
Can money buy happiness?

Project Report: Data Analysis and Modeling for Public Affairs - Spring 2019

Akshay Rathii || Agastya Teja || Devanshi Mittal || Linh Truong

Introduction to the Problem

Irrespective of where we come from, happiness is one of the most important virtues we all aim to possess in life. We also, almost always, hear that money can't buy happiness, but we unanimously agree that money is important. Our project aims at analyzing happiness index panel data from 85 countries over the years 2009 to 2018 and exploring the relationship between happiness and various other factors, including GDP per capita (purchasing power parity) of the country. Happiness is also significant from a government point of view, since happy citizens means they are more productive and in turn aid in development of the country.

Null Hypothesis (H_0) – Increase in GDP per capita does not increase the happiness of that country.

Here, we are interested in “happiness”, which is driven by a variety of factors, so to draw meaningful causal inferences, we need a way to isolate the effects of each factor. Regression analysis helps us understand how the value of our dependent variable (happiness) changes when one of the independent variables (money) is varied, while all the other independent variables are held fixed. The regression analysis is the best statistical tool here, since it allows us the flexibility to model multiple independent variables, include continuous/categorical variables and assess interaction terms to determine whether the effect of one independent variable depends on the value of another.

The Model

Our study includes panel data from 85 countries over the years 2009 to 2018, with Happiness Index being our dependent variable and GDP per capita being our independent variable in study. Since, the GDP per capita for all the countries together have a huge variance, we get skewed data and hence, apply a log transformation. Thus, we end up studying the relationship of log(GDP per capita) with happiness. The data has been sourced from the World Happiness Report Data - Helliwell, J., Layard, R., & Sachs, J. (2017). World Happiness Report 2017, New York: Sustainable Development Solutions Network - <https://data.world/laurel/world-happiness-report-data>.

The independent variable – “happiness” has been measured through the Gallup World Poll with a typical sample size of 1000+ people in each country comprising of the following question – “Please imagine a ladder with steps numbered from 0 at the bottom to 10 at the top. Suppose we say that the top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you.”

Different measures have been employed in calculating the independent variables as follows –

Predictor	How was the predictor measured?
GDP per capita	Purchasing Power Parity (PPP) adjusted to constant 2011 international dollars, taken from the World Development Indicators (WDI).
Healthy Life Expectancy at Birth	Constructed based on data from the World Health Organization (WHO) Global Health Observatory data repository.
Social Support	National average of the binary responses (either 0 or 1) to the Gallup World Poll (GWP) question “If you were in trouble, do

	you have relatives or friends you can count on to help you whenever you need them, or not?"
Freedom to make life choices	National average of binary responses to the GWP question "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?"
Generosity	Residual of regressing the national average of GWP responses to the question "Have you donated money to a charity in the past month?" on GDP per capita.
Perceptions of Corruption	Average of two GWP questions: "Is corruption widespread throughout the govt or not?" and "Is corruption widespread within businesses or not?"
Confidence in national govt.	Scale: 1 to 5 (No opinion – None at all – Not very much – Fair amount – Great deal)

Table 1: Data Description

Before implementing the model, during our data analysis we encountered a specification error with GDP per capita (non-linearity with happiness index) and hence log transformed the same to maintain linearity between the independent and dependent variable. Another potential specification error might occur due to omitted variable bias – since happiness is a very subjective term, it cannot be accurately measured by survey polls and it depends on many other factors than the ones listed above. However, many of these variables might be really difficult to assess. An added potential specification error might occur due to multi-collinearity between perceptions of corruption and confidence in national government, since the two intuitively seem to be highly correlated. Panel data also hints potential heteroskedasticity.

Since our data is panel data, we have fixed variables of countries and years whose values do not change across time. Also, after comparing the results from Pooled OLS, One-way fixed effect and Two-way fixed effects model and running the nested F-test, we prefer the two-way fixed effects model –

lm (Happiness ~ Independent Variables + as.factor(year) + as.factor(country))

$$Y = \beta_0 + \beta_1 * \text{LogGDPpercapita} + \beta_2 * \text{Socialsupport} + \beta_3 * \text{Healthylifeexpectancyatbirth} + \\ \beta_4 * \text{Freedomtomakelifechoices} + \beta_5 * \text{Generosity} + \beta_6 * \text{Perceptionsofcorruption} + \\ \beta_7 * \text{Confidenceinnationalgovt} + \beta_8 * \text{as.factor(Year)} + \beta_9 * \text{as.factor(Country)} + e \text{ (error terms)}$$

With fixed effects model, we do not estimate the effects of variables whose values do not change across time. Rather, we control for them.

Discussion of similar relevant statistical analysis –

Abadi et al. (2018) implemented study in Indonesia to evaluate factors on life happiness by using multiple linear regression. The result shows that health, safety, family household harmony, leisure time, home and amenities and affection have significant impact on life happiness while education, employment, income, social relation, transportation and environmental condition do not have relation to happiness. These results provide a contrast to our findings and we believe that the authors should include fixed effect variable in the model to capture time-invariance observed and unobserved characteristics within unit fixed. In addition, the limitation of the sample size and scope could cause a threat to external validity of the research.

https://www.researchgate.net/publication/329787956_Factors_Affecting_of_Lifes_Happiness

Howel et al. (2008) studied the relationship between subjective well-being and economic status in developing countries. In this study, the authors increase strength of internal validity by including additional covariates and using fixed effect techniques in the model. They found that money has weak effect on happiness, however, the effect size is larger in low-income people in developing countries.

<https://www.ncbi.nlm.nih.gov/pubmed/18605819>

Regression Analysis Results

Independent Variables	Coefficient	P - values
Log (GDP per capita)	1.196	3.722*10-9***
Social Support	-0.039	0.00015***
Health Life Expectancy at Birth	1.235	0.088
Freedom to Make Life Choices	0.790	0.00032***
Generosity	0.474	0.0138***
Perceptions of Corruption	-0.417	0.1316
Confidence in National Government	0.419	0.016***

Table 2: Regression Analysis Results

Interpretation done at **95 percent** confidence interval –

Adj. R²: 0.91 – Our model explains 91 percent of the variance in the data.

Log(GDP per capita): This variable is statistically significant and since this is a linear-log regression, we say that 1% increase in “GDP per capita” leads to a 0.01196 unit increase in “Happiness” on an average, while all other variables are held constant.

Social Support: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Social Support” leads to a 1.2353 unit increase in “Happiness” on an average, while all other variables are held constant.

Health Life Expectancy at Birth: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Healthy Life expectancy at Birth” leads to a 0.039 unit decrease in “Happiness” on an average, while all other variables are held constant. This is in contrast to our intuitive conclusion, however on deeper understanding, we realized that the cultural opinions in Latin American countries causes this data to be a bit skewed.

Freedom to Make Life Choices: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Freedom to Make Life Choices” leads to a 0.79 unit increase in “Happiness” on an average, while all other variables are held constant.

Generosity: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Generosity” leads to a 0.474 unit increase in “Happiness” on an average, while all other variables are held constant.

Perceptions of Corruption: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Perceptions of Corruption” leads to a 0.417 unit decrease in “Happiness” on an average, while all other variables are held constant. This is contrary to our belief, however, we could not find a suitable explanation to justify this effect.

Confidence in National Government: This variable is statistically significant and can be interpreted as – every 1 unit increase in “Confidence in National Government” leads to a 0.419 unit increase in “Happiness” on an average, while all other variables are held constant.

Assumptions Check:

We check for outliers first and realize that we encounter this issue on account of missing data for specific countries over the years. Thus, we drop these countries and only include 85 countries between the years 2009 to 2018 while implementing our model.

We then begin by measuring the “vif” scores and find multi-collinearity between log GDP per capita and healthy life expectancy at birth, their vif scores being greater than 10. However, we believe that these two variables are of utmost importance in measuring happiness and dropping them would create an omitted variable bias which is worse than multi-collinearity.

We then employ the Breusch-Pagan test and conclude that heteroscedasticity exists in our data. Then, we include robust standard errors (Huber-White Standard Errors). Since, we are working with panel data, heteroskedasticity is not a problem.

Conclusion

The Log GDP per capita is a statistically significant variable and leads us to reject the null hypothesis – “Increase in GDP per capita does not increase the happiness of that country” and conclude that money can indeed buy happiness.

In the graphs plotted below, it can be seen that the United States has a high happiness index which is a result of having a high social support factor of 0.9 and a high log GDP per capita.

On comparing the slopes of Log GDP per capita and Social Support, it can be understood that the slope of Social Support v/s Happiness is higher than that of Log GDP per Capita v/s Happiness. This helps us conclude that to increase happiness, it is more important to invest in community building programs that help improve the social support factor as compared to investing in improving the GDP per capita of a country.

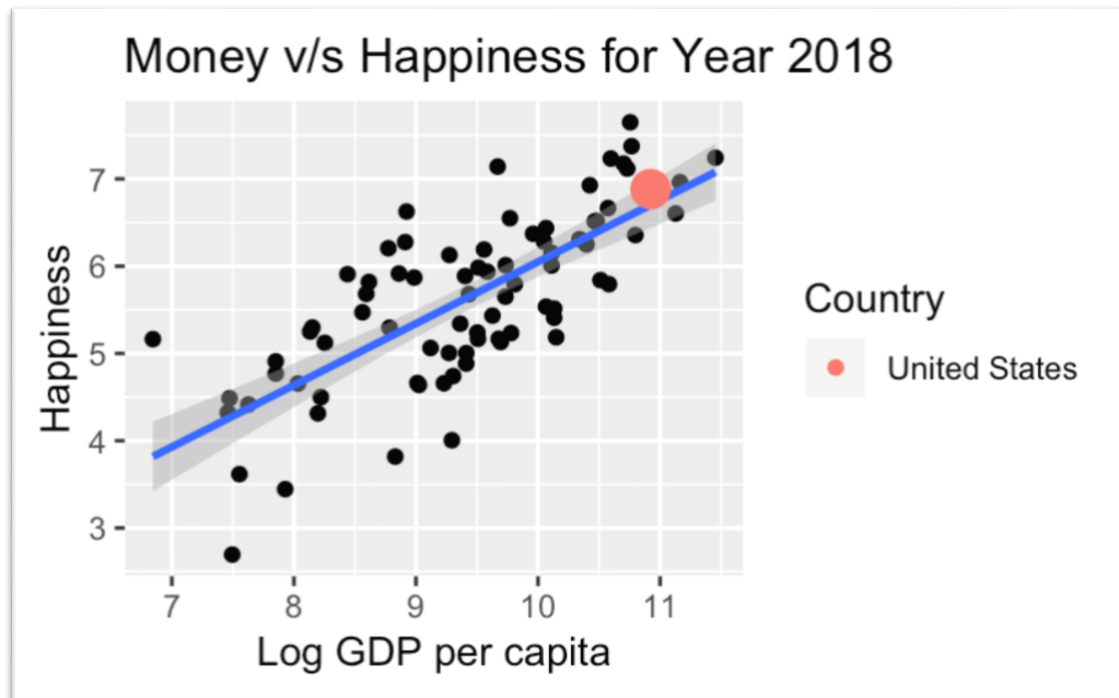


Diagram 1: Happiness increases as Log GDP per capita increases

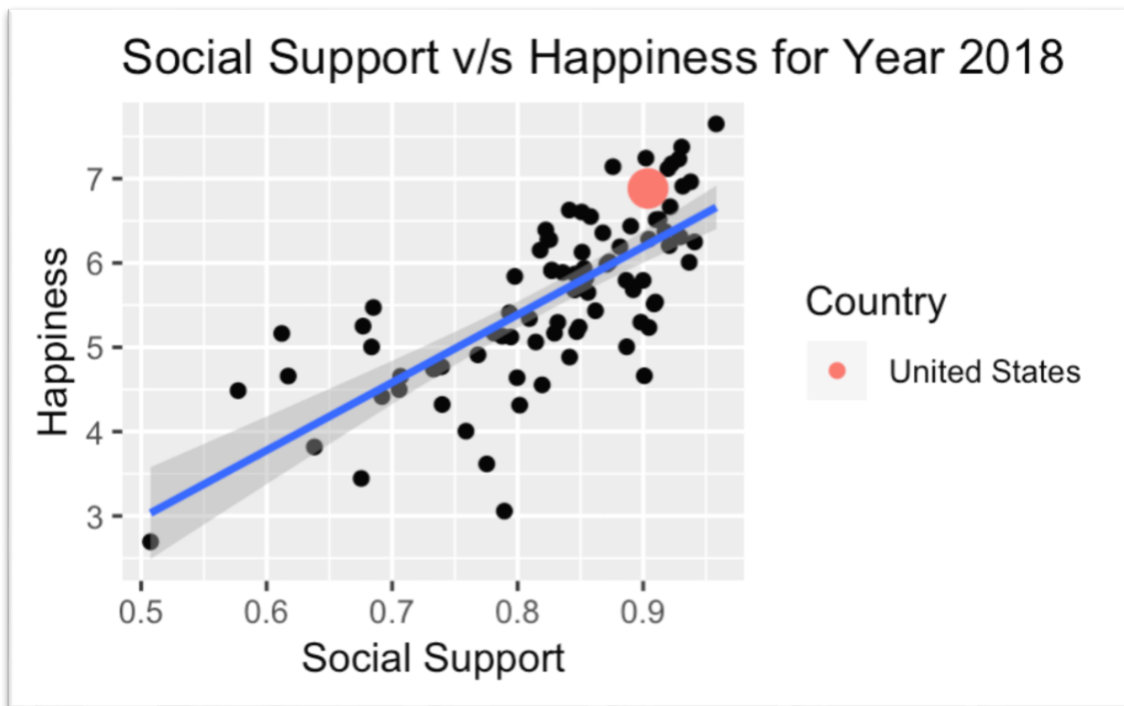


Diagram 2: Happiness increases as Social Support increases

Apart from money and social support, we can help promote happiness by being more generous (donations to charity, volunteering for a social cause), investing more time and money in maintaining good personal health and hygiene.

Practical significance of the analysis and suggestions for policy makers include –

Working on policies that help build a strong community, in turn improving the social support factor for its citizens. The freedom to make life choices positively affects the happiness index and contributes significantly and can be improved by providing right to life and liberty, freedom from slavery and torture, freedom of opinion and expression, the right to work and education, and many more. Investing in health and education is essential for a country to have a high GDP. Also, since health expectancy and GDP per capita have a high correlation, investing to improve one factor will directly affect the other factor. Governments must work with constant feedback from its citizens to develop a positive rapport and confidence among them.

Additional research would include, but not be limited to – assessing the psychological aspects of happiness and developing ways to measure them. Factors such as freedom to make life choices are highly subjective and vary from person to person. Happiness is highly dependent on the time frame one is measuring it in. Same people at different points of time (age) may have different virtues of happiness. These factors also need to be considered and there is a need to devise ways to include them in such kind of regression (statistical) analysis.

Technical Appendix

Exploring the data –



Diagram 3: Frequency of happiness scores worldwide in 2018

Multi-collinearity test –

Predictors	VIF Scores
Log GDP per capita	18.124777
Healthy Life Expectancy at Birth	13.137312
Social Support	2.951401
Freedom to make life choices	2.641130
Generosity	2.666562
Perceptions of Corruption	3.874899
Confidence in national govt.	2.757005
Country (as.factor)	1.118087
Year (as.factor)	1.099308

Table 3: VIF scores (Multi-collinearity)

Breusch-Pagan test –

BP	df	p-value
194.26	96	1.229-08

Table 4: pBtest (Heteroskedasticity)

Result for Hub-White standard errors –

Coefficients	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.950872	2.303887	-1.715	0.08685 .
LogGDPpercapita	1.196136	0.266499	4.488	8.5e-06 ***
Socialsupport	1.235379	0.426845	2.894	0.00393 **
health	-0.038786	0.026446	-1.467	0.14296
Freedomtomakelifecoices	0.789764	0.282370	2.797	0.00531 **
Generosity	0.473688	0.204468	2.317	0.02083 *
Perceptionsofcorruption	-0.416568	0.331668	-1.256	0.20958
Confidenceinnationalgovernment	0.418673	0.207727	2.015	0.04427 *
Residual standard error: 0.3157 on 645 degrees of freedom (108 observations deleted due to missingness) Multiple R-squared: 0.9228, Adjusted R-squared: 0.9114 F-statistic: 195.1 on 96 and 645 DF, p-value: < 2.2e-16 Note: Heteroscedasticity-consistent standard errors using adjustment hc3				

Table 5: Hub-White Robust Standard Errors

References –

<https://ourworldindata.org/happiness-and-life-satisfaction>

<https://worldhappiness.report/ed/2019/>