Investigating the Relationship Between Economic Freedom and Human Development: A Canonical Correlation Analysis

Bruce Nguyen
793663

1 Introduction

Economic freedom, or economic liberty, is the ability of people of a society to control their labor and property. [1]. Individuals in an economically free society are free to work, produce, consume, and invest as they see fit. Governments in economically free societies allow labor, capital, and goods to move freely and refrain from coercion or restriction of liberty beyond what is required for ensuring liberty itself.

Economic freedom has been shown to not only brings about prosperity but also human development in terms of health, education, and more [2]. For example, economically freer countries fare much better in health care quality indicators such as infantile mortality rate, life expectancy, access to clean water and sanitation, etc. They also tend to be happier in general compares to those with less economic freedom. Such countries, perhaps surprisingly, have also been shown to pay more attention to ensuring their environmental quality [3]. The list goes on. Thus, the question is not whether economic freedom is a meaningful concept, but how it affects other aspects of societies.

In this statistical analysis, we aim to analyze the relationship between the factors that contribute to economic freedom and those which determine the human development. Our research question is: how do the variables that make up economic freedom correlate with those that determine human development in countries? In the next section, we will showcase the data used by the analysis. Section 3 will demonstrate the analysis itself: the univariate, bivariate, and ultimately the canonical correlation analyses. In the end, the results will be shown, and conclusions be made.

2 Data Acquisition

To obtain the assessment of the economic freedom of nations, we turn to the Index of Economic Freedom by the Heritage foundation. The index is measured following four principal components of economic freedom:

1. Rule of Law: Property Rights, Government Integrity, Judicial Effectiveness
2. Government Size: Government Spending, Tax Burden, Fiscal Health
3. Regulatory Efficiency: Business Freedom, Labor Freedom, Monetary Freedom
4. Open Markets: Trade Freedom, Investment Freedom, Financial Freedom

Thus, there are 12 variables in total. Each is graded on a scale of 0 to 100. Regarding the measure of human development, we obtained the dataset that made up the Human Development Index (HDI) created by the United Nations. The dimensions of the measurement are:

1. Long and healthy life: Life Expectancy at Birth
2. Knowledge: Expected Years of Schooling, Mean Years of Schooling
3. A Decent Standard of Living: GNI per capita (PPP $)

It follows that there are 4 variables in total. Each with different scale. All the data collected was from 2019, as we find that this is the most recent year where the component data from both indices are available. After processing the data and performing imputations, we are left with 178 data points representing 178 countries.

3 Data Analysis

3.1 Univariate analysis

Descriptive statistics shows nothing out of the ordinary with the data. Most of the average scores across the economic freedom variables lie between 40-70.

	Property Rights	Government Integrity	Judicial Effectiveness	Tax Burden	Government Spending	Fiscal Health
count	178.000000	178.000000	178.000000	178.000000	178.000000	178.000000
mean	52.464607	41.865169	45.250000	77.611966	64.257303	66.160112
std	19.646700	19.729173	17.778157	11.894733	22.764870	31.333981
min	7.600000	7.900000	10.000000	42.000000	0.000000	0.000000
25%	37.200000	28.100000	31.075000	71.050000	51.500000	40.775000
50%	50.350000	36.400000	42.900000	78.350000	68.700000	80.150000
75%	65.900000	50.300000	54.675000	85.375000	81.750000	90.050000
max	97.400000	96.700000	92.400000	99.800000	96.600000	100.000000

Figure 1: Univariate Descriptive Statistics (Economic Freedom) (1)

	Business Freedom	Labor Freedom	Monetary Freedom	Trade Freedom	Investment Freedom	Financial Freedom
count	178.000000	178.000000	178.000000	178.000000	178.000000	178.000000
mean	63.946067	59.805056	75.532022	74.644045	57.528090	48.527753
std	15.067555	13.974105	9.662288	10.782074	21.922096	19.249450
min	17.700000	20.000000	0.000000	45.000000	0.000000	10.000000
25%	54.625000	50.800000	72.000000	66.650000	45.000000	30.000000
50%	65.150000	60.000000	77.800000	76.100000	60.000000	50.000000
75%	75.075000	69.075000	81.775000	83.950000	75.000000	60.000000
max	96.400000	91.000000	88.000000	95.000000	95.000000	90.000000

Figure 2: Univariate Descriptive Statistics (Economic Freedom) (2)

	Life expectancy at birth	Expected years of schooling	Mean years of schooling	Gross national income (GNI) per capita
count	178.000000	178.000000	178.000000	178.000000
mean	72.658427	13.370787	8.705618	19807.522472
std	7.462218	2.923457	3.120059	19883.001253
min	53.300000	5.000000	1.600000	754.000000
25%	67.150000	11.425000	6.325000	4875.500000
50%	74.000000	13.150000	8.950000	12577.000000
75%	77.975000	15.200000	11.375000	29542.750000
max	84.900000	22.000000	14.200000	92418.000000

Figure 3: Univariate Descriptive Statistics (Human Development)

We can better visualize the distribution of the variables by plotting their histograms with the Kernel Density Estimation (KDE) superimposed.

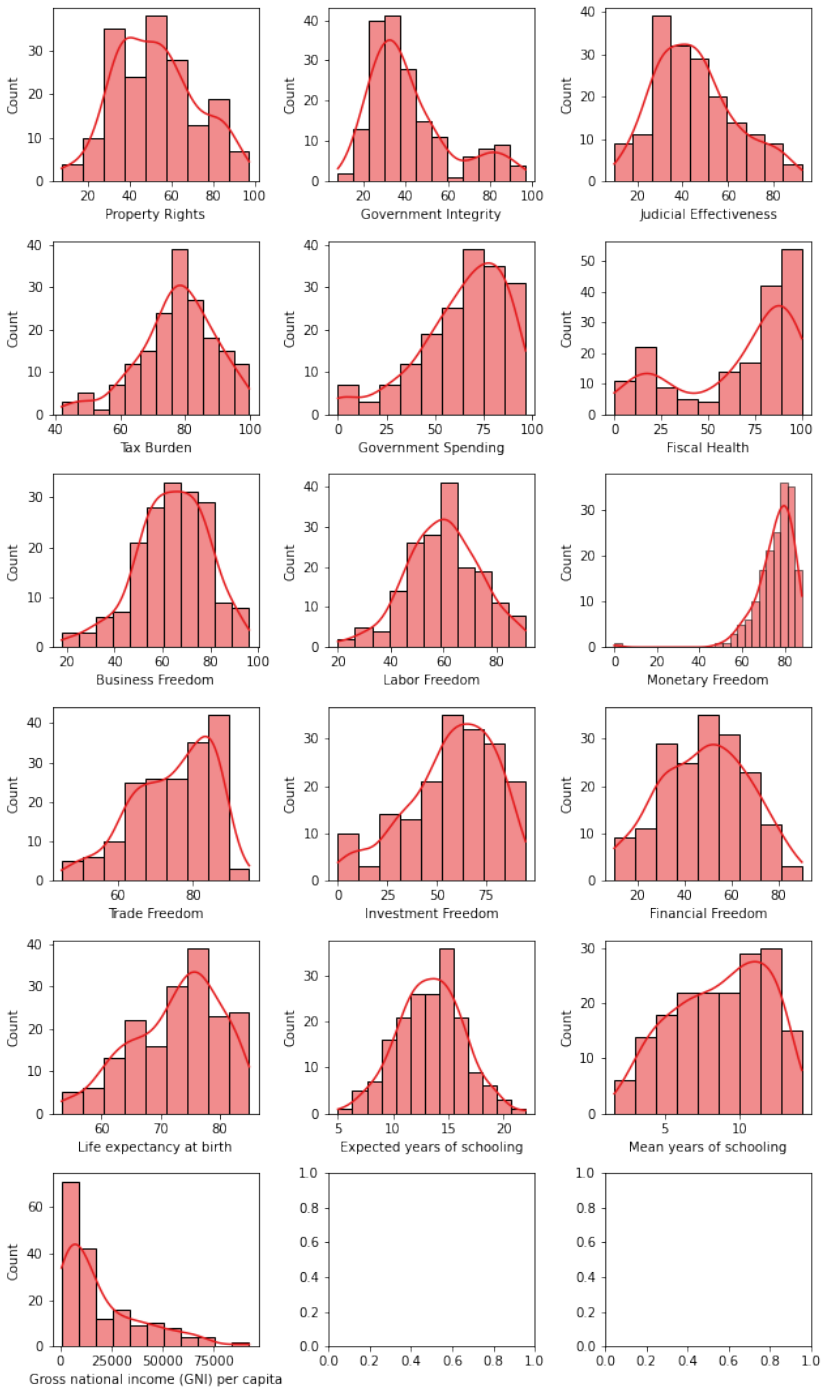

Figure 4: Histograms with Superimposed KDE

It can be seen that most of the variables are not distributed normally. Those with skewed-right distributions are: `Property Rights`, `Government Integrity` and `Judicial Effectiveness` (index 1, 2, 3) . This simply means the majority of countries did not score high along these measures. Most of the rest of the variables saw themselves skewed-left distributed, which means the opposite in the context of scores, except for `Labor Freedom` (index 8) and `Expected Years of Schooling` (index 14), where their KDE plot very much resemble the normal distribution.

Regarding `Fiscal Health`, `Government Integrity` and, to some degrees, `Investment Freedom` (index 6, 2, 11), their distributions seem to be multi-modal. This means that we can form distinct groups of countries according to how they perform on these indices.

3.2 Bivariate Analysis

We conduct the bivariate analysis by creating a scatter plot for each pair of variables. Within each plot, the countries are color-coded with its World Bank income classification based on `GNI per capita` (index 16). Red represents countries with low income, blue lower-middle income, green upper-middle income, and purple high income. Due to the sheer number of variables, the visualization might be difficult to make sense of within the margin of this document. Nevertheless, we can still observe the overall interactions between the different variables from the plots.

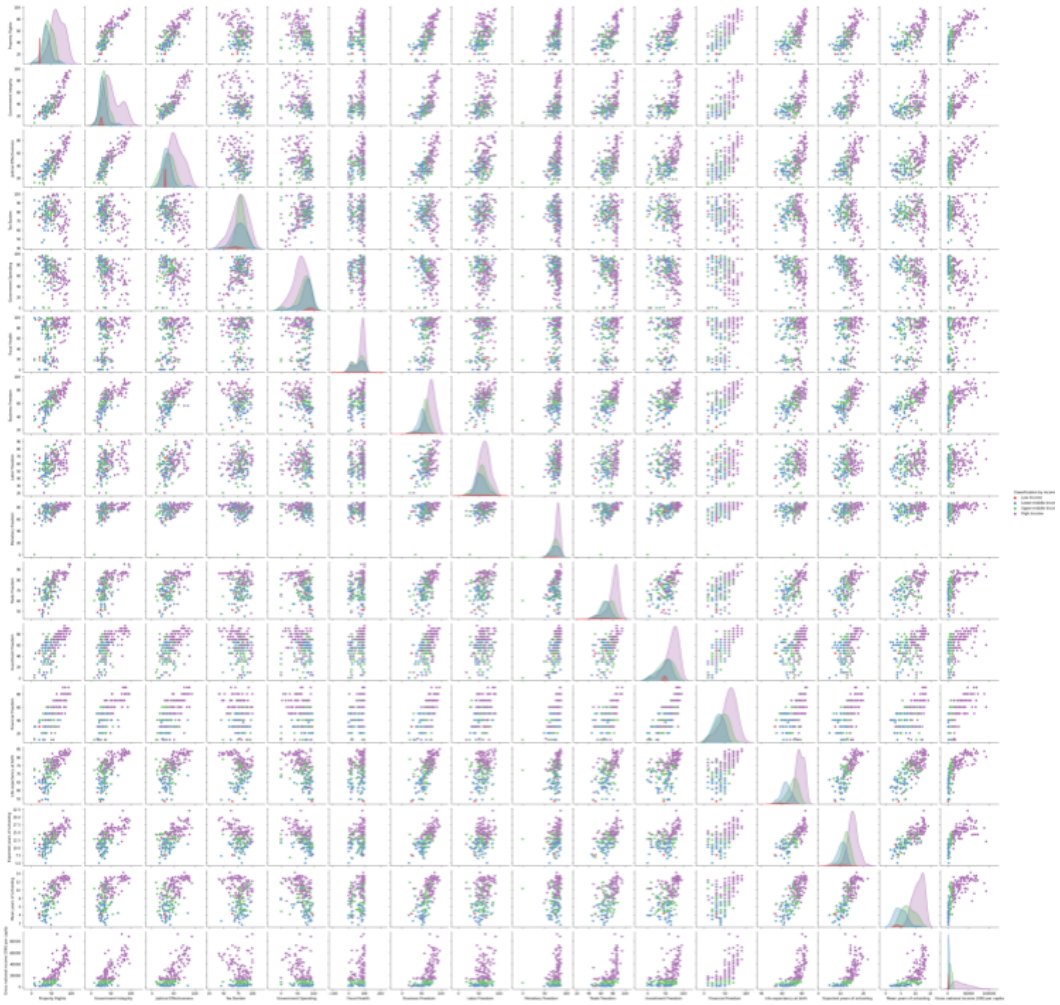

Figure 5: Scatter Plot of All Variables

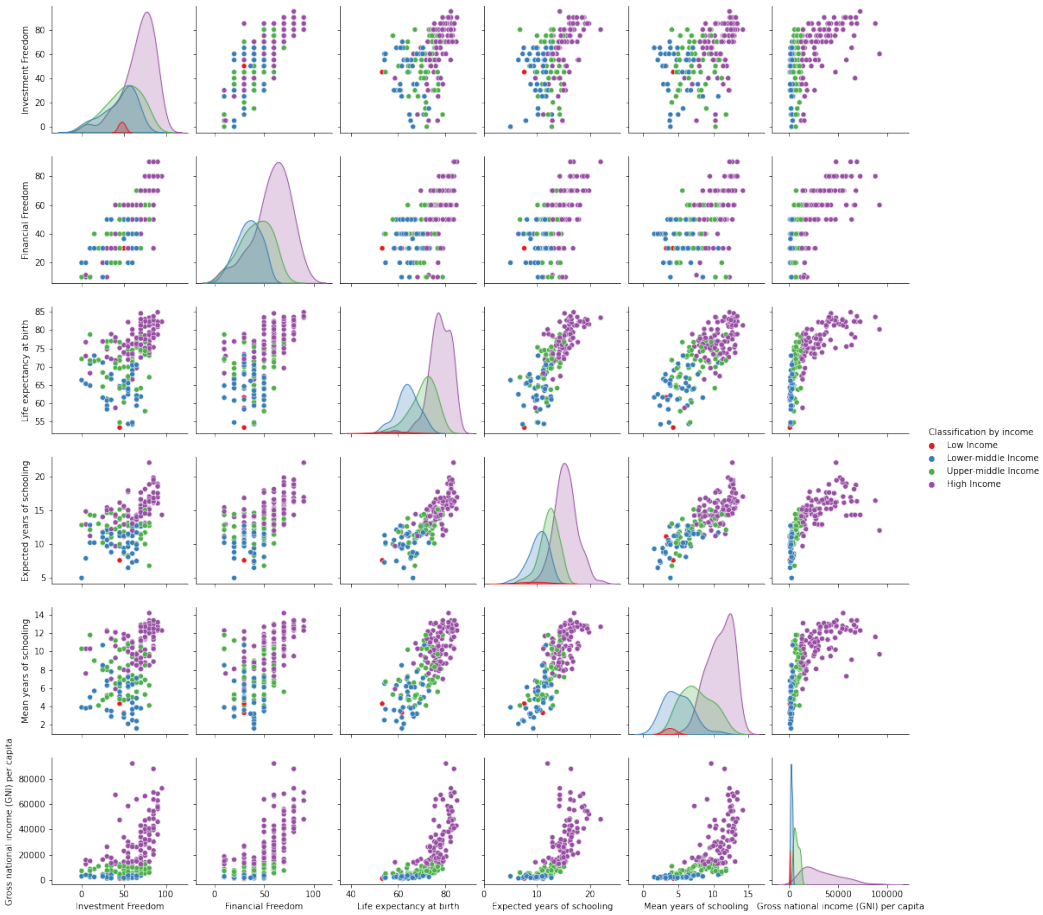

Figure 6: Scatter Plot of Selected Variables

In general, most of the pairs of variables show positive linear trends, with the exception of those involving `GNI per capita` which shows nonlinear interactions. Therefore, we can get a hint that there is a positive correlation between the two set Index of Economic Freedom and HDI and between the variables within each set as well. This can be further examined using the correlation matrix plots.

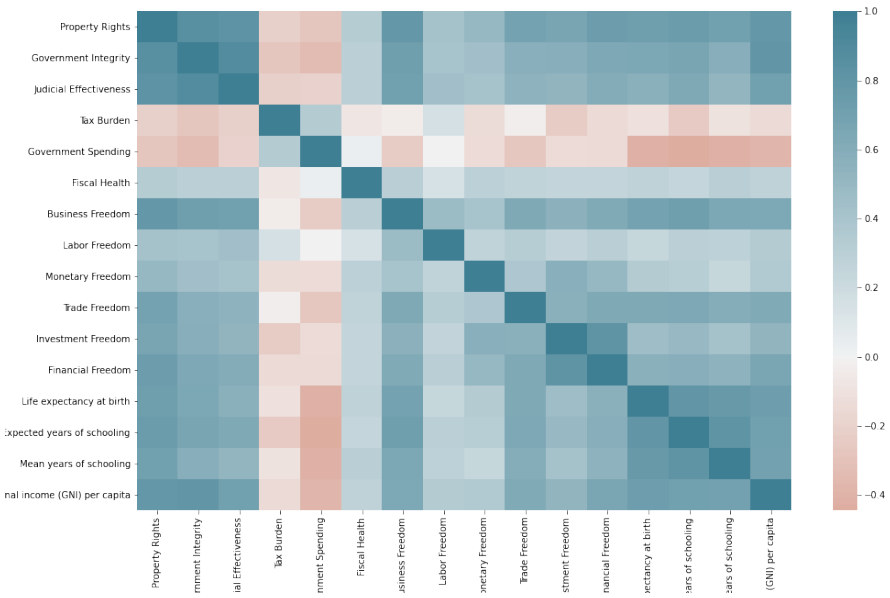

Figure 7: Correlation Matrix Plot

As expected, we can see that most of our speculations are verified. However, there are 2 variables that stand out with negative correlation with all the other variable but themselves: `Tax Burden` and `Government Spending` (index 4, 5). Furthermore, the next variable - `Fiscal Health` (index 6) - shows weak correlations with most of the other indices. Thus, we can see that the component of economic freedom that those 3 constitute - Government Size - has an adverse effects on the performance of countries on the rest of the variables. `Labor Freedom` and `Monetary Freedom` (index 8, 9) also show similar behaviors to `Fiscal Health`.

Our analysis so far suggests that there is a positive correlation between the variables within the sets of economic freedom and human development. Nevertheless, it is time that we perform our main analysis to extract more accurate multivariate inference.

3.3 Canonical Correlation Analysis

Canonical Correlation Analysis (CCA) is a technique for determining and quantifying the relationships between two sets of variables. Theoretically, it aims to find linear combinations of two random variables vector X and Y that results in maximum correlation. Its computations involve using singular value decomposition on a correlation matrix [4].

In our case, we are interested in the interactions between economic freedom variables and those that form the HDI, we can formulate our X set as the variables that make up the Index of Economic Freedom, Y set as those that make up the Human Development Index. In particular,

`X = (Property Rights, Government Integrity, Judicial Effectiveness, Government Spending, Tax Burden, Fiscal Health, Business Freedom, Labor Freedom, Monetary Freedom, Trade Freedom, Investment Freedom, Financial Freedom),`

`Y = (Life Expectancy at Birth, Expected Years of Schooling, Mean Years of Schooling, GNI per capita).`

3.3.1 Independence Testing

Before any analysis can be performed, we need to test whether the two sets are independent. If it is already so, there is no meaning in performing CCA altogether. However, statistical tests for this

require the multivariate normality assumption. Therefore, we first check for this in our data using the Mardia's Tests of Multinormality. In `R`, this equates to running the function `mvn` from the `MVN` library with the argument `mvnTest = "mardia"`.

	Test	Statistic	p-value	Result
1	Mardia Skewness	2026.40477838127	2.89536163382093e-104	NO
2	Mardia Kurtosis	14.4282717124555	0	NO

Table 1: Mardia's Test Results

The results very strongly reject the null hypothesis, thus our data is highly non-normal. In this case, we resort to bootstrapping for independence testing. Our test statistic is

$$T = -(n - \frac{1}{2}(p + q + 3)) \ln(\prod_{k=1}^{m}(1 - \hat{\rho}_k^2).$$

Using the Efron and Tibshirani algorithm for bootstrap hypothesis testing, we obtained a p-value of 0.061 which is, surprisingly, only marginally significant. Nevertheless, we can reject the null hypothesis where X and Y are independent at 90% confidence level ($\alpha = 0.1$).

3.3.2 Technical Implementation

The analysis was conducted with the help of the `CCA` library, in particular the `cc` function. Certainly, the method returns $k = 4$ canonical correlations when we input the data.

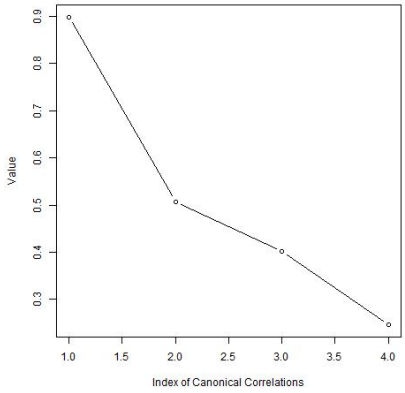

Figure 8: Values of Different Canonical Correlations By k

We can see that the first and second canonical correlations ρ_1 and ρ_2 explain much more the correlations between X and Y than the rest. Therefore, we mainly focus on them during inference. We can view the Canonical Correlation Coefficients (CCC) α and β using the `xcoef` and `ycoef`, respectively, on the return value of the `cc` function.

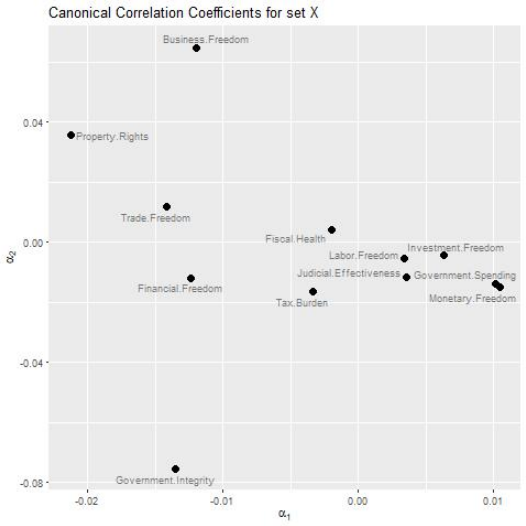

Figure 9: Canonical Correlation Coefficients of Set X For the First 2 Canonical Correlations (α_1 and α_2)

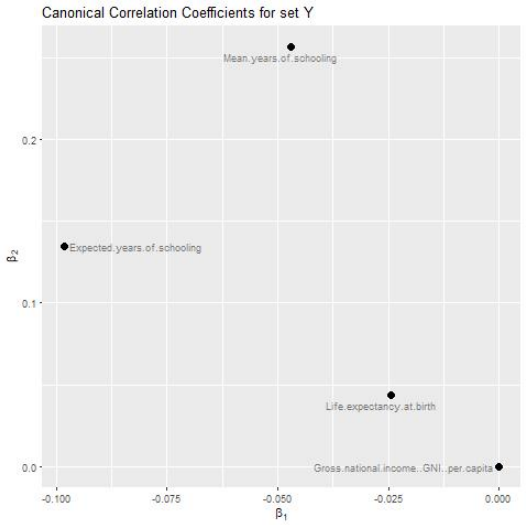

Figure 10: Canonical Correlation Coefficients of Set Y For the First 2 Canonical Correlations (β_1 and β_2)

Looking at the CCC for the Y set, we can see that v_1 has negative correlation with most of the Y variables. Therefore, the interpretation for the first pair of canonical variates u_1 and v_1 is the opposite of their signs: variables in the set X which show *negative* correlations with u_1 have *postive* effects on those that are in the Y set. With regards to the the CCC of the X set, u_1 has negative correlation with half of the variables - $x_1, x_2, x_7, x_{10}, x_{12}$. This means that u_1 can be thought to represent the `Property Rights, Business Freedom, Trade Freedom, Financial Freedom, Government Integrity`, and, to a very small extent, `Fiscal Health` and `Tax Burden`. Subsequently, they, especially `Property Rights`, affect Human Development positively, in contrary to the rest of the X variables.

From the second canonical correlation ρ_2, we can find a much different situation. The most notable levels of correlation with u_2 can be found among the 3 variables: x_1, x_7 and x_2. The first two

demonstrate positive correlation, while the last one shows the contrary. On the other set, v_2 has strong positive correlation with y_2 and y_3, but little with y_1 and y_4. All of this shows that `Business Freedom` and `Property Rights` impact `Mean Years of Schooling` and `Expected Years of Schooling` positively, while `Government Integrity` does the opposite in ρ_2.

Looking at both ρ_1 and ρ_2 holistically, it is evident that x_1 and x_7, or `Property Rights` and `Business Freedom`, strongly affect human development positively. Meanwhile, x_2 has a more complicated relationship with both u_1 and u_2, creating uncertainty in inference. On the other hand, x_5 and x_9 show a positive correlation with u_1 and a negative one with u_2. This shows that `Government Spending` and `Monetary Freedom` have an adverse effect on human development.

Finally, we can plot the scores of countries on u_1 and v_1 by using the method `xscores` and `yscores`, respectively.

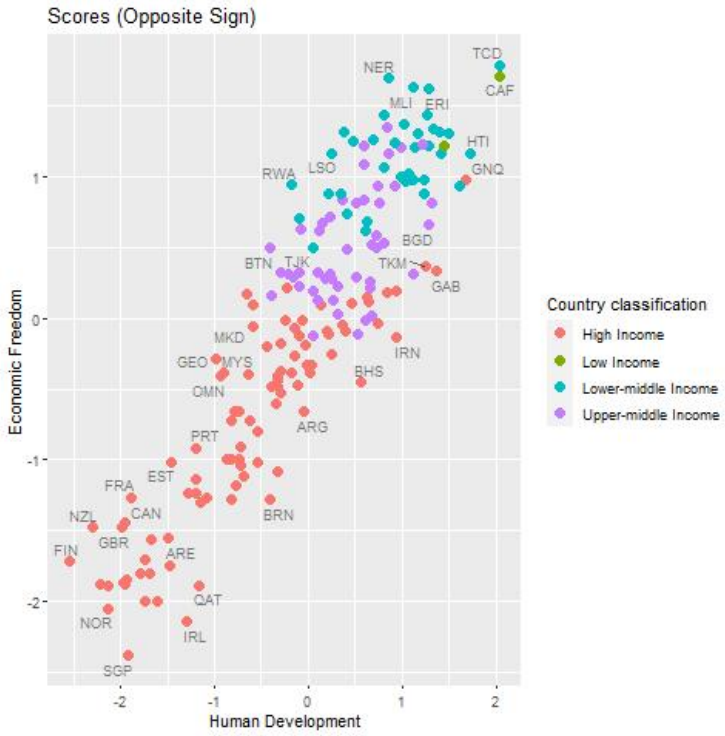

Figure 11: Scores of Countries on the First Pair of Canonical Variables

Regarding the x-axis of the plot, we have countries to the leftmost enjoying the highest expected years of schooling, mean years of schooling and life expectancy at birth. The extent of the effects follow the listed order of the variables. On the y-axis, to the bottom are those with the highest property rights, trade freedom, government integrity, financial freedom, business freedom with magnitude in the listed order.

4 Conclusion and Future Works

This paper present a statistical analysis on the relationship between economic freedom and human development. Among the 12 economic freedom variables, `Business Freedom`, `Property Rights` and `Trade Freedom` have the biggest positive effects on human development. `Labor Freedom`, `Investment Freedom`, `Judicial Effectiveness`, `Government Spending`, `Monetary Freedom`, on the other hand, have the opposite, albeit smaller, effects. The variables that have little or no effects on human development are `Fiscal Health`, `Labor Freedom`, `Judicial Effectiveness`. There remains 3 variables which has complicated effects which makes inference difficult are `Financial Freedom`, `Government Integrity` and `Tax Burden`.

On the other side, half of the parameters (`Mean Years of Schooling` and `Expected Years of Schooling`) have significant positive response to economic freedom. `Life Expectancy at Birth` also show some positive correlation, but to a lesser extent. Surprisingly, `GNI per capita` shows no correlation with economic freedom in both the pairs of canonical variates. This may be due to the fact that it has a nonlinear relationship with the other variables, as shown in bivariate analysis. For future studies, the focus can be shifted from absolute values of the data to their difference with the past ones, or changes over time. This can help us see the immediate effects of the variables.

5 Critical Evaluation

The foremost problem lies in the construct of the Index of Economic Freedom itself. Many of its measures are inherently subjective, as they are purely human interpretation of reality. For example, *foreign investment code* was a criteria used to examine the investment freedom of a country following this rubric:

- No transparency and burdensome bureaucracy: 25 points deducted
- Inefficient policy implementation and bureaucracy: 15 points deducted
- Some investment laws and practices non-transparent or inefficiently implemented: 5 points deducted

The question this prompts is how we can divide the thin line between having up to 15 points and only 5 points deducted from the score of a country. What is the single tipping point that differentiates "inefficient policy implementation and bureaucracy" and only having "some investment laws and practices non-transparent or inefficiently implemented"? The same problem applies to many other criterion used in the index.

References

[1] 2021 index of economic freedom. https://www.heritage.org/index/. Accessed: 2021-04-10.

[2] Edgars Rožāns. *The Socioeconomic Impact of Economic Freedom*. Number ID 2683493. Oct 2015.

[3] Angel Hsu and Alisa Zomer. Environmental performance index. *Wiley StatsRef: Statistics Reference Online*, page 1–5, 2016.

[4] Daniel Hsu, Sham M. Kakade, and Tong Zhang. A spectral algorithm for learning hidden markov models. *Journal of Computer and System Sciences*, 78(5):1460–1480, 2012. JCSS Special Issue: Cloud Computing 2011.