality does not necessarily invalidate the model but does suggest that predictions may be less reliable at the extremities of the data range. It might be beneficial to explore whether a transformation of the response variable or the use of robust regression techniques could yield a more homoscedastic and normally distributed set of residuals. The model could also be enhanced by investigating potential leverage points or influential observations that may disproportionately impact the regression results.

Despite these considerations, the central mass of the data adheres closely to the line, reinforcing confidence in the model's findings for the majority of the data. Consequently, while the Q-Q plot indicates minor concerns, the overall diagnostic evaluation suggests that the model provides a generally reliable representation of the data.

6 Discussion

6.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

6.2 Second discussion point

6.3 Third discussion point

6.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

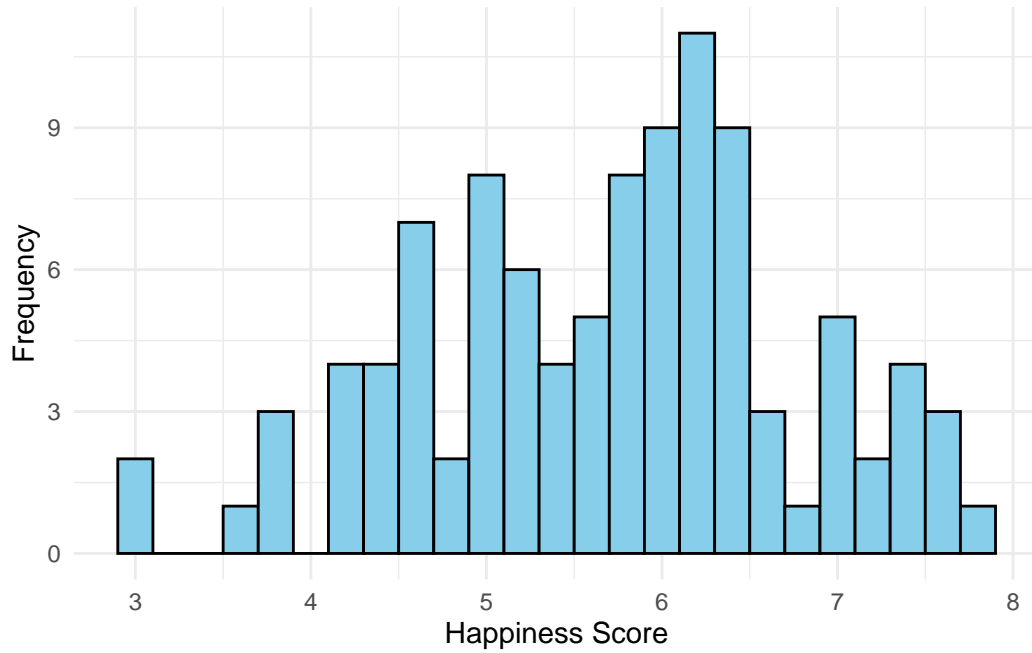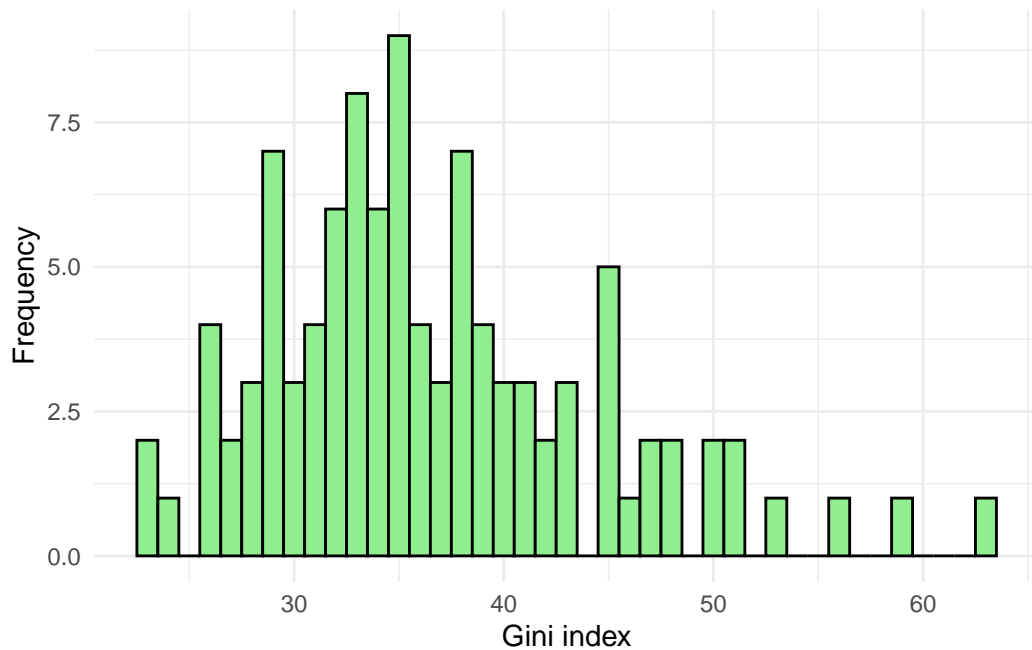## A Tables and figures

Figure 1: Distribution of Happiness Scores

Figure 2: Distribution of Gini Index
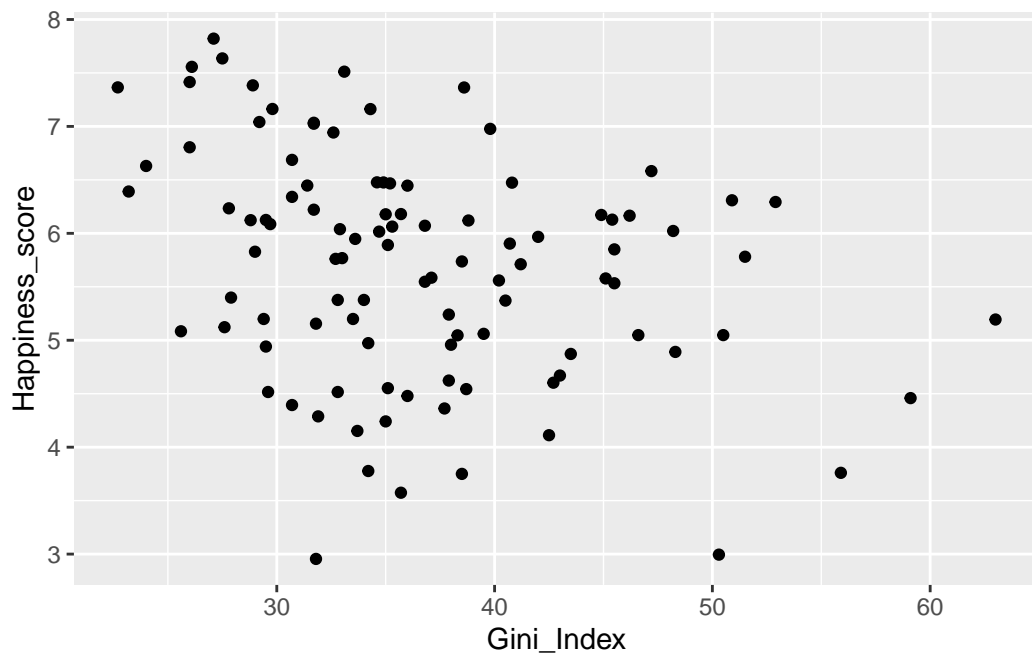
Figure 3: Relationship between Gini Index and Happiness

Table 1: Linear Regression

	First model
(Intercept)	7.263
	(0.476)
	(15.265)
	(<0.001)
	[6.319, 8.207]
Gini_Index	−0.043
	(0.013)
	(−3.390)
	(0.001)
	[−0.068, −0.018]
Num.Obs.	102
R2	0.103
R2 Adj.	0.094
AIC	295.2
BIC	303.1
Log.Lik.	−144.613
F	11.489
RMSE	1.00

Table 2: Multiple Linear Regression

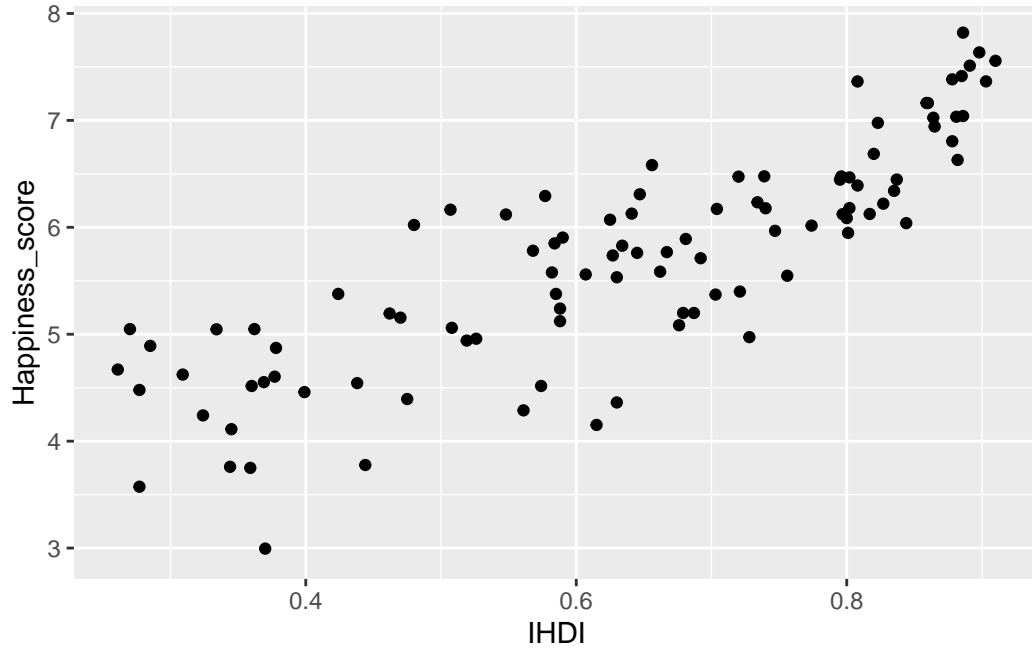|	Second model
(Intercept)	2.016
	(0.468)
	(4.310)
	(<0.001)
	[1.087, 2.944]
Democracy_index	0.096
	(0.040)
	(2.385)
	(0.019)
	[0.016, 0.176]
Gini_Index	0.013
	(0.009)
	(1.551)
	(0.124)
	[−0.004, 0.030]
IHDI	4.112
	(0.515)
	(7.987)
	(<0.001)
	[3.090, 5.134]
Num.Obs.	100
R2	0.735
R2 Adj.	0.727
AIC	166.6
BIC	179.6
Log.Lik.	−78.311
F	88.755
RMSE	0.53

Figure 4: Relationship between IHDI and Happiness

B Model details

In preparation for the OLS regression, the dataset was meticulously inspected for multi-collinearity to prevent inflated variance in the coefficient estimates. The IHDI, adjusted for income disparities, was chosen specifically to avoid overlap with the Gini index. The assumptions of residual normality, constant variance across predictions (homoscedasticity), and the independence of observations were each tested. Any deviations from these conditions prompted corrective measures such as variable transformations to achieve normality and homoscedasticity, and the use of robust standard errors to counteract potential heteroscedasticity, ensuring the robustness of our regression results.
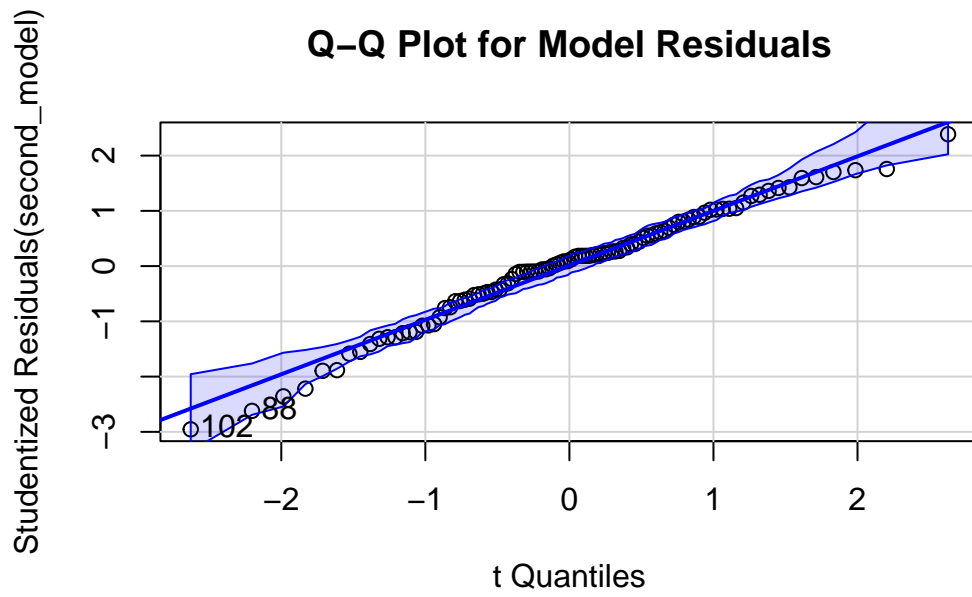
B.1 Diagnostics

```
[1]  88 102
```

Figure 5

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.