**Investigating Clean Energy Consumption**

Group Members: Yash Soogoor, Macon Magee, Ryan O'Connell

ECON 3161

Instructor: Prof. S. Dhongde

**Abstract**

In order to provide insight into the factors that are related to the clean energy consumption of a nation, simple and multiple regression models were developed to assess the relationship between these factors. The simple regression relating a country's GDP to clean energy production found that countries more economically well off are more likely to consume non-clean energy. The multiple regression analysis showed a similar result for GDP's relation to clean energy production and also showed a negative relationship between a country's percent of GDP contributed from industry and clean energy production.

## Introduction

As a greater number of countries in the world set their eyes on shifting towards using cleaner, more sustainable, and more modern sources of energy, it is becoming increasingly important to determine what factors influence a country's usage of clean energy. In this research project, we seek to study explanatory variables such as a country's GDP and its level of fossil fuel reserves in order to describe a correlation or lack thereof with the percentage of the country's total energy consumption that is clean energy consumption, a ratio which will be the explained variable.

The topic under study is related to the United Nations' Sustainable Development Goal Seven, which is to "ensure access to affordable, reliable, sustainable, and modern energy for all." The reliance on fossil fuels to drive the development of many nations has been a key contributor to the phenomenon of climate change, and as such, a key goal of the UN is to reduce the carbon intensity of energy alongside pursuing a modernization of energy consumption away from usage of sources such as coal, oil, and natural gas. Naturally, the prescription for countries is to increase their level of clean energy consumption as a proportion of their total energy consumption. However, there are different factors based on a country's level of development, resources, and higher education that determine how viable it is for that country to shift to clean energy and how far along in the shift that country has come. Learning about the correlation between these factors such as a country's gross domestic product and level of fossil fuel reserves and a country's level of clean energy consumption can help experts and economists prescribe natural means for a country to shift to clean energy consumption as opposed to forcing policies that may hurt these countries' economies.

For the purpose of this research, clean energy is considered alternate and nuclear energy that is not energy derived from fossil fuels. The underlying assumptions of the research are that most if not all countries of the world are not yet at a place at which clean energy consumption can happen solely due to the free market and that attempts to induce clean energy consumption will have some sort of detrimental effect on the short term economies of countries. Thus, there is a need to study factors that can promote natural tendencies towards consumption of clean energy. It is hypothesized that a country's GDP has a positive correlation with clean energy consumption and that a country's level of fossil fuel reserves has a negative correlation with clean energy consumption.

## Literature Review

The spreading of environmental concerns related to fossil fuel usage has generated interest in the economic implications of switching to more renewable energy sources. Results have varied with differing geopolitical contexts around the globe. Asafu-Adjaye et al (2016) sought to control for this by sorting countries by net importers versus net exporters of fossil fuel, and developed versus developing countries. Asafu-Adjaye examined both short run and long run relationships between both fossil fuel and non-fossil fuel consumption and Gross Domestic Product in both directions for each category of country. For developed exporters, they found bidirectional causality between fossil fuel consumption and real GDP in the short run and long run, and they found a long run bidirectional causality between non-fossil fuel consumption and real GDP. Developed importers saw positive, bidirectional causality in both short term and long term between both fossil fuels and non-fossil fuels, and real GDP. The results of developing exporters showed a positive bidirectional causality between fossil fuel consumption and real GDP and a negative long run causality from real GDP to non-fossil fuel use. Developing importers showed a bidirectional causality between fossil fuel use and real GDP that was positive in the short run and negative in the long run and a positive long term causality from non-fossil fuel use to real GDP. In all scenarios examined, fossil fuel use correlated with real GDP between 1990 and 2012, implying that any restriction of their use by a country may have risked hurting their economy.  As of 2012, we may be too dependent on fossil fuels to quit using them. The relationship between non-fossil fuel use and real GDP varied significantly by category of country, so these divisions in data were helpful. For developed countries increases in income resulted in increases in non-fossil fuel use, but for developing countries this was not the case. Developing exporters decreased their use of non-fossil fuel when real GDP increased and developing importers only saw an increase in real GDP from non-fossil fuel use in the long run. Despite it being more sustainable in the long run, not all countries had an economic incentive to increase non-fossil fuel use as of 2012. Developed countries seem to benefit more economically from switching to fossil fuels than developing countries, which may mean that more funding and access to newer technology may allow clean energy to achieve economic parity with fossil fuels. It appears that it would take more than just the free market to achieve a transition to clean energy at the moment. Asafu-Adjaye does not account for subsidies in the study which leaves room for further research to see if government intervention can help to accelerate a transition to clean energy.

Destek and Aslan (2017) examined the relationship between renewable energy consumption, nonrenewable energy consumption, and real GDP from 1980 to 2012 for 17 emerging countries. In their literature review, they summarized well the results of past studies. Results had been mixed, but mostly pointed to there being causality from total energy use to real GDP, and were mixed about there being causality from renewable energy use to real GDP and from real GDP to total energy use. They found there to be cross-sectional dependence from country to country and there to be country specific heterogeneity. Their study also differed from Asafu-Adjaye et al (2016) in that they did not include nuclear energy in their data. Nuclear energy is interesting because it does not significantly contribute to climate change which is the main motivation for clean energy; however it is not a renewable source of energy. Destek and Aslan found a bidirectional causality between renewable energy consumption and real GDP for Greece and South Korea, a causality from renewable energy consumption to economic growth for Peru, and a causal relationship from real GDP to renewable energy consumption for Colombia and Thailand, and no relationship between renewable energy consumption and real GDP for the other 12 countries. They found a causal relationship from nonrenewable energy use to real GDP for China, Colombia, Mexico, and the Philippines, a causal relationship from real GDP to nonrenewable energy use for Egypt, Peru, and Portugal, a bidirectional causality for Turkey, and no relationship for the remaining 9 countries. This supports the ideas that non-renewable energy is still more important to the economies of most emerging countries examined, and that there are other uncontrolled variable that are affecting the impact of renewable energy. Reducing non-renewable energy consumption could still damage the economies of many emerging countries until they have further expanded their capacity to use renewable sources.

Pao et al (2013) studied the relationship between clean and non-clean energy consumption and and economic growth in the countries of Mexico, Indonesia, South Korea, and Turkey. The results of the paper showed that there is a “positive short-term unidirectional causality” from fossil fuel energy consumption to economic growth with a “bidirectional long-term causality.” They also indicated that there is a long-run unidirectional causality from renewable consumption to fossil fuel consumption alongside positive short-term bidirectional causality with “negative short-run feedback effects”. Lastly, the researchers found that there is a “long-run bidirectional causality between nuclear energy consumption and economic growth and a long-run causality from fossil fuel energy consumption to nuclear energy consumption with positive short-run feedback effects.” The results of this paper seem to suggest that changes in energy consumption away from traditional fossil fuel sources seem to have depressing effects on the economic growth of these nations. One issue with the paper is the question of whether or not the results can be extrapolated to other developing nations or even any other nations in general.

In these papers, the researchers primarily address bidirectional causality between non-clean energy consumption and economic growth. Through our research, we intend to focus more specifically on the factors that affect clean energy consumption, and economic growth of a nation is only a subset of the factors involved in the multiple regression analysis. We intend to focus on GDP as well as study the relationship of clean energy consumption to natural resource reserve levels of a country. While most current literature bases their assessment on the effect of clean or non-clean energy consumption on the economic growth of a country, we are looking at it from the opposite perspective and treating clean energy consumption as the explained variable. The goal is to conduct an analysis that might allow economists and experts to prescribe methods for increasing clean energy consumption, which is a major sustainable development goal for the United Nations.

**Data**

To construct this analysis, a simple linear regression is performed to estimate the percent of non-fossil fuel energy consumption in 2013 (energy). This is the explained variable, or dependent variable, which represents the percentage of a country's total energy consumption that comes from renewable energy or nuclear energy, which together are considered clean energy for the purposes of this analysis. The explanatory, or independent, variable in the simple linear regression is a measure of the corresponding countries' GDP per capita in 2013 (lngdp) in US dollars. A natural logarithm of gross domestic product per capita is used here to clearly indicate the effect of changing GDP by a given percentage on non-fossil fuel energy consumption. The use of lngdp as the independent variable is justified because doing so should precipitate results that provide an understanding of how a nation's development (in terms of GDP per capita) will affect their relative usage of clean energy.

A more rigorous multiple linear regression is performed with more independent variables to estimate the same dependent variable, energy. In this case, lngdp is the first independent variable, and another is added: lnresources. This new independent variable is the natural logarithm of the 2013 sum of a country's rents for coal, natural gas, and oil divided by GDP. Natural resource rents seek to measure the economic profit that a country generates from the resources found within their borders. The World Bank calculates it annually as a percentage of GDP. It is used here to represent impact of being a country that possesses more fossil fuels and will likely have less of a need for alternative energy sources and should impact the percent of non-fossil fuel energy consumption.

A second multiple linear regression was performed using the above mentioned variables and adding edu, which represents the percentage of a country's population with at least some college education. People who are more educated may be more likely to be aware of the dangers of fossil fuels and more capable of finding solutions. A country needs talented scientists and engineers to build modern, clean power plants.

Another independent variable that was tested was ind_perc which represents the percent of value added to gdp from the energy industry in a country. If a country has a strong energy industry then they may be less interested in overhauling it to make room for alternative energy, while on the other hand, they might have more capital with which they can invest in clean energy.

The last independent variable, urban, describes the percentage of a country's population living in urban areas. Urban populations often consume a greater amount of energy than rural ones, and these greater power needs are easier to meet with traditional fossil fuel sources.

The data for all of these variables is taken or derived from the World Bank and is for the year 2013. This year was chosen because it had data available for the most countries while still being recent. energy is taken directly from the data given for percentage of total energy consumption from renewable and nuclear energy for 178 countries listed as an appendix. lngdp uses data given for GDP per capita, and lnresources uses the data given for each of the coal, natural gas, and oil rents added together. The data used to derive lnresources is available for 155 of those 178 countries. edu is taken directly from the World Bank's statistics for 114 of these countries. Provided below are the descriptive statistics of the variables taken from our STATA data set compiling all of these sources.

Table 1. Descriptive Statistics

| Variable Name | Observations | Mean | Std. Deviation | Min | Max |
|---|---|---|---|---|---|
| energy | 178 | 31.21 | 25.82 | 0 | 95.82 |
| lngdp | 178 | 8.87 | 1.35 | 6.03 | 11.64 |
| lnresources | 155 | 0.49 | 2.27 | -7.04 | 4.04 |
| edu | 114 | 41.45 | 26.15 | 3.65 | 95.35 |
| ind_perc | 170 | 31.01 | 11.29 | 3.33 | 72.02 |
| urban | 177 | 60.39 | 20.86 | 8.67 | 100 |

Source: World Bank

In order to conduct a multiple linear regression with the independent and dependent variables given, the data should meet the Gauss-Markov Assumptions. The first assumption, that the multiple regression model is linear in parameters, is met since the ensuing multiple regression is of the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + u$$

The second assumption, that the sampling is random, is only partially met since the data used is only for the countries for which data was available from the World Bank for 2013, so some bias may be caused if a country's likeliness of having their data provided in the database is correlated to any of the variables. However, if this is assumed to not be the case, then the random sampling condition is met.

Table 2: Correlation between Independent Variables

| | lngdp | lnresources | edu | ind_perc | urban |
|---|---|---|---|---|---|
| lngdp | 1 | | | | |
| lnresources | -0.15 | 1 | | | |
| edu | 0.16 | -0.13 | 1 | | |
| ind_perc | 0.049 | 0.23 | -0.45 | 1 | |
| urban | 0.16 | 0.75 | -0.47 | 0.26 | 1 |

Furthermore, none of the independent variables are constant as evidenced by the non-zero standard deviations in Table 1, and there is no exact linear relationship between the independent variables lngdp and lnresources. Therefore, the third Gauss-Markov assumption is met. The fourth assumption holds that the value of the explanatory variables do not contain any information about the mean of the unobserved factors. Since GDP per capita is likely to be correlated to many potential sources of unobserved variables, this assumption is not definitively met given the current selection of independent variables. In order to eliminate this deficiency in the data, more independent variables may be considered later. For example, urbanization rates and college enrollment rates may be significant but are yet to be tested. Lastly, the fifth assumption of homoscedasticity is likely met because the values of the two independent variables are unlikely to be correlated to the variance of unobserved factors. However, this is difficult to justify.

**Results**

The primary relationship studied was the effect that national GDP per Capita (gdp) had on percent of Non-Fossil Fuel Energy Consumption in 2013 (energy). This was a skewed relationship however and energy is already a percentage variable, so studying the relationship of lngdp on energy made more sense. This resulted in a regression of:

$$\widehat{energy} = 120.0454 - 10.01248 * lngdp$$

This means that for every 1% increase in a country's GDP, the model predicts that their alternative energy usage decreases by 10 percentage points.

Statistics:

Table 3. Simple Regression Results

| **Independent Variables** | **Model (1): Simple Regression** |
|---|---|
| lngdp<br>(std error) | -10.012***<br>(1.228) |
| Intercept | 120.045***<br>(11.015) |
| No. of obs. | 178 |
| R-square | 0.274 |

This was still a somewhat weak relationship, especially for wealthier countries, so other variables were accounted for to get a better picture. The other variables thought to have a significant effect on cleanenergy were education level, edu, and the total rent of coal, natural gas, and oil as a percentage of GDP in 2013 (resources). Considering the skew of this relationship, it made more sense to use lnresources instead. This upgraded regression is as follows:

$$energy = 127.450 - 11.326 * lngdp - 2.057 * lnresources + 0.000316 * edu$$

Statistics:

Table 4. Multiple Regression Results

| Independent Variables | Model (2):<br>3 Ind. Variables |
|---|---|
| lngdp<br>(std error) | -11.326***<br>(1.645) |
| lnresources | -2.057**<br>( 0.893) |
| edu | 0.000316<br>( 0.082) |
| Intercept | 127.450***<br>(14.464) |
| No. of obs. | 88 |
| R-square | 0.374 |

The simple linear regression above predicts that for each 1% increase in a country's GDP per capita, the percentage of their energy consumption to come from clean sources would decrease by about 10 on average. This shows that countries that are better off are more reliant on fossil fuels. As of 2013, it is likely that most industry runs on fossil fuel energy and has not switched to cleaner energy sources. Hopefully in the future with new investment this may change. Since the model predicts a percentage, which can only be between 0% and 100%, it can be physically meaningful when a nation's GDP is between $7 and $161,067 per capita in current, October 2017, USD. All nations studied fell well within this range.

The multiple linear regression predicts that for each 1% increase in a country's GDP per capita, the percentage of their energy consumption to come from clean sources would decrease by about 11 on average. This is just slightly steeper than before. It also predicts that for each 1% increase in fossil fuel rents per GDP results in an average decrease in percentage of clean energy consumption by 2. It is not surprising that there would be a negative relationship between the two as countries with more fossil fuel production would likely use more. The fact that it was significantly less steep than the relationship between GDP and clean energy consumption was initially surprising, but is realistic considering that their fossil fuel consumption is divided by GDP in the variable and suggests that much of the fossil fuels produced are exported. Considering that the fossil fuel production of the world is highly concentrated this is not surprising.

The statistical significance of this model should be taken into account, however, in order to provide a clearer context for the results. For example, Table 5 below justifies the classification of edu as statistically insignificant since it does not fall within a ten percent level of significance. Moreover, notice the P-value of zero for lngdp that indicates it falls within a one percent confidence level, and the P-value of 0.024 for lnresources indicates that it falls within a five percent confidence level. Overall, one could say with this evidence that the edu variable could be removed from the regression if no other variable were considered.

Table 5. Statistical Inference Results for Multiple Linear Regression Model (2)

| Independent Variables | T-statistic | P-value | 95% Confidence Intervals |
|---|---|---|---|
| lngdp | -6.885 | 0.000 | (-14.957, -8.054) |
| lnresources | -2.303 | 0.024 | (-3.832, -0.281) |
| edu | 0.003856 | 0.997 | (-0.163, 0.163) |
| Intercept | 8.81 | 0.000 | (98.686, 156.214) |

**Extensions**

In order to solve some of the issues with the preceding model, a new functional form is considered now to reflect the impact of two more explanatory variables: ind_perc and urban. The addition of more variables here should constitute a more thorough model. The updated regression is:

$$energy = 132.100 - 9.478 * lngdp - 0.004 * lnresources + 0.053 * edu - 0.685 * indperc - 0.016 * urban$$

Statistics:

Table 6. Final Regressions with additional variables

| Independent Variables | Model (3): Unrestricted | Model (4): Restricted |
|---|---|---|
| lngdp<br>(std error) | -9.478***<br>(2.449) | -8.977***<br>(1.227) |
| lnresources | -0.004<br>(1.145) | |
| edu | 0.053<br>(0.081) | |
| ind_perc | -0.685***<br>(0.232) | -0.645***<br>(0.144) |
| urban | -0.016<br>(0.163) | |
| Intercept | 132.100***<br>(15.649) | 130.931***<br>(11.390) |
| No. of obs. | 83 | 170 |
| R-square | 0.450 | 0.332 |

The unrestricted model in Table 6 provides the coefficients for all of the explanatory variables, but it is evident that the introduction of the two new variables has caused a discrepancy in terms of which variables are statistically significant. This is reflected in the restricted model in Table 5 which only contains coefficients for lngdp and ind_perc. In other words, the unrestricted model shows that lnresources, edu, and urban are not statistically significant, so they are omitted from the final restricted model which is represented by the following equation:

$$energy = 130.931 - 8.977 * lngdp - 0.645 * indperc$$

The following statistical inference tables justify the classification of lnresources, edu, and urban as statistically insignificant variables in the unrestricted model as well as the classification of lngdp and ind_perc (and the intercept constant) as statistically significant in both the unrestricted and restricted models. Notice the P-values of less than 0.01 for lngdp and ind_perc, indicating better than one percent significance levels for these variables.

Table 7. Statistical Inference Results for Unrestricted Model (3)

| Independent Variables | T-statistic | P-value | 95% Confidence Intervals |
|---|---|---|---|
| lngdp | -3.87 | 0.000 | ( -14.354, -4.602) |
| lnresources | 0.66 | 0.513 | (-2.284, 2.275) |
| edu | 0.003856 | 0.997 | (-0.108, 0.214) |
| ind_perc | -2.95 | 0.004 | (-1.147, -0.223) |
| urban | -0.10 | 0.920 | (-0.342, 0.309) |
| Intercept | 8.44 | 0.000 | (100.939, 163.260) |

Table 8. Statistical Inference Results for Restricted Model (4)

| Independent Variables | T-statistic | P-value | 95% Confidence Intervals |
|---|---|---|---|
| lngdp | -7.32 | 0.000 | ( -11.399, -6.556) |
| ind_perc | -4.47 | 0.000 | (-0.930, -0.360) |
| Intercept | 11.50 | 0.000 | (108.444, 153.418) |

In order to test overall significance, an F-test was conducted. The unrestricted model used lngdp, lnresources, urban, and ind_perc as the independent variables and energy as the dependent variable. The restricted model was the same except for the fact that lngdp was omitted. The resulting F-statistic was 15.466, which is very high, indicating that the omitted variable is significant. Other restricted models and thus multicollinearity could not be tested because of varying sample sizes.

**Conclusion**

The initial hypotheses were that as a country's GDP increases, its level of clean energy consumption would increase and that as a country's level of fossil fuel reserves increases, the level of clean energy consumption decreases. It was also hypothesized that a country's level of college education has a positive correlation with clean energy consumption. The first multiple linear regression model supported the second of these two hypotheses. It predicted that each one percent increase in fossil fuel rents per GDP resulted in an average decrease in percentage of energy consumption by two. The education variable had a very high p-value and was therefore not significant. The surprising result was that the first hypothesis was not supported. The results indicated that wealthier countries in terms of GDP are actually less likely to rely on clean energy consumption.

In order to improve the accuracy of the model, two new variables were added: percentage of a country's population living in an urban area and percentage of value added to a country's GDP by the energy industry of the country. In this regression, level of fossil fuel reserves, education, and percent of population living in an urban area were found to be not significant. GDP once again showed a negative correlation with clean energy consumption as did the ind_perc variable.

The result which once again does not support the original hypothesis is that as a country's GDP grows, its level of clean energy consumption actually decreases. Therefore, the implication for the United Nations sustainable development goal is that other avenues must likely be pursued such as government investment or subsidies to decrease fossil fuel consumption.

## Sources

Asafu-Adjaye, John, et al. “Dataset for analysing the relationships among economic growth, fossil fuel and non-Fossil fuel consumption.” *Data in Brief*, vol. 10, 2017, pp. 17–19., doi:10.1016/j.dib.2016.11.076.

Destek, Mehmet Akif, and Alper Aslan. “Renewable and non-Renewable energy consumption and economic growth in emerging economies: Evidence from bootstrap panel causality.” *Renewable Energy*, vol. 111, 2017, pp. 757–763., doi:10.1016/j.renene.2017.05.008.

Pao, H.T., Li, Y.Y. and Fu, H.C. (2014) Clean Energy, Non-Clean Energy, and Economic Growth in the MIST Countries. Energy Policy, 67, 932-942. http://dx.doi.org/10.1016/j.enpol.2013.12.039

Unrestricted Regression for F-test

```
. regress energy urban lnresources lngdp ind_perc

      Source |       SS           df       MS      Number of obs   =       136
-------------+----------------------------------   F(4, 131)       =     19.02
       Model |  29388.3246         4  7347.08116   Prob > F        =    0.0000
    Residual |  50613.9721       131  386.366199   R-squared       =    0.3673
-------------+----------------------------------   Adj R-squared   =    0.3480
       Total |  80002.2968       135  592.609606   Root MSE        =    19.656

------------------------------------------------------------------------------
      energy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       urban |  -.0393988   .1447332    -0.27   0.786    -.3257156     .246918
 lnresources |  -.7645419   .9120295    -0.84   0.403    -2.568754     1.03967
       lngdp |  -8.971853   2.208421    -4.06   0.000    -13.34064   -4.603069
    ind_perc |  -.5173247   .1827189    -2.83   0.005    -.8787862   -.1558632
       _cons |   127.9037   14.03172     9.12   0.000     100.1456    155.6618
------------------------------------------------------------------------------
```

Restricted Regression for F-test

```
. regress energy lnresources ind_perc urban

      Source |       SS           df       MS      Number of obs   =       136
-------------+----------------------------------   F(3, 132)       =     17.77
       Model |  23011.5723         3  7670.52409   Prob > F        =    0.0000
    Residual |  56990.7245       132  431.747913   R-squared       =    0.2876
-------------+----------------------------------   Adj R-squared   =    0.2714
       Total |  80002.2968       135  592.609606   Root MSE        =    20.779

------------------------------------------------------------------------------
      energy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 lnresources |  -.0957278   .9482684    -0.10   0.920    -1.971496    1.780041
    ind_perc |  -.6158353   .1914435    -3.22   0.002    -.9945294   -.2371412
       urban |  -.4948361   .0967692    -5.11   0.000    -.6862551   -.3034171
       _cons |   78.49735   7.398946    10.61   0.000      63.8615    93.13319
------------------------------------------------------------------------------
```

Countries included in final model:

| | | | | | |
|---|---|---|---|---|---|
| Albania | Congo, Rep. | Iceland | Malta | Russian Federation | Ukraine |
| Algeria | Costa Rica | India | Mauritius | Saudi Arabia | United States |
| Armenia | Cote d'Ivoire | Indonesia | Moldova | Senegal | Uruguay |
| Austria | Croatia | Ireland | Mongolia | Singapore | Uzbekistan |

| | | | | | |
|---|---|---|---|---|---|
| Azerbaijan | Cuba | Italy | Morocco | Spain | Venezuela, RB |
| Bahrain | Cyprus | Jamaica | Mozambique | Sri Lanka | Zimbabwe |
| Bangladesh | Ecuador | Japan | Namibia | Sudan | |
| Belgium | Egypt, Arab Rep. | Jordan | New Zealand | Suriname | |
| Benin | Ethiopia | Kenya | Oman | Sweden | |
| Bolivia | Finland | Kyrgyz Republic | Panama | Tajikistan | |
| Botswana | France | Latvia | Peru | Tanzania | |
| Cameroon | Germany | Lebanon | Philippines | Thailand | |
| Chile | Ghana | Lithuania | Poland | Togo | |
| Colombia | Honduras | Luxembourg | Portugal | Trinidad and Tobago | |
| Congo, Dem. Rep. | Hungary | Malaysia | Qatar | Tunisia | |