A Study of Factors Which Affect Migration

TABLE OF CONTENTS

1. Background

Migration is the movement of people from one place to another. This paper is intended to discuss international migration, which is when people migrate from one country to another, for example, moving from Mexico to Canada. There are two key migration terms: Emigration and immigration. Emigration is when someone leaves a country and immigration is when someone enters a country. One can know whether the country is an emigration country or an immigration country from the data such as, if the migration number is positive the country should be an immigration country, vice versa.

"More people than ever are living abroad. In 2013, 232 million people, or 3.2 per cent of the world's population, were international migrants, compared with 175 million in 2000 and 154 million in 1990. The new estimates include breakdowns by region and country of destination and origin, and by sex and age. The North, or developed countries, is home to 136 million international migrants, compared to 96 million in the South, or developing countries. Most international migrants are of working age (20 to 64 years) and account for 74 per cent of the total. Globally, women account for 48 per cent of all international migrants."(EMBARGOED (2013, Sep 11), 232 million international migrants living abroad worldwide –New UN global migration statistics reveal, UN PRESS RELEASE, retrieved from http://esa.un.org/unmigration/wallchart2013.htm)

There are many reasons that people choose to migrate to other countries, such as lack of services, lack of job opportunities, war in their motherland, high crime, poverty, or some other environmental factors such as crop failure, drought, flooding and so on.

"Economic factors provide the main motivation behind migration. In fact, according to the International Labor Organization, approximately half of the total population of current international migrants, or about 100 million migrant workers, have left home to find better job and lifestyle opportunities for their families abroad (International Labor Office of the Director-General, 2008). In some countries, jobs simply do not exist for a great deal of the population. In other instances, the income gap between sending and receiving countries is great enough to warrant a move. India, for example, has recently experienced a surge in emigration due to a combination of these factors (Index Mundi 2012)."(Globalization 101 (n.d.), Push Factors, retrieved from http://www.globalization101.org/push-factors)

All these reasons can be classified as economic, social, political or environmental:

"Economic migration - moving to find work or follow a particular career path; Social migration - moving somewhere for a better quality of life or to be closer to family or friend; Political migration - moving to escape political persecution or war; Environmental causes of migration include natural disasters such as flooding."(BBC (n.d.), Migration trends, retrieved from http://www.bbc.co.uk/schools/gcsebitesize/geography/migration/migration_trends_rev2.shtml)

2. Introduction

BBC[4] reports four reasons why people migrate which are, “economic migration, social migration, political migration and environmental causes of migration.” (BBC (n.d.), Migration trends, http://www.bbc.co.uk/schools/gcsebitesize/geography/migration/migration_trends_rev2.shtml) In this study we want to use statistical data analysis to check these factors.

This report is intended to check the factors that affect migration which can be used for common people, since only a small fraction of people in the world face political persecution problems and few countries are still fighting wars right now. Factors of political migration cannot stand for the common people, which is why we dropped this factor in our research.

By searching on the internet, there are three indexes which can fit the condition to reflect the truth for the three factors from BBC (Reference 5, 6, 7): income level stands for the economic factor, human development index is used for the social influence and environmental causes can be based off of environmental performance index. Within the three variables, income level and human development index are categorical variables, there are four income levels and four human development index levels, and level 1 to 4 are from high to low; environmental performance index is a continuous variable.

Income level (IL): “Income levels show the income category of a particular country as identified by the World Bank. Income levels themselves just have a name and an id code.”(THE WORLD BANK (n.d.), Income Level Queries, retrieved from http://data.worldbank.org/node/207)

Human development index (HDI): “The Human Development Index (HDI) is a composite measure of health, education and income that was introduced in the first Human Development Report in 1990 as an alternative to purely economic assessments of national progress, such as GDP growth. It soon became the most widely accepted and cited measure of its kind, and has been adapted for national use by many countries. HDI values and rankings in the global Human Development Report are calculated using the latest internationally comparable data from mandated international data providers. Previous HDI values and rankings are retroactively recalculated using the same updated data sets and current methodologies, and are presented in Table 2 of the Statistical Annex of the 2013 Report.”(UNITED NATIONS DEVELOPMENT PROGRAMME (n.d.), Human Development Index, retrieved form http://hdr.undp.org/en/statistics/hdi/)

Environmental performance index (EPI): “The environmental performance index countries on performance indicators tracked across policy categories that cover both environmental public health and ecosystem vitality. These indicators provide a gauge at a national government scale of how close countries are to established environmental policy goals.”(IChemE (n.d.), Organisations, https://www.icheme.org/communities/special-interest-groups/sustainability/resources/organisations.aspx)

We gathered data from 127 countries from all over the world. The data for population, migration and income level comes from World Data Bank, the data for human development index (HDI) is from United Nations Development Program, and the environmental performance index (EPI) information was from Yale University.

3. Methods

Due to the fact that populations are different in different countries, if we just compare the population of migration to analyze the factors, the results may be inaccurate. To counter this in this study we use the ratio = (number of migration / population of countries) as the outcome variable which is continuous to analyze the data.

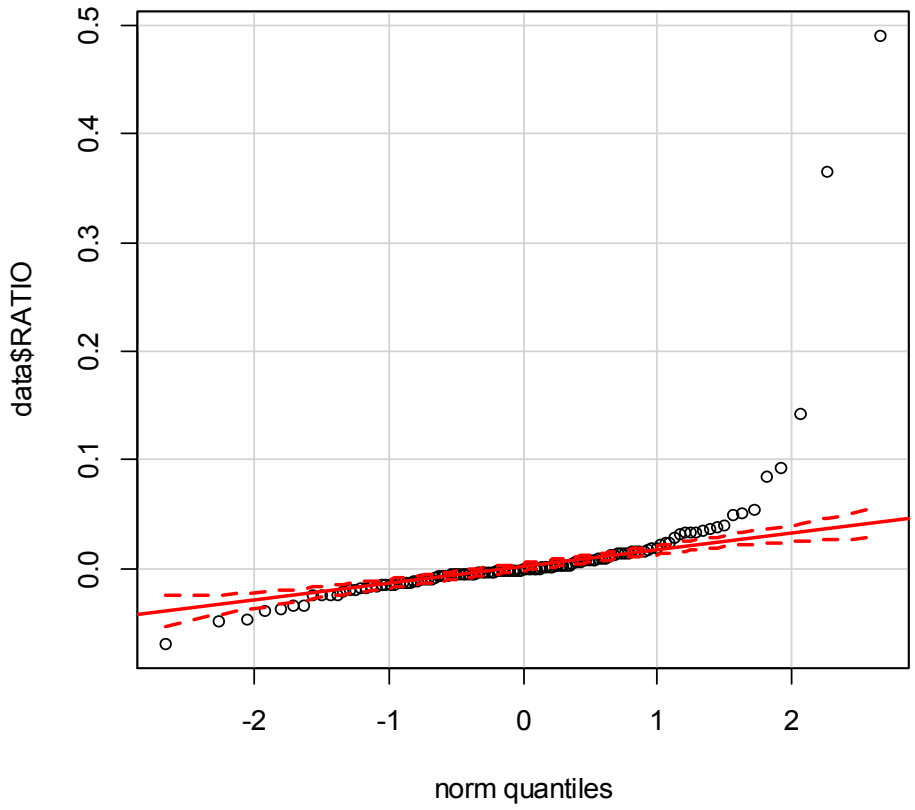
We want to analyze three factors that affect the ratio: income level (IL) and human development index (HDI) are two categorical variances, with every one being divided the into four groups for the countries; environmental performance index (HPI) is a kind of continuous data.

After data transformation, we put all the three factors into a multiple linear regression, and then analyzed the factors by different models depending on the model's results. To explore the effect of IL and HDI we used two-way ANOVA to compare the means in the different groups of each factor, and make sure whether the means in the different levels of countries are significantly different. If so we can say that the factor is a variable affecting migration for common people in the world. Then by using dummy variables analysis we can compare the actual differences between different groups in different factors. As EPI is continuous, from the plots we can make sure it is appropriate to use quadratic regression to analyze the relationship between the environmental performance index and ratio of migration.

4. Data transformation

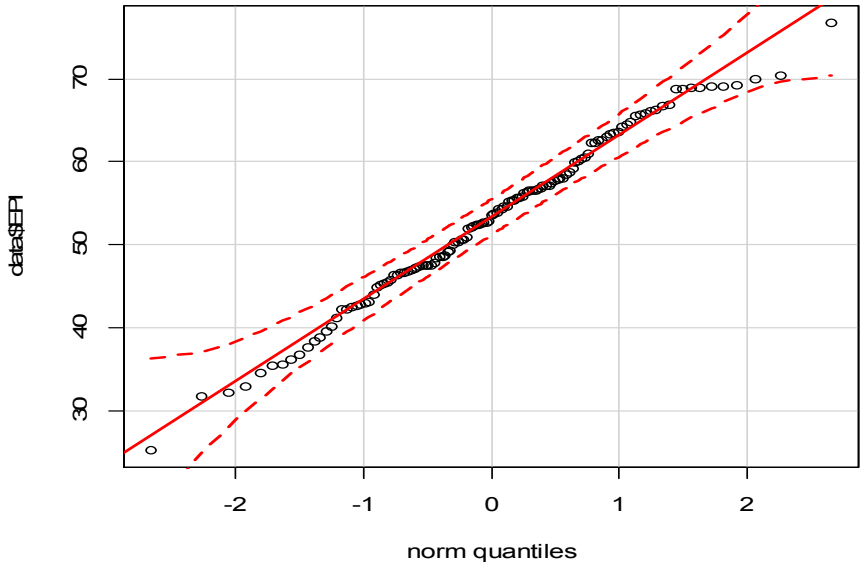
From the QQ plot in Figure 1 (The solid line is a reference line automatically, and the broken lines are 95% confidence intervals for the plot), it is obvious to see that RATIO is not normal, p-value < 2.2e-16 in the Shapiro-Wilk normality test, which means RATIO is far from normal. There are a few outliers in this plot and we cannot delete these outliers. To get the better results of the analysis, we need to transform RATIO before doing other statistical analysis.
RATIO = migration / population

Figure 1: QQ plot of RATIO



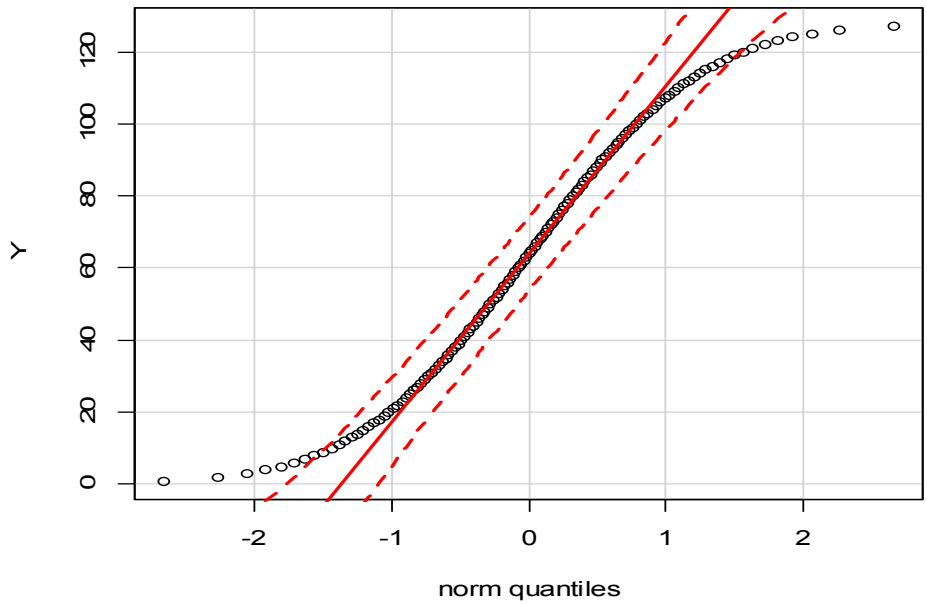
In Figure 2, the p-value of normality test is 0.5382, it is larger than 0.05, we can say EPI is normal and we can use EPI without transformation in the further model analysis.

Figure 2: QQ plot of EPI



For reducing the outliers' effect, we use rank transformation for RATIO (Y=rank RATIO) and EPI (EPI1=rank EPI), as EPI is normal in Figure 2, we can compare the different models by using EIP and rank EPI. The transformed data in the QQ plot is shown by Figure 3 below. In rank RATIO, higher rank numbers on behalf of higher immigration populations than migration populations in these countries.

Figure 3: QQ plot of rank transformation of RATIO



6. Data analysis

6.1 The analysis for the three factors together

6.1.1 Descriptive summary statistics

For exploring the relationship between different income levels by rank ratio, we can use a box plot to give a more descriptive view of the data which is more intuitive. From Figure 4, it is obvious that the rank ratio in group 1 (102.78) is higher than other groups, there are more people immigrating to these countries than migrating from these countries. The range of rank ratio is the biggest in group 2 (110), because the situations are very different in these countries in the second income level, some are similar to the countries in group 1 and more countries are the same as group 3 and group 4. For the last two groups, the range of group 4 is larger than group 3, the strange result is that the mean of group 3 is the smallest, even less than group 4. The migration trend in the third group is the worst of the four groups. Maybe there are other conditions that stop the people in group 4 from migrating from their countries, they can be religion, political systems or any other conditions.

Figure 4: Box plot of rank ratio in the different income levels

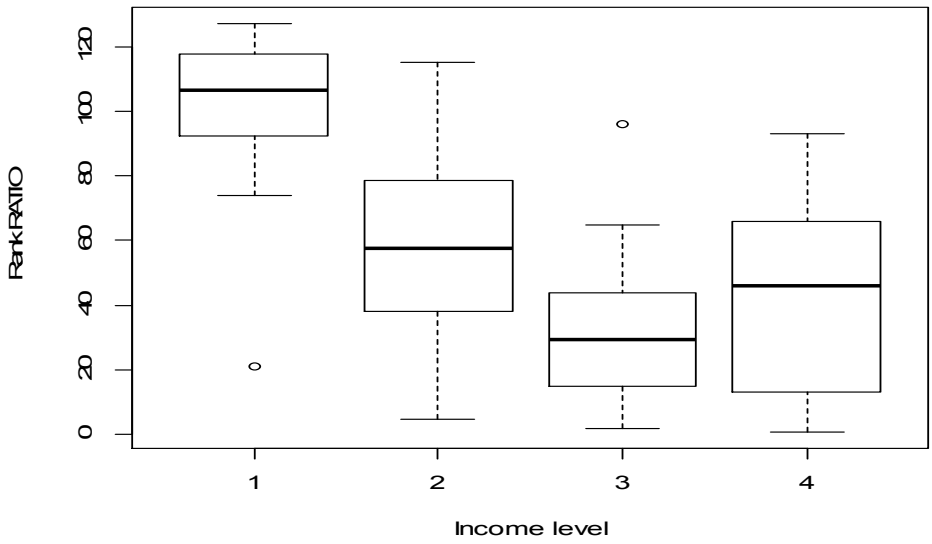


Table 1: Summary statistic of rank ratio in the different income levels

Income level	Number	Mean	Median	Min	Max	Range	Sd
1	40	102.78	106.5	21	127	106	20.03
2	40	57.98	57.5	5	115	110	26.99
3	30	32.97	29.5	2	96	94	22.03
4	17	41.71	46	1	93	92	29.17

The results in the different human development index levels are shown by Figure 5 and Table 2. The average in the group 1 is the largest by 98.03 and the least mean is in the third group (41.43). The ranges of group 2 and group 3 are very large which are respectively 119 and 113. The differences between group 2 and group 4 are not large, the mean of the three groups are 54.17, 41.47 and 48.5, and the second strong tendency of immigrants is the second group. The least tendency of immigrants is still group 3, which is similar to the income level analysis.

Figure 5: Box plot of rank ratio in the different human development index levels

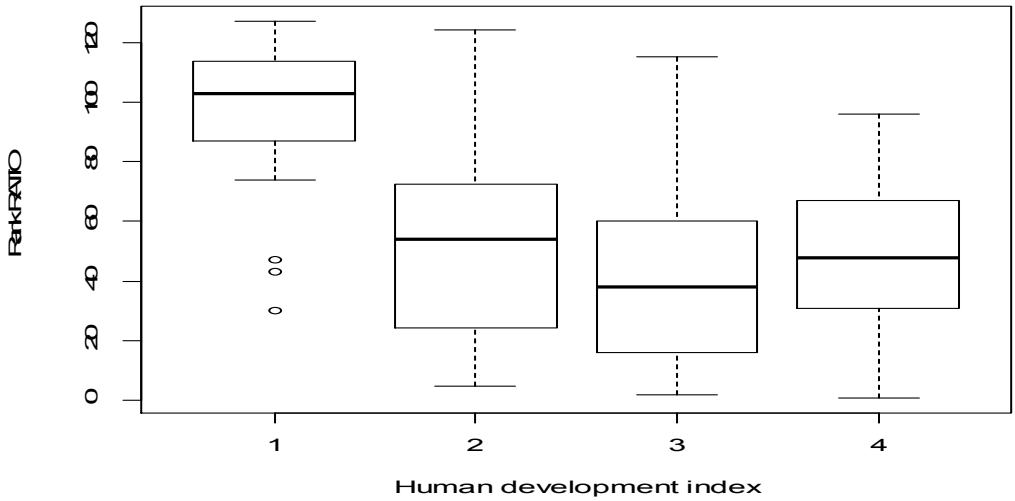


Table 2: Summary statistic of rank ratio in the different human development index levels

Income level	Number	Mean	Median	Min	Max	Range	Sd
1	40	98.03	103	30	127	97	22.51
2	35	54.17	54	5	124	119	33.08
3	30	41.47	38	2	115	113	30.93
4	22	48.5	48	1	96	95	27.4

By comparing Figure 4 and Figure 5 overall the results for different levels of human development index can be seen as similar to the results of the income levels. Group 1 is the immigration countries group, group 2 is the steady group, and group 3 and 4 are the migration group with more people in the third group choosing to immigrate abroad than the last group.

Table 3: Summary statistic of environmental performance index

N	Mean	Sd	Median	Min	Max	Range
127	53.04	9.94	53.55	25.32	76.69	51.37

The summary information is shown in Table 3, the mean of EPI is 53.04, it is similar to the median 53.55, and the standard deviation is 9.94.

6.1.2 Multiple linear regression

To analyze the three factors that affect rank ratio we use a multiple linear regression at first and the summary of the model information as below in Table 3.

Table 4: Multiple linear regression of rank ratio

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	102.3927	19.9319	5.137	1.06E-06	***
IL	-32.0755	4.865	-6.593	1.14E-09	***
HDI	9.3973	4.9565	1.896	0.0603	.
EPI	0.1981	0.2958	0.67	0.5042	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 26.96 on 123 degrees of freedom

Multiple R-squared: 0.4761, Adjusted R-squared: 0.4634

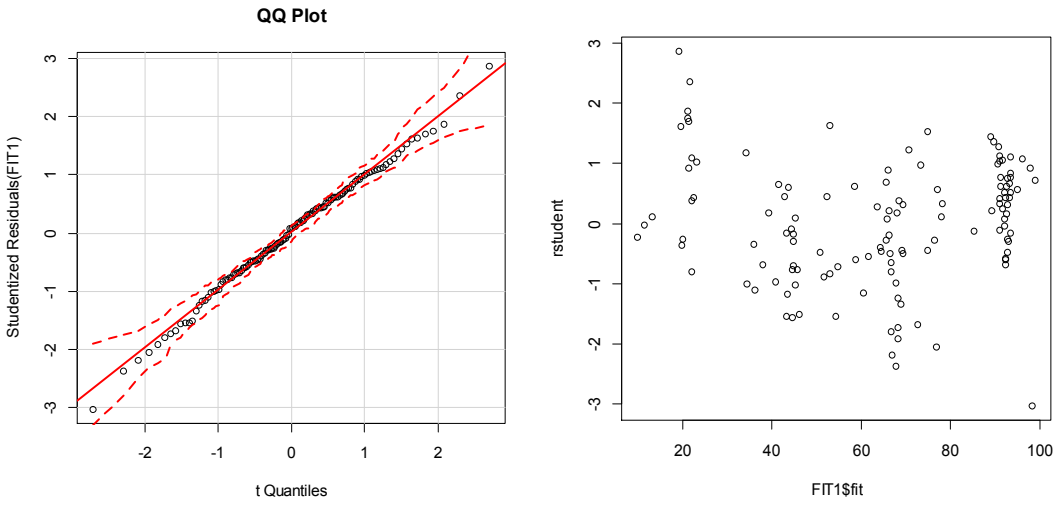
F-statistic: 37.26 on 3 and 123 DF, p-value: < 2.2e-16

The estimated model: E (estimated rank ratio) = 102.39 – 32.08*IL + 9.40*HDI + 0.20*EPI

If the other factors are fixed, the estimated mean of rank ratio will decrease 32.08 for every unit income level increase; if IL and EPI are fixed, the estimated mean of rank ratio will increase 9.40 for every unit of human development index level increase; if IL and HDI are fixed, the estimated mean of rank ratio will increase 0.20 for every unit of environmental performance index increase.

From Table 3 we can see that the t-test results for the three coefficients are very different. IL is the most statistically significant followed by HDI, the p-value of EPI is 0.50>0.1, we can say that income level and human development index can influence migration with statistical significant and environmental performance index cannot influence migration apparently.

Figure 6: QQ plot for studentized residual and studentized residuals plot



The p-value of Shapiro-Wilk normality test is 0.97, so the standardized residual is normal. From the outliers test there is no studentized residuals with Bonferonni p < 0.05, and from the two plots in Figure 6 we can say that the model can fit the data well.

On the other hand, we want to check whether the rank transformed EPI (EPI1) in the same model can get the same results. In Table 4, the results are the multiple linear regression of rank ratio with rank EPI.

The estimated model: E (estimated rank ratio) = 110.24 – 32.01*IL + 9.22*HDI + 0.04*EPI1

This model is similar to the multiple linear regression of rank ratio with EPI. IL is the most statistically significant followed by HDI. EPI rank is not a significant factor in the model and in this model the coefficient of EPI rank (0.04) is smaller than the coefficient of EPI (0.20) in the former model.

Table 5: Multiple linear regression of rank ratio & rank EPI

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	110.2429	10.50514	10.494	< 2e-16	***
IL	-32.0123	4.86474	-6.58	1.21E-09	***
HDI	9.22189	4.97725	1.853	0.0663	.
EPI1	0.04561	0.08125	0.561	0.5756	

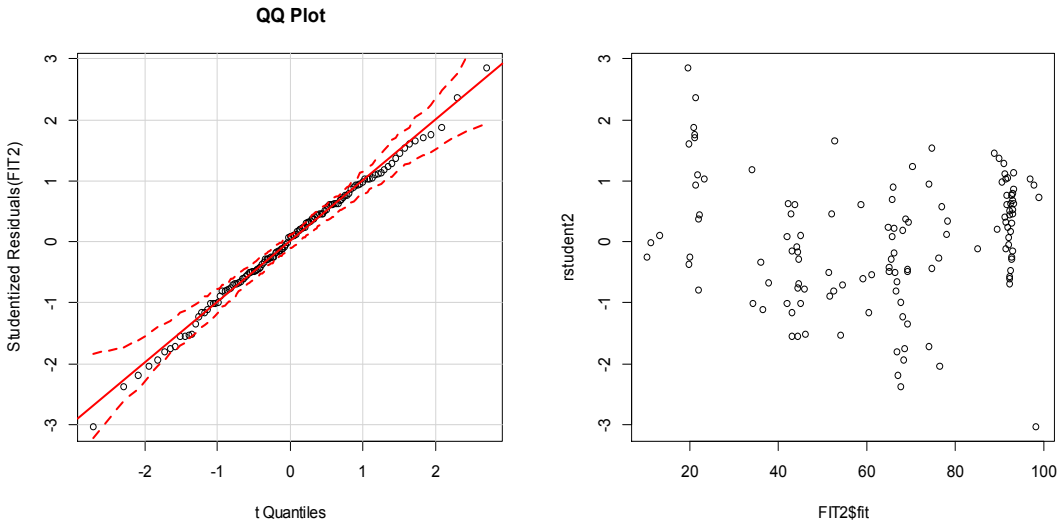
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 26.98 on 123 degrees of freedom

Multiple R-squared: 0.4756, Adjusted R-squared: 0.4628

F-statistic: 37.18 on 3 and 123 DF, p-value: < 2.2e-16

Figure 7: QQ plot for studentized residual and studentized residuals plot



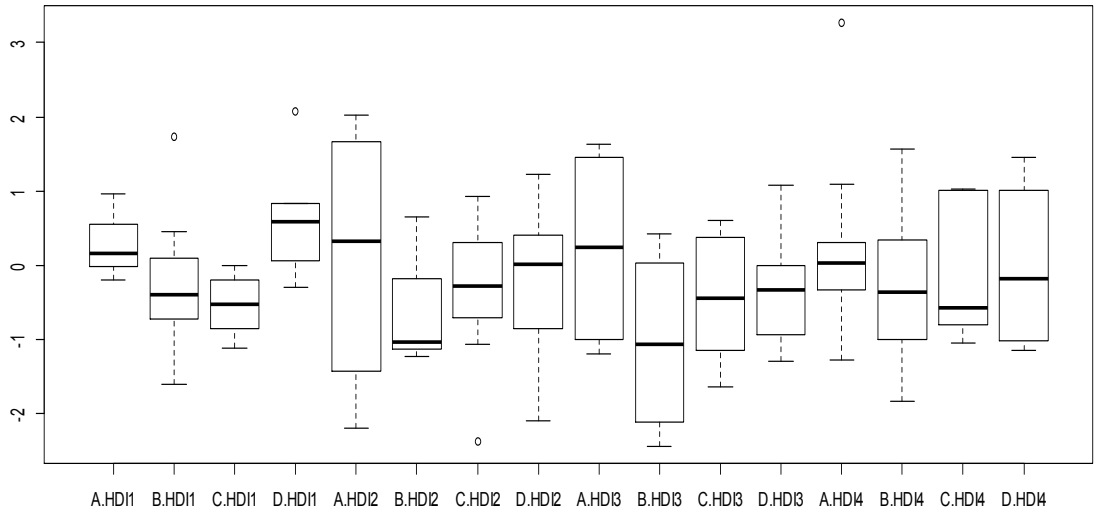
Therefore, we want to analyze the three factors into two parts in the future models. First we want to know the exact effects of the two significant variables on migration, then we want to test if there any other relationship between EPI and migration.

6.2 Income level and human development index level

6.2.1 Two-way ANOVA

To compare the means of rank ratio in different income levels and different human development index levels, two-way ANOVA is the best to solve this problem. In Figure 8, we can see that in every HDI level, the mean rank ratio of IL 1 is the biggest, but other trends are not significant as seen from the plot. To assess whether the mean rank ratio in income level 1, 2, 3, 4 is different based on levels of human development index are statistically different, we can use more models and statistical analysis as below.

Figure 8: Box plot of rank ratio of income levels and human development index levels (A: IL1 B: IL2 C: IL3 D: IL4)



The p-value for IL and HDI are <0.0001 and 0.0424 (Table 6). Since they are all smaller than 0.1, we can say that in the different income level groups the values of rank ratio are significantly different, and human development index can effect rank ratio significantly. In addition, the interaction between IL and HDI is significant (p-value<0.0001). From this it can be seen that the analysis supports the claim that the income level factor and the human development factor are two significant factors effecting migration for countries in the world.

Table 6: Two-way ANOVA with interaction

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
IL	1	78655	78655	144.514	<2e-16	***
HDI	1	2289	2289	4.205	0.0424	*
IL*HDI	1	22799	22799	41.890	2.07e-09	***
Residuals	123	66945	544			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.2.2 Dummy variables analysis

To analyze how the two factors in different levels influence migration, we create six dummy variables in order to do the multiple regression.

We create three dummy variables for income level: IL1, IL2, IL3 (Table 7), and three dummy variables HID1, HDI2, HDI3 (Table 8) for human development index level. Then they are used in the regression procedure to fit a two-way ANOVA main effects model.

Table 7: Dummy variables of income level

Income level	IL1	IL2	IL3
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1

Table 8: Dummy variables of human development index level

Human development index level	HDI1	HDI2	HDI3
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1

At first we fit the model with all interactions:

E (estimated rank ratio) = β0+β1IL1+β2IL2+β3IL3+β4HID1+β5HID2+β6HID3+β7IL1*HDI1+β8IL1*HDI2+β9IL1*HDI3+β10IL2*HID1+β11IL2*HID2+β12IL2*HID3+β13IL3*HDI1+β14IL3*HDI2+β15IL3*HDI3

From the output of R, it is not appropriate to use this model because of singularities. To solve this problem, we define each of IL and HDI into two groups, the HI part (level 1 and level 2) and the LO part (level 3 and level 4), then:

Z1 = 1 if the individual is HI for IL and HI for HDI

Z2 = 1 if the individual is HI for IL and LO for HDI

Z3 = 1 if the individual is LO for IL and HI for HDI

As any one of these Z's can be perfectly predicted (R^2=1) from the six IL and HDI variables, together with any one of the other Z's, we can get the same result by fitting any one of the Z's.

Here we can see the result of the model by using Z1:

E (estimated rank ratio) = β0+β1IL1+β2IL2+β3IL3+β4HID1+β5HID2+β6HID3+β7Z1

Table 9: Multiple regression of dummy variables with interaction

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	119.001	14.92	7.976	1.04E-12	***
IL1	-48.995	8.766	-5.589	1.47E-07	***
IL2	-92.106	12.449	-7.399	2.12E-11	***
IL3	-95.721	14.032	-6.822	3.99E-10	***
HDI1	-1.495	8.766	-0.171	0.865	
HDI2	2	14.569	0.137	0.891	
HDI3	21.946	15.513	1.415	0.16	
Z1	-16.076	14.36	-1.12	0.265	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23.3 on 119 degrees of freedom

Multiple R-squared: 0.6214, Adjusted R-squared: 0.5991

F-statistic: 27.9 on 7 and 119 DF, p-value: < 2.2e-16

The p-value for the interaction term is 0.265 (Table 9), since Z1 is not significant we can delete the interaction term in the future studies. For the regressive procedure these six dummy variables are the predictor variables, the formula for the regression model is:

E (estimated rank ratio) = β0+β1IL1+β2IL2+β3IL3+β4HID1+β5HID2+β6HID3

Table 10: Multiple regression of dummy variables

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	102.8446	3.7908	27.13	< 2e-16	***
IL1	-48.1964	8.7465	-5.51	2.08E-07	***
IL2	-85.2161	10.8327	-7.867	1.79E-12	***
IL3	-90.0325	13.0938	-6.876	2.96E-10	***
HDI1	-0.6964	8.7465	-0.08	0.9367	
HDI2	13.2139	10.5901	1.248	0.2146	
HDI3	32.2537	12.4985	2.581	0.0111	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

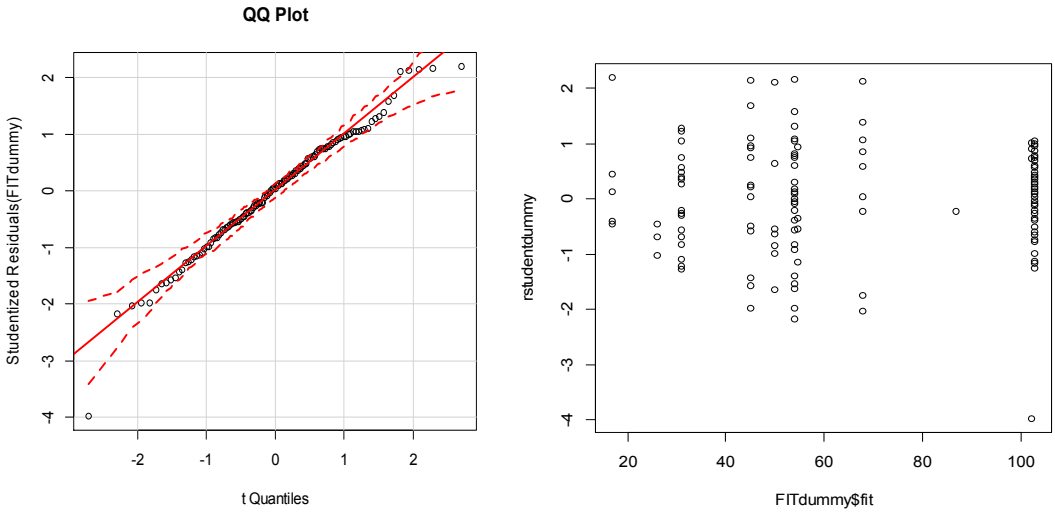
Residual standard error: 23.33 on 120 degrees of freedom

Multiple R-squared: 0.6174, Adjusted R-squared: 0.5983

F-statistic: 32.27 on 6 and 120 DF, p-value: < 2.2e-16

The p-value of Shapiro-Wilk studentized residual normality test is 0.08 > 0.05. It is seen that the residual is normal and the diagnostic plots are shown in Figure 9. Except for one point the plot can be seen as a model that fits the data well.

Figure 9: QQ plot for studentized residual and studentized residuals plot



In Table 9, the coefficients of constant, IL1, IL2, IL3, HDI1, HDI2, HDI3 are β0= 102.84, β1= -48.20, β2= -85.22, β3= -90.03, β4= -0.70, β5= 13.21, β6= 32.25, as H0: β=0. The p-values for IL1, IL2, IL3 HDI3 and constant are all smaller than 0.05 which means we should reject H0. They are statistically significant influencing rank ratio in the model; the p-values for HID1, HDI2 (β4, β5) are 0.94 and 0.21 which are bigger than 0.05. This means that these two variables have no significant effect for rank ratio. For the full model, there are four significant variables out of six, and three are from income level variables. From those results we can say that income level is important to the action of migration. The estimated mean of rank ratio for each income level in every human development index level is shown in Table 10.

Table 11: Mean of income level for human development index level

MEAN	Human development index level			
Income level	1	2	3	4
1	β0	β0+β4	β0+β5	β0+β6
	102.84	102.14	116.05	135.09
2	β0+β1	β0+β1+β4	β0+β1+β5	β0+β1+β6
	54.64	53.94	67.85	86.89
3	β0+β2	β0+β2+β4	β0+β2+β5	β0+β2+β6
	17.62	16.92	30.83	49.47
4	β0+β3	β0+β3+β4	β0+β3+β5	β0+β3+β6
	12.81	12.11	26.02	45.06

From the table we can see that β1= -48.20, β2= -85.22, β3= -90.03 are the mean difference of migration in the income level 2, 3, 4 relative to the level 1; β4= -0.70, β5= 13.21, β6= 32.25 are the mean difference of migration in the human development index level 2, 3, 4 relative to the level 1. As β4 and β5 are not significant, the mean differences of migration in the human development index 2, 3 relative to the level 1 are not statistical significant different.

It is obvious that people would like to immigrate into the countries of income level 1 and human development index level 4; more people choose to migrate from the countries of income level 4 and human development index level 1. For the same income level, more people choose to immigrate into the lower human development index level countries, but the difference is not large in the income level 1 and 2; for the same human development index level, the immigration trends countries in the income level 1 is significant different from other income levels, almost all countries in other three levels show the migration trends, and the situations are not too much difference between income level 3 and 4.

6.3 Environmental Performance Index

6.3.1 Scatter plot

Based on the fit line in the scatter plot (Figure 10), the relationship between rank ratio and environmental performance index can be considered as quadratic regression. Compared with Figure 11, the two plots show almost the same smooth fit line. That is why we try to use quadratic regression to fit rank ratio by EPI and rank EPI (EPI1) in following model analysis.

Figure 10: Scatter plot of Signed root of the ratio vs. environmental performance index

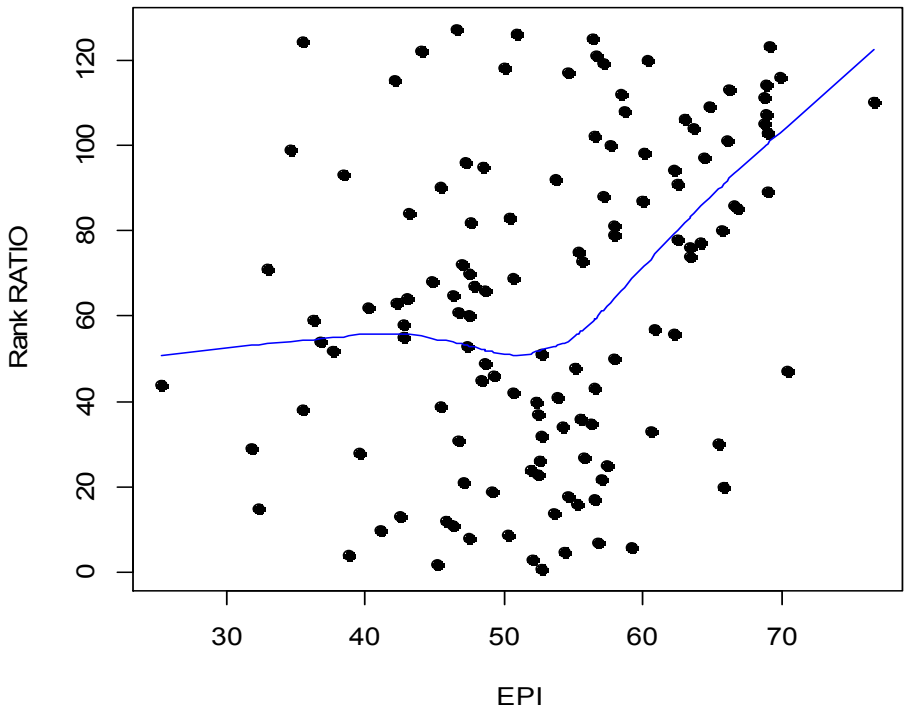
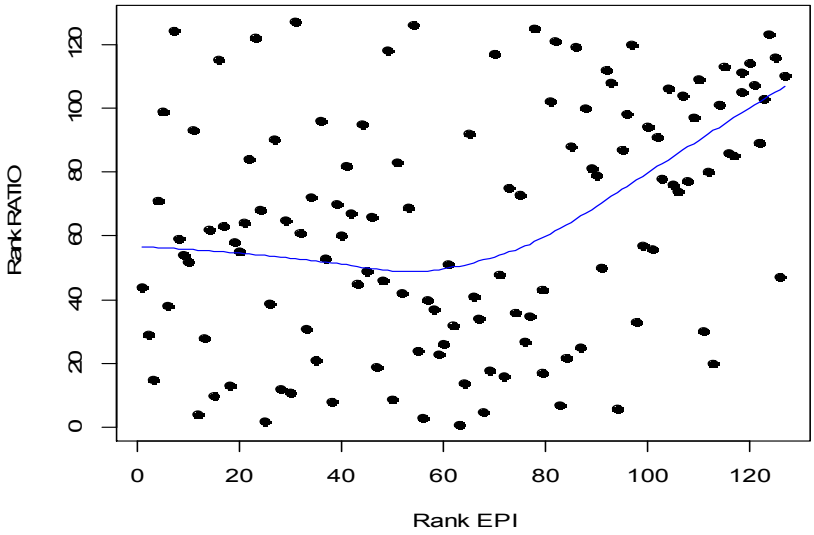


Figure 11: Scatter plot of Signed root of the ratio vs. Rank EPI



6.3.2 Quadratic regression model

Quadratic regression model: Y = Ax + BX^2 + C (Y is estimated mean of rank ratio, X is EPI)

A = -5.40

B = 0.06

C= 166.91

E (estimated rank ratio) = 166.91 -5.40*EPI + 0.06*EPI^2

The estimated quadratic regression model is above. The p-value for the model is 5.328e-05 which is smaller than 0.05, so that we can say this quadratic regression model can explain the transformed data properly. As R-squared is 0.15, only 15% of rank ratio is explained by the model, which means that the data can be explained by more factors together.

Table 12: Simple linear regression of Signed root of the ratio & environmental performance

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	166.9101	65.89637	2.533	0.0126	*
data$EPI	-5.39779	2.55571	-2.112	0.0367	*
I(data$EPI^2)	0.06299	0.02432	2.59	0.0107	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 34.27 on 124 degrees of freedom

Multiple R-squared: 0.1468, Adjusted R-squared: 0.133

F-statistic: 10.66 on 2 and 124 DF, p-value: 5.328e-05

E (estimated rank ratio) = 64.28 -0.65*EPI1 + 0.01*EPI1^2

The results of the model with rank EPI are similar to the model with EPI, except the coefficients are different, R-squared is still small, so the conclusion for this model is the same as the model with EPI: this model fit the data properly.

Table 13: Simple linear regression of Signed root of the ratio & Rank EPI

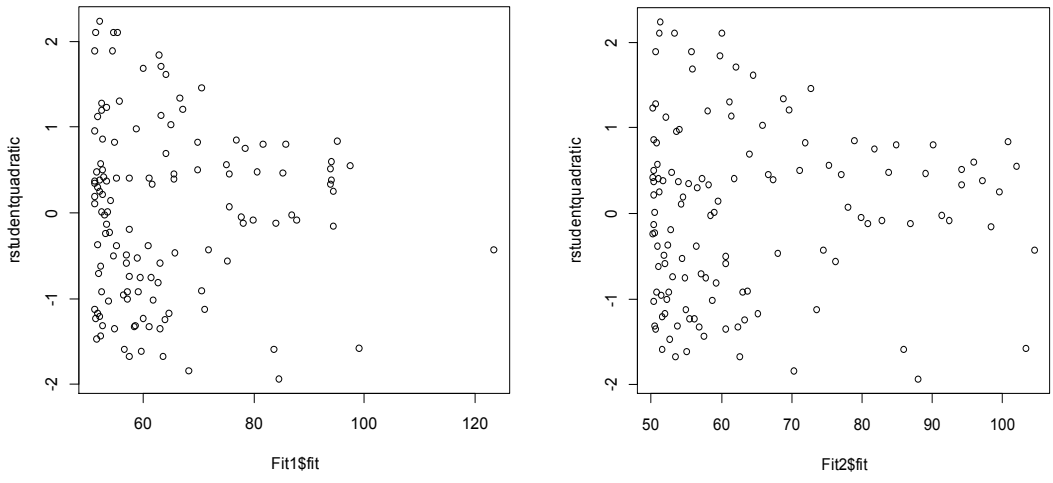
Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	64.282805	9.150945	7.025	1.25E-10	***
EPI1	-0.654846	0.330049	-1.984	0.04945	*
I(EPI1^2)	0.007652	0.002498	3.063	0.00268	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 33.84 on 124 degrees of freedom
Multiple R-squared: 0.1683, Adjusted R-squared: 0.1549
F-statistic: 12.55 on 2 and 124 DF, p-value: 1.09e-05

6.3.3 Examine the assumptions of linear regression via Diagnostic Plots

After fitting the models, one can check the studentized residuals plots in Figure 12 as below, the two models can be used to fit the data, the second one is better. The p-values of normality test show that the first model is 0.0353 and the second model is 0.1294, so that in the second model the studentized residuals are normal, but the first one are not. According to the comparison, we think the model with rank EPI is better than the model with EPI to fit rank ratio in a quadratic regression model.

Figure 12: Diagnostic plots of the two quadratic regression models



7. Conclusions

Based on these analysis,

- Income level and human development index are two statistically significant factors for rank ratio, they can influence the differences between the countries' migration status.
- People would like to immigrate into the countries of income level 1 and human development index level 4; more people choose to migrate from the countries of income level 4 and human development index level 1. For the same income level, more people choose to immigrate into the lower human development index level countries, but the difference is not large in the income level 1 and 2; for the same human development index level, the immigration trends countries in the income level 1 is significant different from other income levels, almost all countries in other three levels show the migration trends, and the situations are not too much difference between income level 3 and 4.
- We can use quadratic regression to estimate the effect of environmental performance on rank ratio.
- The model to estimate the effect of environmental performance on rank ratio by using rank environmental performance index is better than the model with environmental performance index.
- The quadratic regression model with rank environmental performance index is E (estimated rank ratio) = 64.28 -0.65*EPI1 + 0.01*EPI1^2, the relationship between rank ratio and environmental performance index is quadratic.

8. References

1. EMBARGOED (2013, Sep 11), 232 million international migrants living abroad worldwide – New UN global migration statistics reveal, UN PRESS RELEASE, retrieved from http://esa.un.org/unmigration/wallchart2013.htm

2. Globalization 101 (n.d.), Push Factors, retrieved from http://www.globalization101.org/push-factors

3. BBC (n.d.), Migration trends, retrieved from http://www.bbc.co.uk/schools/gcsebitesize/geography/migration/migration_trends_rev2.shtml

4. The British Broadcasting Corporation (BBC) is a British public service broadcasting statutory corporation. Its main responsibility is to provide impartial public service broadcasting in the United Kingdom, the Channel Islands, and the Isle of Man. The BBC is headquartered at Broadcasting House in London and has major production centres in Salford Quays, Belfast, Birmingham, Bristol, Cardiff and Glasgow, and smaller production centres throughout the UK. The BBC is the world's oldest national broadcasting organisation and the largest broadcaster in the world by number of employees, with about 23,000 staff. http://en.wikipedia.org/wiki/BBC

5. THE WORLD BANK (n.d.), Income Level Queries, retrieved from http://data.worldbank.org/node/207

6. UNITED NATIONS DEVELOPMENT PROGRAMME (n.d.), Human Development Index, retrieved form http://hdr.undp.org/en/statistics/hdi/

7. IChemE (n.d.), Organisations, retrieved from https://www.icheme.org/communities/special-interest-groups/sustainability/resources/organisations.aspx

8. Levitt, Peggy, and Ninna Nyberg-Sørensen. "Global Migration Perspectives."Global migration perspectives (2004).

9. Hagen-Zanker, Jessica. "Why do people migrate? A review of the theoretical literature." A Review of the Theoretical Literature (January 2008). Maastrcht Graduate School of Governance Working Paper No (2008).

10. Massey, Douglas S., et al. "Theories of international migration: a review and appraisal." Population and development review (1993): 431-466.

11. Johnson, Richard Arnold, and Dean W. Wichern. Applied multivariate statistical analysis. Vol. 5. No. 8. Upper Saddle River, NJ: Prentice hall, 2002.

12. Conover, William J., and Ronald L. Iman. "Rank transformations as a bridge between parametric and nonparametric statistics." The American Statistician35.3 (1981): 124-129.

13. Rousseeuw, Peter J., and Annick M. Leroy. Robust regression and outlier detection. Vol. 589. John Wiley & Sons, 2005.

14. Iman, Ronald L., and William J. Conover. "The use of the rank transform in regression." Technometrics 21.4 (1979): 499-509.

15. Cox, Trevor F., et al. An introduction to multivariate data analysis. London: Hodder Arnold, 2005.

9. Appendix

Data:

	Country	IL	RATIO	HDI	EPI
1	ALB	3	-0.0152	2	65.85
2	DZA	2	-0.00378	2	48.56
3	AGO	2	0.004195	4	47.57
4	ARG	2	-0.00495	1	56.48
5	ARM	3	-0.02531	2	47.48
6	AUS	1	0.050969	1	56.61
7	AUT	1	0.019071	1	68.92
8	AZE	2	0.005883	2	43.11
9	BGD	4	-0.01924	4	42.55
10	BLR	2	-0.00527	2	53.88
11	BEL	1	0.018356	1	63.02
12	BEN	4	0.005258	4	50.38
13	BOL	3	-0.01626	3	54.57
14	BIH	2	-0.0026	2	36.76
15	BWA	2	0.009511	3	53.74
16	BRA	2	-0.00256	2	60.9
17	BRN	1	0.008738	1	62.49
18	BGR	2	-0.00664	2	56.28
19	KHM	4	-0.01775	3	55.29
20	CMR	3	-0.00092	4	42.97
21	CAN	1	0.032187	1	58.41
22	CHL	2	0.001749	1	55.34
23	CHN	2	-0.00141	3	42.24
24	COL	2	-0.00258	2	62.33
25	COD	4	-0.00039	4	47.49
26	COG	3	0.012129	4	47.18
27	CRI	2	0.01619	2	69.03
28	CIV	3	-0.01897	4	53.55
29	HRV	1	0.002264	1	64.16
30	CUB	2	-0.01685	2	56.48
31	CYP	1	0.040017	1	57.15
32	CZE	1	0.022858	1	64.79
33	DNK	1	0.01628	1	63.61
34	DOM	2	-0.01398	3	52.44
35	ECU	2	-0.008	2	60.55

36	EGY	3	-0.00444	3	55.18
37	SLV	3	-0.04691	3	52.08
38	ERI	4	0.00958	4	38.39
39	ETH	4	-0.00344	4	52.71
40	FIN	1	0.013543	1	64.44
41	FRA	1	0.007689	1	69
42	GAB	2	0.003213	3	57.91
43	GEO	3	-0.03369	2	56.84
44	DEU	1	0.006726	1	66.91
45	GHA	3	-0.00211	3	47.5
46	GRC	1	0.01362	1	60.04
47	GTM	3	-0.01395	3	51.88
48	HTI	4	-0.02425	4	41.15
49	HND	3	-0.01312	3	52.54
50	HUN	1	0.0075	1	57.12
51	ISL	1	0.032754	1	66.28
52	IND	3	-0.00249	3	36.23
53	IDN	3	-0.00537	3	52.29
54	IRN	2	-0.00249	2	42.73
55	IRQ	3	-0.00485	3	25.32
56	IRL	1	0.02235	1	58.69
57	ISR	1	0.035893	1	54.64
58	ITA	1	0.033049	1	68.9
59	JAM	2	-0.03702	2	54.36
60	JPN	1	0.002118	1	63.36
61	JOR	2	0.033563	3	42.16
62	KAZ	2	0.000428	2	32.94
63	KEN	4	-0.00463	4	49.28
64	KWT	1	0.092803	2	35.54
65	KGZ	4	-0.02415	3	46.33
66	LVA	2	-0.00447	1	70.37
67	LBN	2	-0.00288	2	47.35
68	LBY	2	-0.00336	2	37.68
69	LTU	2	-0.0108	1	65.5
70	LUX	1	0.083773	1	69.2
71	MKD	2	0.000951	2	46.96
72	MYS	2	0.002988	2	62.51
73	MLT	1	0.012019	1	48.51
74	MEX	2	-0.01531	2	49.11
75	MDA	3	-0.04822	3	45.21
76	MNG	3	-0.00553	3	45.37
77	MAR	3	-0.02133	3	45.76

78	MOZ	4	-0.00083	4	47.82
79	MMR	4	-0.00963	4	52.72
80	NAM	2	-0.00069	3	50.68
81	NPL	4	-0.00372	4	57.97
82	NLD	1	0.00301	1	65.65
83	NZL	1	0.014883	1	66.05
84	NIC	3	-0.03435	3	59.23
85	NER	4	-0.00179	4	40.14
86	NOR	1	0.035022	1	69.92
87	OMN	1	0.05459	2	44
88	PAK	3	-0.01155	4	39.56
89	PAN	2	0.002991	2	57.94
90	PRY	3	-0.00619	3	52.4
91	PER	2	-0.02478	2	50.29
92	PHL	3	-0.0132	3	57.4
93	POL	1	0.001457	1	63.47
94	PRT	1	0.014101	1	57.64
95	QAT	1	0.489846	1	46.59
96	ROU	2	-0.00466	2	48.34
97	RUS	2	0.007976	2	45.43
98	SAU	1	0.038723	2	49.97
99	SEN	3	-0.01026	4	46.73
100	SGP	1	0.142167	1	56.36
101	SVK	1	0.006756	1	66.62
102	SVN	1	0.010739	1	62.25
103	ZAF	2	0.014002	3	34.55
104	ESP	1	0.048838	1	60.31
105	LKA	3	-0.0121	2	55.72
106	SWE	1	0.028326	1	68.82
107	CHE	1	0.023362	1	76.69
108	SYR	3	-0.00259	3	42.75
109	TJK	4	-0.03882	3	38.78
110	TZA	4	-0.00667	4	54.26
111	THA	2	0.007413	3	59.98
112	TGO	4	-0.00086	4	48.66
113	TTO	1	-0.01491	2	47.04
114	TUN	2	-0.0019	2	46.66
115	TUR	2	-0.00069	2	44.8
116	TKM	2	-0.01081	3	31.75
117	UKR	3	-0.00087	2	46.31
118	ARE	1	0.364464	1	50.91
119	GBR	1	0.016383	1	68.82

120	USA	1	0.016018	1	56.59
121	URY	2	-0.01483	2	57.06
122	UZB	3	-0.01815	3	32.24
123	VEN	2	0.001377	2	55.62
124	VNM	3	-0.00495	3	50.64
125	YEM	3	-0.00593	4	35.49
126	ZMB	3	-0.00643	4	55.56
127	ZWE	4	-0.06882	4	52.76