Empirical Analysis of Alternative Human Development Index Metrics

Robert Mepham
Middle Tennessee State University, Murfreesboro, TN

Paper submitted to complete course requirements in Econometrics & Forecasting (ECON 4620), December 2021

Abstract

The United Nation's Human Development Index (HDI) is often criticized for its oversimplification of the concept of human development into a concise quantitative value. Many critics have proposed alternative theories to the HDI, whether it be additional components to alternative mathematical procedures to calculate a quantitative summary of human development. This paper sought to evaluate the legitimacy of many alternative human development metrics and test how they perform as additional components to the current HDI model. The most effective alternative metrics at predicting the portions of human development that are not readily covered by the components of the current HDI are Columbia University's "Environmental Performance Index", the Economic Freedom Index's "Business Freedom" measure, and the Social Progress Index's "Personal Freedom and Choice" value. These three metrics capture about 66.5% of the unexplained variability from the base HDI model, thus cutting the error of the base model by a factor of three.

Key words: Human Development, Linear Regression Analysis

JEL category: O11

Robert Mepham, Student, Department of Economics and Finance, Middle Tennessee State University, Murfreesboro, TN 37132, phone: 224-279-4156, email: rgm3a@mtmail.mtsu.edu

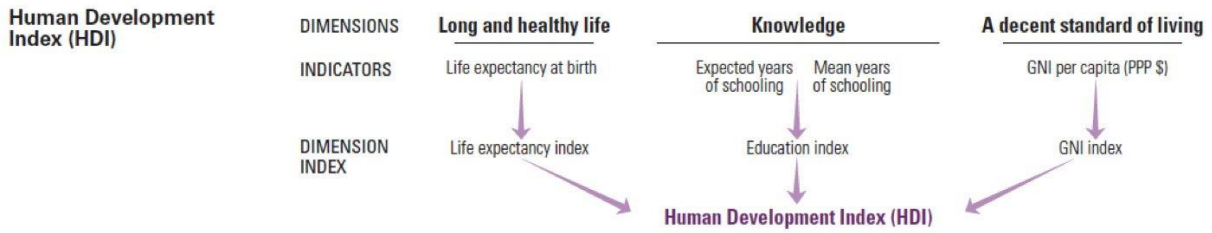
Table of Contents

Introduction

The Human Development Index has been used by the United Nations as an objective standard of evaluating nation's progress in elevating the quality of life for their people for 30 years. The UN refers to the HDI when forming global-scale initiatives such as their Sustainable Development Goals.

The United Nations first published their Human Development Index (HDI) in 1990. As a measure, the HDI is calculated as follows.

Figure 1: Visual Representation of the Human Development Index & Its Components



Source: http://hdr.undp.org/en/content/human-development-index-hdi

The HDI is a very simple index, but it does not come without its critics. This paper will introduce various alternative methodologies for measuring Human Development proposed by critics of the UN's current model. After introducing these alternative theories, their metrics will be put to the test to see if they empirically capture variation in the real HDI score that is not captured by the basic UN-approved measures themselves.

A simple linear regression will be used wherein the various alternative Human Development measures will be regressed against the residual values from the HDI regression model. From the results of this regression, three alternative measures will be selected to be included in this paper's proprietary metric the "newHDI". Then, the newHDI scores for each

nation will be compared to the 2017 HDI scores to see which nations benefit the most from the inclusion of these new metrics, and which were disadvantaged the most by including the additional measures.

Literature Review

History of the Human Development Index

In the 1980's, Dr. Mahbub ul Haq was chosen by the United Nations Development Project to head up a new era of quantitative analysis of human development. Before Dr. Haq arrived at the UNDP, many efforts had been made to measure an objective standard of human development across the globe, but none had been ultimately successful in satisfying the UNDP's administrators (Stanton, 2007). Dr. Haq lead the empirical revolution of Human Development-based economics with the breakthrough index, the Human Development Index (HDI). The HDI is simple by nature, and Dr. Haq intended it to be that way as a measure that was able to be derived even from the most isolated nations. The HDI is composed of three simple indices: Life Expectancy, Education, and Gross National Income (UNDP, 2014). As a result of the revolution that the HDI created in Human Development-based economics and anthropology, the United Nations founded the "Mahbub al Haq Award for Outstanding Contribution to Human Development" in 2014 (UNDP, 2014).

Alternative Mathematical Methodologies

Since its first publication in 1990, many critics have spoken out against the HDI, either for its mathematical composition, or the merits of the components on which it is based. Many critics have condemned the HDI for its simplistic mathematical composition. A simple linear regression is bound to oversimplify the true complexities present in the analysis of human development, thus losing efficiency as a mathematical model (Wu et al., 2013). Some critics, including D K Despotis believe that the HDI's ranking of countries is arbitrary at best, since it does not do enough to adjust for factors such as spatial autocorrelation (Despotis, 2005). Other proponents of alternative

methodologies go so far as to condemn the United Nations themselves for choosing measures of Human Development that are too Westernized.

From a mathematical perspective, the HDI is very simple. Proponents of an alternative methodology have suggested using a model that is capable of reducing selection bias such as a Least Absolute Shrinkage and Selection Operator (LASSO) regression model (Noorbakhsh, 1998).

Categories of Alternative Metrics

In terms of alternative metrics that have been proposed to be additions to the HDI, the United Nations has formally recognized six categories of alternative components: Subjective Metrics, Goals-Based Measures, Sustainable-Indexes, Comprehensive Measures, Specific Measures, and Economic Freedom Measures (Comim, 2016). In addition to the recognition of these alternative metrics that have a reasonable contribution to human development outside of the current HDI measures, many believe that the original three sets of metrics included within the HDI are "redundant" and "ambiguous" (McGillivray, 1991). These claims stem from the notion that Human Development itself is a measure that cannot be quantitatively defined across such diverse cultures in an objective and unbiased way (Rannis et al., 2006).

Diving deeper into the alternative categories can reveal the prevailing hypotheses that lead economists around the globe to theorize that there exist some measures that are not included within the HDI that ought to significantly contribute to the Index as it is implemented into the future.

First up is the economic side of the argument. The HDI only recognizes Gross National Income (GNI) as a measure that contributes to human development. While GNI is an objective measure of the purchasing power of individuals within a state, other subjective measures of economic development have a significant contribution to a nation's economy (Osberg & Sharpe,

2010). It is no secret that people's confidence in an economy is another great indicator of economic wellbeing within a country, which is why some critics of the HDI have included additional economic measures that account for people's perception of the economy as a significant contributor to alternative models for human development (Mangaraj & Aparajita, 2020).

On the other side of the additional metrics argument is environmental-based metrics. Environmental factors are noticeably absent from the HDI's component list. Economists have been arguing this point since the 1990's, stating that the absence of environmental factors disregards a society's position on its environment as a limiting factor in economic-based measures like the GNI (Sagar & Najam, 1998). Essentially, a country that does value environmental sustainability over economic strength in the traditional western sense, might be quantitatively hurt in the HDI rankings without recognition that human development does not objectively suffer when a legitimate alternative to economic development is pursued. This revisits the claims made earlier that the HDI is an overly Western model because the GNI itself is a metric that many have argued reflects a western perspective on economic development, instead of a world view. In recognition of this tradeoff between economic development and environmental sustainability, the Wellbeing Economy Alliance, and the New Economics Foundation have jointly developed and sponsored the adoption of the "Happy Planet Index" as a mid-point between an economic index and an environmental index (Marks).

Despite these many alternative theories, the HDI itself serves a utilitarian role as a simplistic measure of human development. This is why this paper is not seeking to refute the HDI, but merely add to its predictive power by integrating the most statistically valid of these alternatives with the current HDI model. Not all of the alternatives introduced above are statistically valid. If a measure does not add significant predictive power on top of the current

HDI model, then there is no point in including it in a new and improved model, since the current HDI factors account for that variation already. It is for that reason that this paper seeks to find the statistical validity of these alternative metric categories as well as determine which of those metrics can function most effectively as additions to the HDI's set of predictors.

Data

Data Structure

The base HDI data has one observation for each country in the dataset. For this reason, the rest of the data that was sourced needed to follow this basic structure, wherein each observation represents an individual country. Once all of the independent variables were sourced, they were all merged together on their respective ISO3 code, allowing for the effective merging of data sources despite inconsistencies in how each source spells the given country and/or territory.

Data Sources & Variable Definitions

Table 1: Data Definitions & Sources – United Nations Human Development Report

Variable Name	Description	Source	Year
HDI	Human Development Index Value	United Nations Human Development Report	2018
lex	Life Expectancy at Birth (Years)	United Nations Human Development Report	2018
mnyrsskl	Mean Years of Schooling (Years)	United Nations Human Development Report	2018
logGNIpc	Logarithm of Gross National Income per Capita	United Nations Human Development Report	2018

Source: Ind_Variable_Descriptions.csv

Table 2: Data Definitions & Sources – World Happiness Report

Variable Name	Description	Source	Year
SocialSupport	Avg. Y/N Response to "If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them?"	World Happiness Report	2020
HLE	Healthy Life Expectancy at Birth (Years)	World Happiness Report	2020
FreedomToMakeLifeChoices	Avg. Y/N Response to "Are you satisfied with your freedom to choose what you do with your life?"	World Happiness Report	2020
Generosity	Avg. Y/N Response to "Have you donated money to a charity in the past month?	World Happiness Report	2020
CorruptionPerceptions	Avg. Y/N Response to "Is corruption widespread?"	World Happiness Report	2020

Source: Ind_Variable_Descriptions.csv

Table 3: Data Definitions & Sources – OECD – Green Growth Report

Variable Name	Description	Source	Year
GPAT_DE_CAP	Development of environment-related technologies, inventions per capita	OECD - Green Growth Report	2016
GPAT_DE_AT	Development of environment-related technologies, % all technologies	OECD - Green Growth Report	2016

Source: Ind_Variable_Descriptions.csv

Table 4: Data Definitions & Sources – Social Progress Index

Variable Name	Description	Source	Year
Nutrition and Basic Medical Care	Index based on five "Nutrition and Basic Medical Care" Factors	Social Progress Index	2018
Personal Safety	Index based on four "Personal Safety" Factors	Social Progress Index	2018
Access to Basic Knowledge	Index based on five "Access to Basic Knowledge" Factors	Social Progress Index	2018
Access to Information and Communications	Index based on four "Access to Information and Communications" Factors	Social Progress Index	2018
Personal Freedom and Choice	Index based on four "Personal Freedom and Choice" Factors	Social Progress Index	2018
Personal Rights	Index based on five "Personal Rights" Factors	Social Progress Index	2018

Source: Ind_Variable_Descriptions.csv

Table 5: Data Definitions & Sources – Economic and Social Rights Fulfillment Index

Variable Name	Description	Source	Year
Right to Education Index	Index Measuring people's access to education	Economic and Social Rights Fulfillment Index	2016
Right to Health Index	Index Measuring people's access to health	Economic and Social Rights Fulfillment Index	2016
Right to Housing Index	Index Measuring people's access to housing	Economic and Social Rights Fulfillment Index	2016
Right to Food Index	Index Measuring people's access to food	Economic and Social Rights Fulfillment Index	2016

Source: Ind_Variable_Descriptions.csv

Table 6: Data Definitions & Sources – OECD - PISA Scores

Variable Name	Description	Source	Year
pisamath	Avg. Mathematics Score for 15-year-olds on the PISA exam	OECD - PISA Scores	2015
pisaread	Avg. Reading Score for 15-year-olds on the PISA exam	OECD - PISA Scores	2015
pisascience	Avg. Science Score for 15-year-olds on the PISA exam	OECD - PISA Scores	2015

Source: Ind_Variable_Descriptions.csv

Table 7: Data Definitions & Sources – Economic Freedom Index

Variable Name	Description	Source	Year
Monetary Freedom	EFI - Monetary Freedom Measure	Economic Freedom Index	2021
Labor Freedom	EFI - Labor Freedom Measure	Economic Freedom Index	2021
Trade Freedom	EFI - Trade Freedom Measure	Economic Freedom Index	2021
Business Freedom	EFI - Business Freedom Measure	Economic Freedom Index	2021

Source: Ind_Variable_Descriptions.csv

Table 8: Data Definitions & Sources – Individual Sources

Variable Name	Description	Source	Year
Happly Planet Index	Index Based on: "Experienced Wellbeing", "Life Expectancy", and "Ecological Footprint"	Happy Planet Index	2016
EPI	Environmental-Based Index by Columbia & Yale	Environmental Performance Index	2020
UNICEF_Median_Infant_Morality_Rate	Median Infant Mortality Rate (IMR)	UNICEF - Infant Mortality Report	2020

Source: Ind_Variable_Descriptions.csv

Descriptive Statistics

The variables have been split into two basic categories:

- Components of the UN's base HDI model
- All alternative variables that this paper will assess in relevance to the residuals of the UN's base HDI mode

Table 9: Descriptive Statistics, Base HDI Components

Variables	n	mean	sd	min	max
HDI	189	0.713	0.151	0.377	0.954
lex	191	72.414	7.510	52.805	84.687
mnyrsskl	189	8.613	3.082	1.586	14.132
logGNIpc	191	9.233	1.173	6.492	11.613

Source: Base_HDI_descrip.csv

Table 10: Descriptive Statistics, HDI Residual Predictors (All)

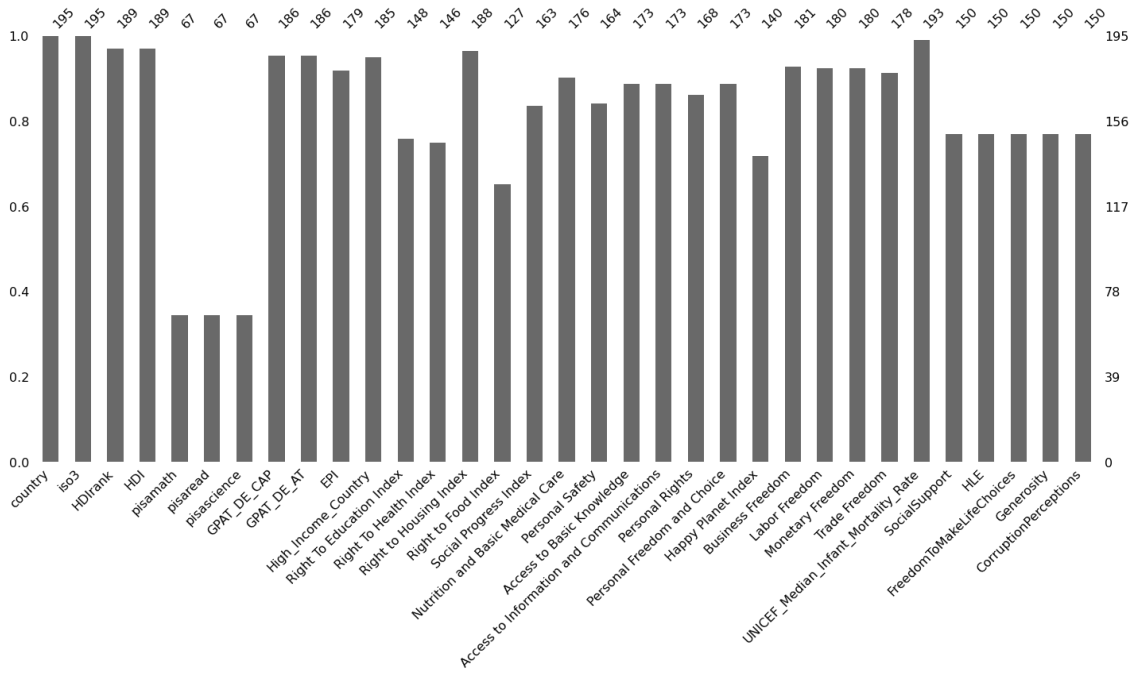
Variable	n	mean	sd	min	max
HDI_resid	156	-0.0692	0.0414	-0.1909	0.0011
pisamath	156	399.2988	86.8542	122.2670	564.0000
pisaread	156	397.1473	87.2001	145.1919	535.0000
pisascience	156	408.3408	82.3872	170.7604	556.0000
GPAT_DE_CAP	156	4.3909	11.1801	0.0000	69.1700
GPAT_DE_AT	156	8.8688	12.5252	0.0000	100.0000
EPI	156	47.2436	16.0800	22.6000	82.5000
Right.To.Education.Index	156	80.9695	14.4170	39.3700	107.0929
Right.To.Health.Index	156	78.1956	12.4914	42.2700	104.2040
Right.to.Housing.Index	156	71.3354	26.3662	11.6500	100.0000
Right.to.Food.Index	156	85.1299	23.1749	32.9400	149.7634
Social.Progress.Index	156	68.2589	15.3063	30.7400	92.8700
Nutrition.and.Basic.Medical.Care	156	84.1461	15.2237	36.3200	98.9600
Personal.Safety	156	68.3501	14.2101	19.5000	96.6500
Access.to.Basic.Knowledge	156	77.2754	18.4556	24.1900	99.1300
Access.to.Information.and.Communications	156	66.5862	20.1064	5.8000	98.3300
Personal.Rights	156	71.0290	21.3853	12.5900	98.0700
Personal.Freedom.and.Choice	156	63.4037	16.7223	19.1300	91.6500
Happy.Planet.Index	156	25.8938	7.1097	12.8000	44.7000
Business.Freedom	156	64.3167	14.9748	17.3000	94.4000
Labor.Freedom	156	59.5333	13.3705	20.0000	91.5000
Monetary.Freedom	156	75.5141	10.0571	0.0000	86.9000
Trade.Freedom	156	72.3629	10.7637	40.8000	95.0000
UNICEF_Median_Infant_Mortality_Rate	156	20.8059	19.7327	1.5531	81.0032
SocialSupport	156	0.8043	0.1211	0.3195	0.9747
HLE	156	64.0706	7.0618	45.2000	76.8046
FreedomToMakeLifeChoices	156	0.7852	0.1129	0.3966	0.9750
CorruptionPerceptions	156	0.7374	0.1724	0.1098	0.9356

Source: HDIresid_descrip.csv

Discussion of Multiple Imputation

Table 2 shows a total of 189 countries having full data, but Table 3 only has 156 full countries. This is due to the fact that the alternative data sources did not have all variables for all countries. Only 88 countries had a value for every single additional attribute. A conscious decision was made to avoid a scenario in which a regression model for every country would only be based on the 88 countries with full data available. This would introduce a massive amount of bias, since a majority of those countries with full data were Westernized countries, which, as discussed earlier, is already a main criticism of the current HDI model, so that same criticism is to be avoided at all reasonable costs.

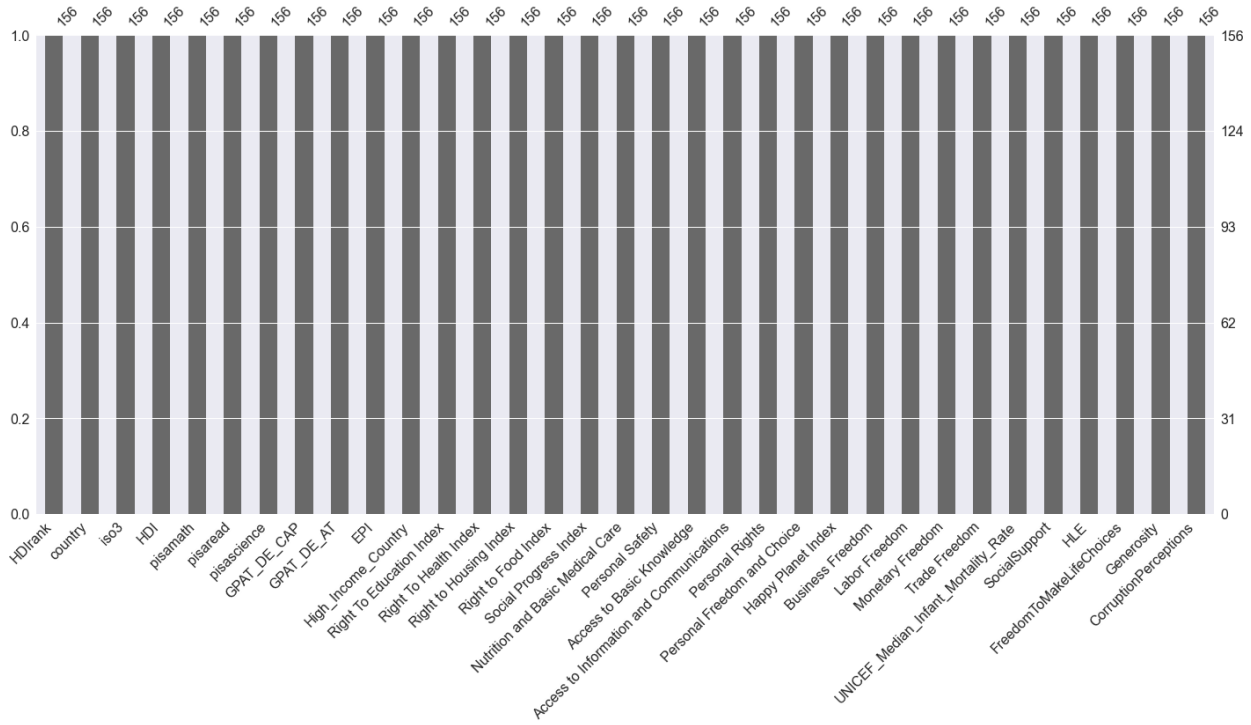
Figure 2: Missing Data Matrix – Raw Data



Source: Missingno Python Package (Bilogur, 2018)

The figure above shows the relative abundance of missing data in the combined raw data. In order to fix the amplitude of missing data, multiple imputation will be applied (Scikit-Learn, 2011).

Figure 3: Missing Data Matrix – Imputed Data



Source: Missingno Python Package (Bilogur, 2018)

The resulting data frame from the multiple imputation process is the raw data that will be used from here on out during the development, execution, and analysis of the R script where the regression models will be defined.

Methods

The first regression that will be performed is the regression that will assess the makeup of the current HDI. The three factors that make up the UN's HDI model will be regressed against the current HDI index value for each nation. This resulting model will be assessed for functional fit, heteroskedasticity, multicollinearity, and spatial autocorrelation. The results of this model will reveal the partitioned-R2 value for each variable, thus showing the relative importance of each of the three components of the current HDI

Once the relative importance is established, and the total R2 value is determined, the remaining variability not captured by the base HDI model will be determined by the formula

$$R^2_{rem} = 1 - R^2 \quad (1)$$

This remaining R2 is what will hopefully be filled by the addition of the best alternative variables into the current HDI model. From this model, the residual values will be saved and added to the raw dataset of alternative metrics to be used as the dependent variable on which the alternatives are regressed onto. This is because the crux of this paper is to specifically capture variation in Human Development that is not currently captured by the HDI, so regressing onto the HDI is pointless when the residuals (HDI_resid) can be used to evaluate the additional predictive power of the alternative measures more effectively.

From these residual values, a new model will be created, consisting of the HDI_resid as the dependent variables, and the entire set of alternative measures defined above as the independent variables. This model will be assessed for multicollinearity, functional fit, heteroskedasticity, and deliver the R2 and partitioned R2 values necessary to evaluate the individual alternative metrics against each other. The partitioned R2 values will be used as the criterium to evaluate the

alternatives against each other in order to determine which three can be added to the HDI model to increase its predictive power most effectively.

Results

The first model that was created was in order to partition the predictive power of the basic three components of the current HDI: Life Expectancy, GNI, and avg. years of schooling.

Table 11: Unrestricted Regression Model, Base HDI Components

Variables	Estimate	Std..Error	t.value	Pr...t..	VIF	r2p
(Intercept)	-0.3649	0.0150	-24.3734	0.0000	NA	NA
lex	0.0056	0.0003	19.2439	0.0000	3.3842	0.3070
logGNIpc	0.0574	0.0020	29.0604	0.0000	3.9090	0.3617
mnyrsskl	0.0165	0.0007	24.6421	0.0000	3.1019	0.3201

Notes: N = 187, R2 = 0.988, RESET = 0.233, NCVTest p-value = 0.129, LM_Lag p-value = 0.6975
Source: Base_HDI_regres.csv

Table 11 shows the basic regression breakdown of the traditional HDI model. This model was tested for functional fit, which is why the GNI per capita value has been replaced with the logarithm of GNI per capita because the original GNI per capita values were orders of magnitude larger than the other independent variables, thus causing predicted HDI to not be a linear function of the full set of predictors. Multicollinearity, heteroskedasticity, and spatial autocorrelation were also tested for, and found not to be significantly present in the base model.

Table 11 also determines that all three variables are significant predictors of HDI (p-values all < 0.1). Using partitioned R2 values to attribute portions of the 98.8% of the variability captured by the model to individual variables shows that logGNIpc captures the most variability (36.17%), with mnyrsskl capturing 32.01%, and lex capturing the least variability of any of the main predictors (30.7%).

The overall R2 for the base model was 98.8%, meaning that there is only 1.2% of variability in the HDI that is not captured by the base three predictors. This is the variability that the alternative metrics are attempting to capture. In order to ensure that the alternatives are being

evaluated on their ability to capture this specific 1.2% of variability in the HDI, the alternative metrics will be regressed onto the residuals of this first model, to avoid bias in variable selection.

Table 12: Unrestricted Regression Model, HDI Residual Predictors (All)

Variable	Estimate	Std..Error	t.value	Pr...t..	VIF	r2p
(Intercept)	-0.2135	0.0258	-8.2691	0.0000	NA	NA
Access.to.Basic.Knowledge	0.0000	0.0002	-0.0317	0.9748	7.7792	0.1467
Personal.Rights	0.0000	0.0001	0.1366	0.8915	2.3526	0.0994
Monetary.Freedom	0.0003	0.0002	1.8997	0.0596	1.4762	0.0813
Labor.Freedom	-0.0001	0.0001	-1.0507	0.2953	1.4023	0.0786
GPAT_DE_CAP	-0.0003	0.0002	-1.5581	0.1215	2.1292	0.0622
GPAT_DE_AT	0.0003	0.0001	2.5479	0.0119	1.1547	0.0549
Happy.Planet.Index	0.0004	0.0003	1.3568	0.1771	2.1518	0.0442
CorruptionPerceptions	0.0113	0.0130	0.8690	0.3864	2.2634	0.0424
Business.Freedom	0.0006	0.0002	3.2276	0.0016	3.2784	0.0381
EPI	0.0009	0.0002	3.6140	0.0004	7.1249	0.0306
FreedomToMakeLifeChoices	-0.0110	0.0184	-0.5963	0.5519	1.9353	0.0290
SocialSupport	0.0086	0.0236	0.3656	0.7153	3.6554	0.0222
Trade.Freedom	-0.0002	0.0002	-0.7998	0.4252	2.6567	0.0214
Right.to.Food.Index	-0.0001	0.0001	-0.8398	0.4025	3.4770	0.0190
Personal.Safety	-0.0002	0.0002	-1.2845	0.2012	3.2396	0.0116
Right.To.Education.Index	0.0002	0.0002	1.0728	0.2853	3.9160	0.0115
Right.To.Health.Index	-0.0007	0.0002	-2.9085	0.0042	3.5466	0.0109
Personal.Freedom.and.Choice	0.0015	0.0003	5.5967	0.0000	9.5645	0.0093
Right.to.Housing.Index	0.0000	0.0001	0.1120	0.9110	3.7593	0.0089

Notes: Notes: N = 187, R2 = 0.8223, RESET = 0.1026, NCVTest p-value = 0.1927
Source: HDIresid_regres.csv

Table 12 is missing 11 alternative predictors from the full list previously introduced. This is purely because of their issues with multicollinearity. One of the drawbacks of using multiple imputation to fill in missing data, is that the multicollinearity of variables that had large portions of missing data can be artificially increased by this process. Therefore, it is no surprise that variables like the PISA scores (Math, Reading, and Science) all had high multicollinearity values (as indicated by high VIFs).

Similar to the Base HDI model, this regression model was checked for heteroskedasticity and functional fit, passing both tests.

Table 13: Restricted Regression Model, Base HDI Components HDI Residual Predictors

Variable	Estimate	Std..Error	t.value	Pr...t..	VIF	r2p
(Intercept)	-0.2097	0.0159	-13.2158	0.0000	NA	NA
Business.Freedom	0.0005	0.0002	3.5097	0.0006	2.2079	0.1698
EPI	0.0006	0.0002	3.2367	0.0015	3.4602	0.2086
GPAT_DE_AT	0.0003	0.0001	2.7695	0.0063	1.0261	0.0184
Monetary.Freedom	0.0003	0.0002	1.5738	0.1177	1.2269	0.0432
Personal.Freedom.and.Choice	0.0015	0.0002	6.9448	0.0000	5.3259	0.2863
Right.To.Health.Index	-0.0005	0.0002	-2.7266	0.0072	1.9974	0.0696

Notes: N = 187, R2 = 0.7959, F-Test p-value = 0.105, RESET p-value = 0.103, NCV-Test p-value = 0.112
Source: Restricted_HDIresid_regres.csv

Table 13 includes the six variables that were deemed to be significant contributors to the HDI_resid value taken from the first linear regression model. Since the HDI is in and of itself a simplistic model, adding all six of these alternative metrics to the model would unnecessarily complicate the computation of human development across all 187 countries for which data is present. For this reason, only the top three variables in terms of partitioned R2 value will be added to the final formula.

Table 14: Alternative Variables to be Added into NewHDI model

Variable Name	% of Residual Var. Captured	% of HDI Var. Captured
Personal_Freedom_And_Choice	28.63%	0.34%
EPI	20.86%	0.25%
Business_Freedom	16.98%	0.20%

Since these three variables will be added to the original HDI equation (Table 11), the HDI_{new} can be defined as:

$HDI_{new} = f$ (lex, logGNIpc, mnyrsskl, Personal_Freedom_And_Choice, EPI, Business_Freedom)

When placing the predicted coefficients into the equation, HDI_{new} can more specifically be defined as

$$HDI_{new} = \text{-0.3649 + 0.0056*lex + 0.0574*logGNIpc + 0.0165*mnyrsskl + 0.0015*Personal\_Freedom\_And\_Choice + 0.0006*EPI + 0.0005*Business\_Freedom}$$

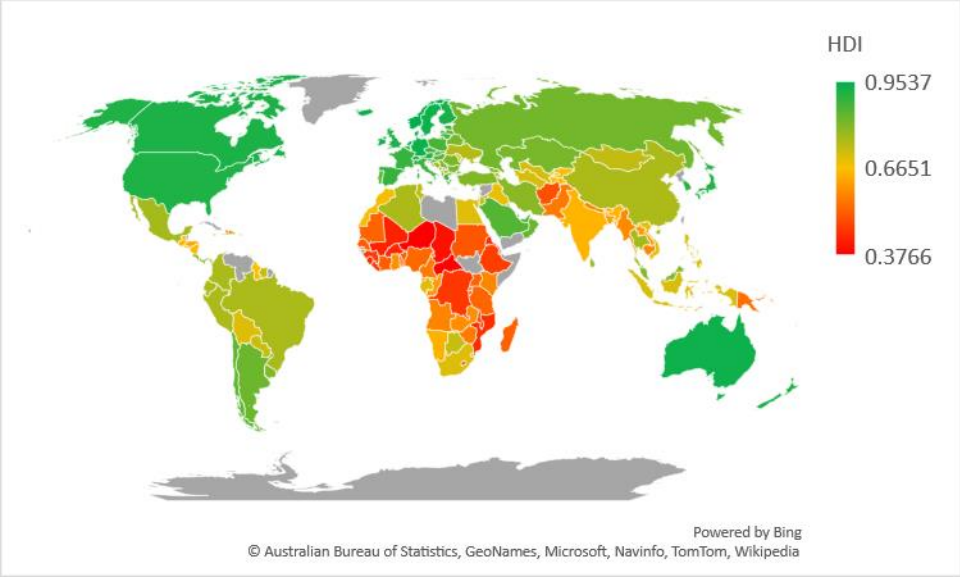
This definition of HDI_{new} is the basis for the rankings that will be discussed in the following sections.

Discussion

The HDI_{new} adds in three components that the UN-approved HDI does not consider. Generally speaking, these include the environment (EPI), freedom (Personal_Freedom_And_Choice), and governmental support of business commerce (Business_Freedom). By adding these three factors into the HDI_{new} there are several countries that were significantly impacted by these additional metrics.

Some countries that perform comparatively well in these new categories will see their HDI_{new} scores significantly improved from their original HDI value and ranking. Conversely, many countries who might've been overvalued by the original HDI, will see their ranking and index value fall by significant margins.

Figure 4: Color-Gradient Map of 2017 HDI Scores

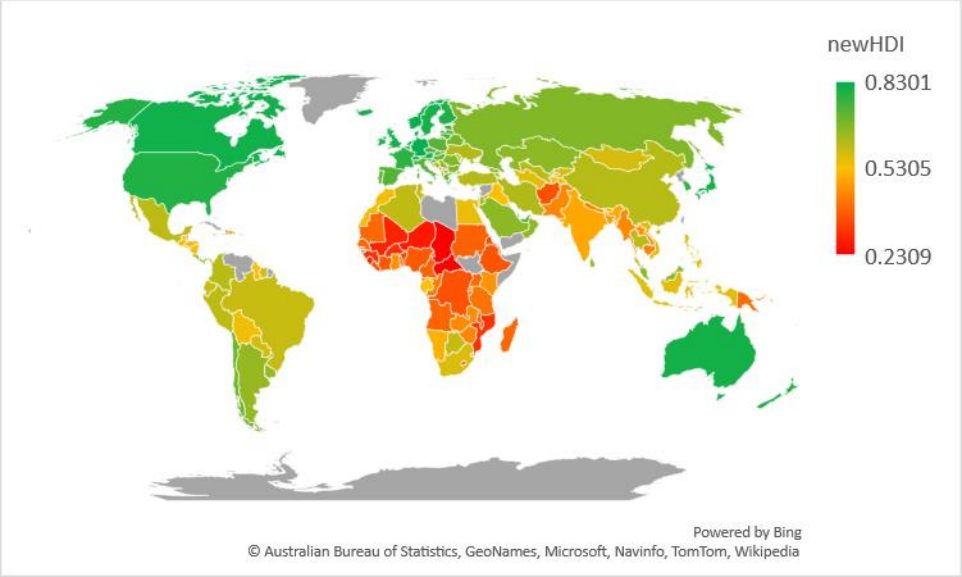


Source: Updated_HDI_Results.xlsx

Figure 4 shows the world's countries ranked by their 2017 UN-Approved HDI values. Regions like Europe, North America, and Australia show a consistent trend of ranking high in the UN's index. This could theoretically be due to the "over-westernization" that was mentioned as a

main critique to the HDI, meaning that the HDI's factors were chosen by primarily western countries, so they would necessarily choose factors that show them in a favorable light compared to regions like Sub-Saharan Africa and Southwestern Asia who do not score as favorably in the original three factors.

Figure 5: Color-Gradient Map of HDI_{new} Scores



Source: Updated_HDI_Results.xlsx

Figure 5 shows the comparative scores for the HDI_{new} values. There are a few countries that have significant changes in their color gradient, meaning that they significantly changed in relative ranking compared to the other countries. With that being said, figures 4 and 5 cannot give the granular detail to assess which countries were particularly effected by this HDI_{new} indexing system. For that type of analysis, the following tables are presented on the following page.

Table 15: Top 10 Countries, HDI_{new}

country	newHDI	Rk	Rk_Change
Switzerland	0.8301	1	1
Germany	0.8235	2	2
Norway	0.8166	3	-2
Canada	0.8096	4	8
Australia	0.8055	5	1
Japan	0.8050	6	11
United Kingdom	0.8047	7	7
Denmark	0.8039	8	2
Sweden	0.8035	9	-2
Iceland	0.8004	10	-5

Source: NewHDI_Ranking_Tables.xlsx

Table 16: Bottom 10 Countries, HDI_{new}

country	newHDI	Rk	Rk_Change
Chad	0.2309	156	-2
Central African Republic	0.2357	155	0
Niger	0.2695	154	2
Mali	0.2851	153	-1
Sierra Leone	0.2928	152	-3
Burkina Faso	0.3020	151	0
Burundi	0.3051	150	3
Guinea	0.3073	149	-4
Mozambique	0.3095	148	0
Gambia	0.3359	147	-3

Source: NewHDI_Ranking_Tables.xlsx

Tables 15 and 16 show the top and bottom 10 countries in the HDI_{new} index, along with their respective ranking change between the original HDI and HDI_{new}. The Rk_Change columns are defined such that a positive value indicates that the country scored comparatively better when the new HDI factors were added to the model versus when only the original HDI factors were considered.

The top of the HDI_{new} rankings are still dominated by mostly western countries from Europe as well as other westernized nations, whereas the bottom HDI_{new} countries still appear to be a plethora of African nations who don't see much change in their rankings even when the new factors are considered.

Table 17: Countries Increasing the Most in NewHDI

country	newHDI	Rk	Rk_Change
Jordan	0.6233	64	22
Uzbekistan	0.6129	72	18
Jamaica	0.6193	67	16
Moldova (Republic of)	0.6078	76	13
Kyrgyzstan	0.5755	89	11
Ukraine	0.6219	66	11
Japan	0.8050	6	11
Armenia	0.6454	60	10
Barbados	0.6857	42	9
Rwanda	0.4374	120	8

Source: NewHDI_Ranking_Tables.xlsx

Table 18: Countries Decreasing the Most in NewHDI

country	newHDI	Rk	Rk_Change
Brazil	0.5962	83	-15
Turkey	0.6148	69	-15
Saudi Arabia	0.6658	49	-15
Dominican Republic	0.5721	91	-13
Congo	0.4241	123	-12
Argentina	0.6535	56	-12
Benin	0.3444	145	-11
Cameroon	0.3839	134	-10
Angola	0.3878	132	-10
Ecuador	0.5952	84	-10

Source: NewHDI_Ranking_Tables.xlsx

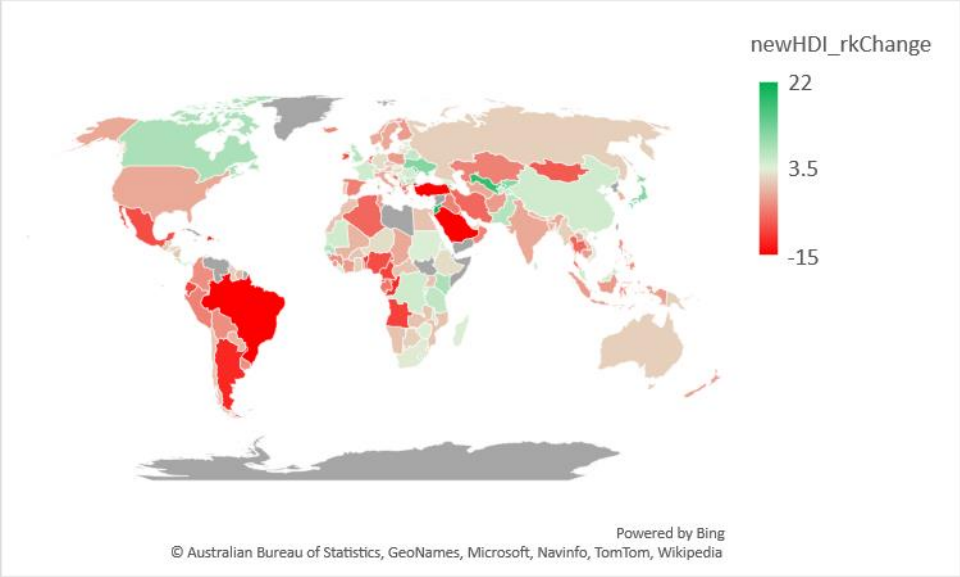
Tables 17 and 18 dive a little deeper into the analysis of HDI_{new} rankings in terms of which states see the biggest increase and decrease in ranking respectively. No definitive trend has been found here, but there appears to be a few former USSR states that see the greatest relative

increase in ranking, as well as some Muslim and South American states that see the greatest relative decrease in ranking.

One proposed theory as to why these particular states are seeing such drastic changes in rankings is two-fold. For the former USSR states, their Personal_Freedom_And_Choice scores are relatively high. This could be due to the fact that there are still significant portions of the population in these nations that grew up under communist rule that will overstate their subjective view of their own personal freedoms because they are comparing it to their freedoms under a former governmental system. For the Muslim states, their Personal_Freedom_And_Choice scores were all relatively low compares to other countries. For the South American states, their Business_Freedom scores were low compared to the world as a whole, likely because of the relative prevalence of perceived corruption within government, leading to a low perception of fair business dealings within the countries.

Surprisingly, of these nations that saw the greatest increases or decreases, none of them saw EPI as the driving force behind their ranking change, indicating that a country's relationship with the environment only amounts to a minuscule driver of human development. This makes sense as many developing countries turn to industries in search of profit, without regard for environmental impact until they reach a later formal stage of development. Even states that went through industrialization hundreds of years ago had this same growth pattern in relation to the environment, so it does level to reason that modern nations experiencing industrialization aren't significantly driven by EPI score.

Figure 6: Color-Gradient Map of Rank Change (NewHDI vs. HDI)



Source: Updated_HDI_Results.xlsx

Figure 6 gives a macro-level view of the relative ranking changes between the original HDI and the HDI_{new} as previously defined. At a macro-level, many South American nations saw decreases in comparative rankings, whereas countries like China, Japan, and Canada saw increases in index rankings.

Summary and Conclusion

The United Nation's HDI model is often criticized for its "arbitrary" selection of components that many view as biased in some regard. However, the HDI serves a key role in understanding the condition of humans across the world, and it is so useful because of its simplicity. Its three-part calculation is what makes it so accessible and translatable across cultural barriers such as language, and region.

For that reason, this paper sought to find three additional components to add the most predictive power to the HDI while taking into account various modeling biases such as heteroskedasticity and multicollinearity. Out of a group of almost 30 alternative metrics, the three that were seen to have the most predictive power on the portions of the HDI left unexplained (residuals) by the base HDI model, were the EPI, Business Freedom, and Personal Freedom & Choice.

The three measures accounted for factors such as environmental sustainability, governmental support of business, and personal perceptions of freedom. These factors were often cited as key factors that were omitted by the UN's original model, and thus by adding these to the new model, coined the HDI_{new}, several nations were significantly impacted in both a positive and negative light when comparing the new rankings to the original HDI rankings.

With these adjusted rankings, no definitive pattern was found as to which countries, regions, or cultural subsets were affected in any particular way by the new rankings, but on a purely individual and anecdotal level, the changes seem to make sense. For example, a country like Brazil, which is often criticized for their treatment of the environment, is severely penalized in the HDI_{new} rankings because it accounts for environmental impact through its inclusion of the EPI as a component.

Based on the empirical analysis above, defining a new HDI that incorporates the EPI, Business Freedom, and Personal Freedom and Choice as contributing metrics along with the three basic metrics that make up the current HDI will improve the HDI as a measurement of human development across the world by capturing 99.6% of the variability in human development with a linear model consisting of only six components.

Works Cited

Bilogur, Aleksey (2018). Missingno: a missing data visualization suite. Journal of Open Source Software, 3(22), 547, https://doi.org/10.21105/joss.00547

Columbia University, & Yale University. (2020). Environmental Performance Index.

Despotis, D. K. (2005). A reassessment of the human development index via Data Envelopment Analysis. *Journal of the Operational Research Society*, *56*(8), 969–980. https://doi.org/10.1057/palgrave.jors.2601927

Comim, Flavio. (2016). (rep.). *Beyond the HDI? Assessing alternative measures of human development from a capability perspective*. United Nations Human Development Report. Retrieved November 15, 2021.

Gallup World Poll, & Lloyd's Register Foundation. (2020). World Happiness Report.

The Hertitage Foundation. (2021). Index of Economic Freedom.

Human Development Reports. Human Development Index (HDI) | Human Development Reports. (n.d.). Retrieved November 15, 2021, from http://hdr.undp.org/en/content/human-development-index-hdi.

Mangaraj, B. K., & Aparajita, U. (2020). Constructing a generalized model of the human development index. *Socio-Economic Planning Sciences*, *70*, 100778. https://doi.org/10.1016/j.seps.2019.100778

Marks, N. (n.d.). (publication). *The Happy Planet Index*. The Wellbeing Economy Alliance (WEAll), The New Economics Foundation (NEF). Retrieved November 15, 2021, from https://happyplanetindex.org/.

McGillivray, M. (1991). The human development index: Yet another redundant composite development indicator? *World Development*, *19*(10), 1461–1468. https://doi.org/10.1016/0305-750x(91)90088-y

Noorbakhsh, F. (1998). A modified human development index. *World Development*, *26*(3), 517–528. https://doi.org/10.1016/s0305-750x(97)10063-8

Organisation for Economic Co-operation and Development (OECD). (2015). Programme for International Student Assessment (PISA).

Organisation for Economic Co-operation and Development (OECD). (2016). OECD - Green Growth Report.

Osberg, L., & Sharpe, A. (2010). The index of economic well-being. *Challenge*, *53*(4), 25–42. https://doi.org/10.2753/0577-5132530402

Sagar, A. D., & Najam, A. (1998). The human development index: A critical review. *Ecological Economics*, *25*(3), 249–264. https://doi.org/10.1016/s0921-8009(97)00168-7

Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

Social Progress Imperative. (2018). Social Progress Index.

Stanton, E. A. (2007). (publication). *The Human Development Index: A History* . Retrieved November 15, 2021, from https://scholarworks.umass.edu/cgi/viewcontent.cgi?article =1101&context=peri_workingpapers.

UNICEF. (2020). Child Mortality.

United Nations Human Development Report. (2018). Human Development Index.

University of Connecticut. (2016). Economic and Social Rights Fulfillment Index.

Wellbeing Economy Alliance, & New Economics Foundation. (2016). Happy Planet Index.

Wu, P.-C., Fan, C.-W., & Pan, S.-C. (2013). Does human development index provide rational development rankings? evidence from efficiency rankings in Super Efficiency Model. *Social Indicators Research*, *116*(2), 647–658. https://doi.org/10.1007/s11205-013-0285-z