A Comparison of Observed Worldwide Data with Commonly Understood Economic Principles

JiHo Bang, Sergio Cervantes, Jacob Kim, Mario Roman, O'Zie Sagbohan, Shan Virani

12/11/2020

I. Introduction

Through our deliberation as Economics apprentices, we have grown to form cognitive understandings of how various macroeconomic variables influence GDP and promote economic growth. Through models that attempt to delineate a relationship between GDP and these variables, one can understand how, theoretically, a nation can utilize policy to promote endogenous growth in an economy. However, one must ponder how strong the relationships between these variables and GDP really are. Testing whether or not these variables act in accordance with preconceived economic notions can inspire further research into developing stronger theories surrounding macroeconomic models. Through a regression analysis of GDP and labor, land, legal protections of property rights, annual population growth, business ease, and the active telephone lines of 120 nations, we seek to determine how much previously understood macroeconomics fundamentals "hold up".

The **Exploratory Data Analysis** portion will give insight into the variables we will be testing, including some neat facets and summarizations that may prove helpful when conceiving the data.

The **Data Analysis** portion will consist of scatterplots and hypothesis tests that relate a macroeconomic variable to GDP, GDP per-capita, or the log of GDP, and will be accompanied by a hypothesis test. The null hypotheses in our hypothesis tests are that the slope of the regression line for each correlation plot is equal to 0, implying no correlation. The alternative hypotheses contend that the slope of the regression line for each correlation plot is not equal to 0, implying some sort of relationship. This will result in **six regression lines** and **six hypothesis tests**, each connecting a macroeconomic variable to some form of GDP that is appropriate to its place in economics. Finally, we will be determining whether the relationships found in our regression lines and hypothesis tests support our previously understood macroeconomic doctrines.

II. Exploratory Data Analysis

World Map Coloured by Legal Rights
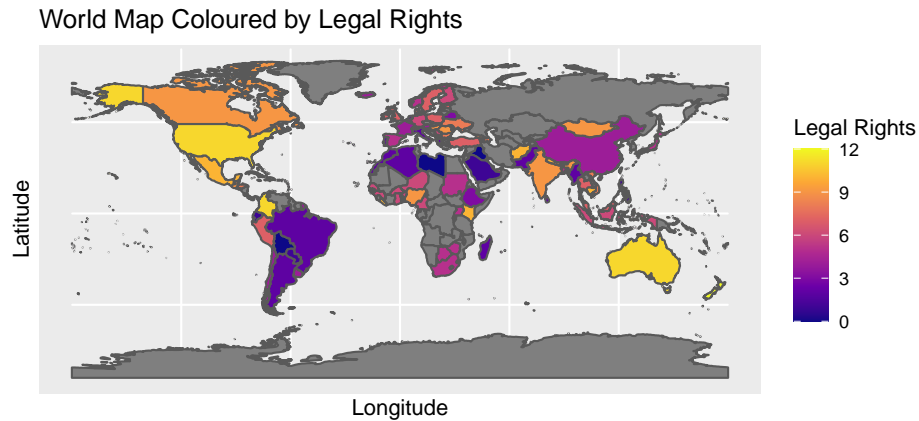
Though it may seem rudimentary to quantify how strong a country's legal rights are, the World Bank attempts to do so by generating an index that measures the degree to which collateral and bankruptcy laws facilitate lending. Showing the magnitudes of legal rights in geographic terms allows for interpretation of which regions generally allow for greater facilitation and issuance of credit. The greyed nations are not accessible through the World Bank's dataset.
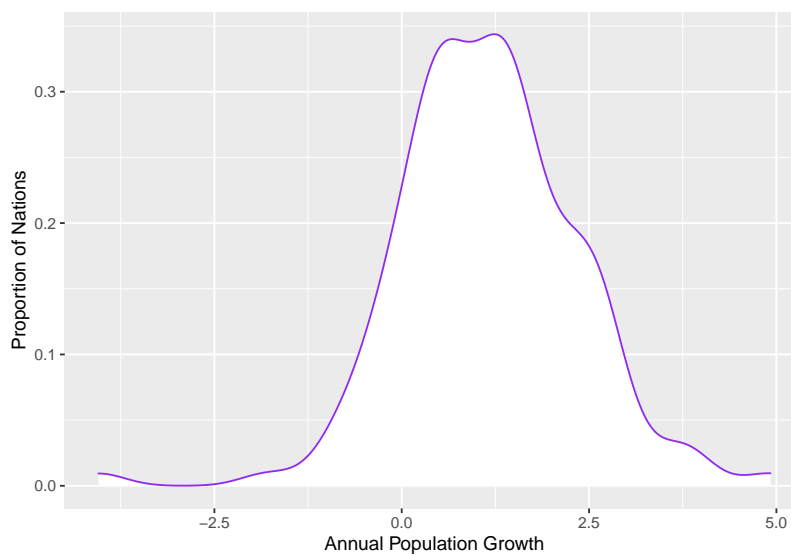
The population growth density graph does, mostly, resemble a normal distribution, but it is slightly skewed left. In the distribution, there is some bulging aiming to the right and a left tail that is a bit longer than the right tail. This means that the majority of countries experience positive population growth with very few

countries experiencing extremely negative population growth, or population decay. An unexpected observation was the somewhat high proportion of nations with a negative annual population growth rate.

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```
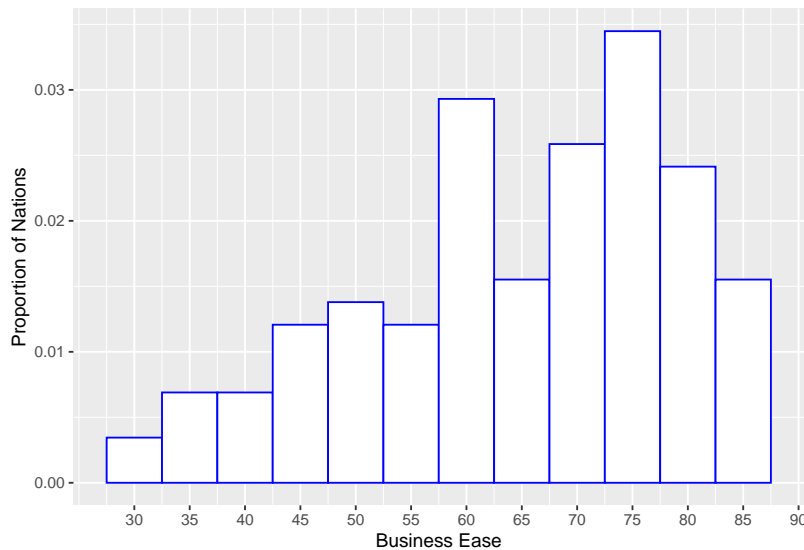

The mean of this data is 64.5, the median is around 67, with an SD of 14.2. This allows us to conclude that our histogram is left skewed, meaning most of the data is greater than the average.

A few small values bring the mean down: countries like Yemen, Venezuela, and Libya have the lowest scores for ease of doing business and are extremely difficult compared to other countries. These countries are notorious for their political instability and are under constant threat from militias and other hostile groups (International Crisis Group, 2020). More, in Venezuela there are 17 procedures for starting a business compared to the Latin and Caribbean average of 9, and Venezuela ranks 181st out of 185 economies for investor protection. When people don't feel that starting a business will be safe or efficient, they will be less likely to create one.
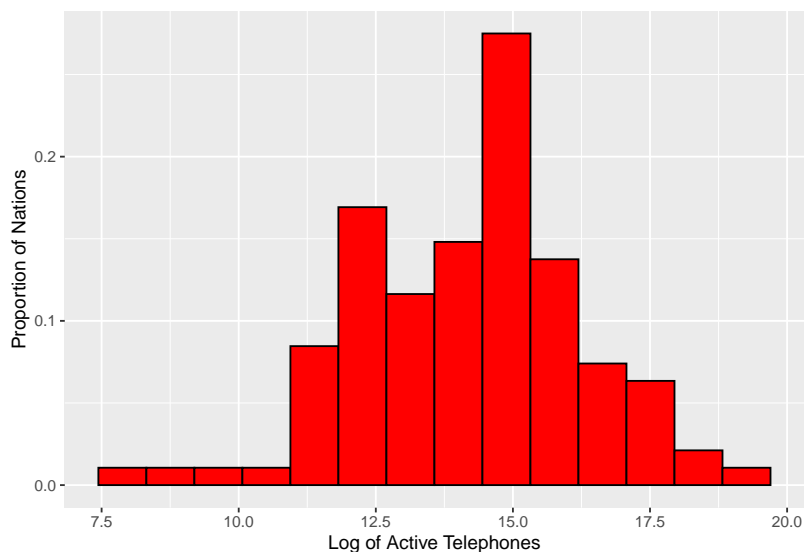
When analyzing the histograms, we can notice that it is left-skewed and come to the realization that the majority of nations have higher access to telephones. With the mean number of telephones in a nation being 8192813, median being 1829166, it is interesting to notice the large Standard Deviation Value of 22642710. This large SD is largely influenced by some outlying countries that have "0" telephones.
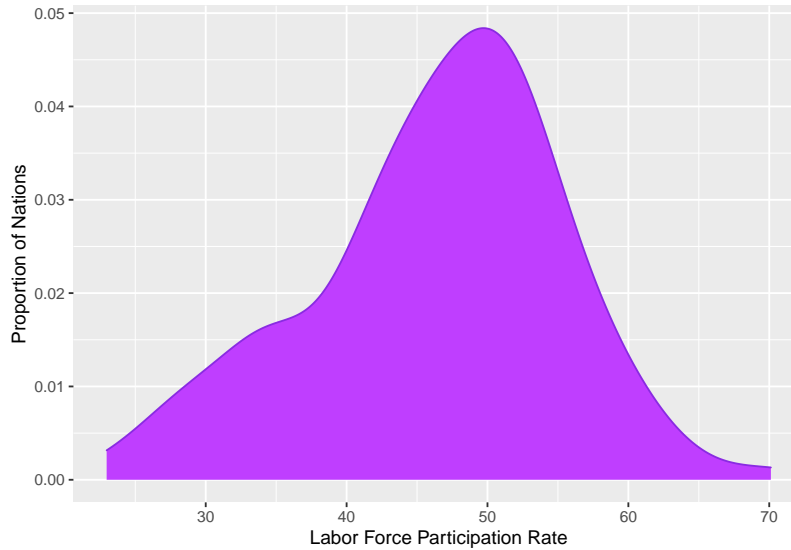
Labor represents the number of people in the country that are either working or looking for work. In the following graph, we examine what percentage of each country's total population is included in the work force. We see a trend with more countries near the 50% threshold for a working population. We will later analyze how the various allocations of working population affect GDP.
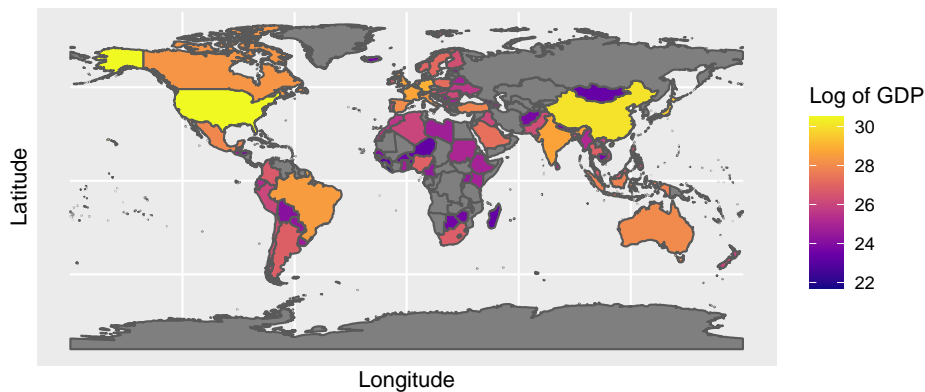

While graphs allow us to understand data, we decided to use a map to allow a better visual representation

of the relationship between land and GDP. By observing the graph and scale, the correlation is clear, larger countries are generally associated with higher GDPs. The scale, log of GDP, matches up with the world bank data, leading to some gray nations on the map. However, this is still a better representation of this specific study, as it is based off of the 120 countries that are actually shown. Ultimately, we found that this was the most interesting way to represent this data because it allows the reader to see the different sizes and GDPs of every country. Moreover, it also provides insight into patterns such as which regions have economic troubles, hurting the correlation from becoming stronger and closer to 1.

III. Data Analysis

One of the more commonly understood concepts in Economics academia is the role the government plays in driving productivity and growth in an economy by facilitating lending and instituting prudential regulation (Hawkins, 2020) (Niwa, 2018). Essentially, we have grown to associate stronger institutions facilitating further credit lending with greater levels of economic growth. Therefore, when such a variable like the strength of legal rights is quantified, one would expect it to form a relationship and correlate with GDP growth.

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 8 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 8 rows containing missing values (geom_point).
```
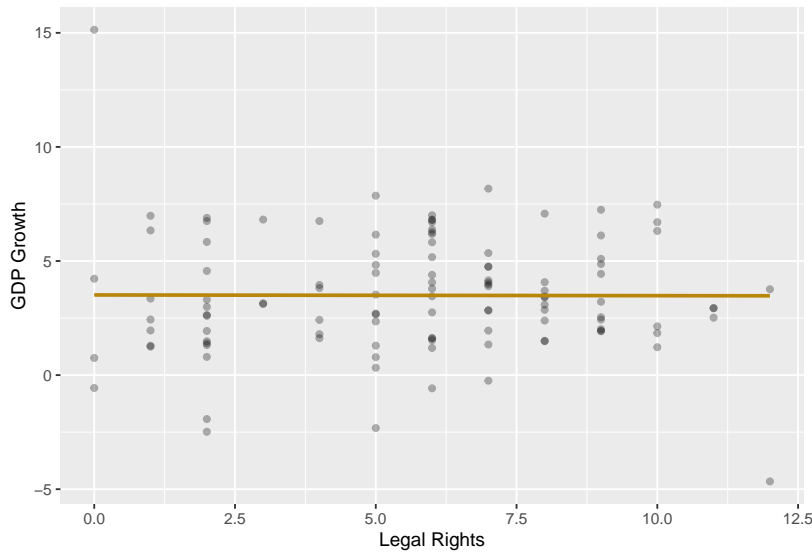

Through our regression analysis, we find no correlation between the two variables; in fact, the correlation coefficient is negative. After further investigation and the conducting of a hypothesis test for the slope of the regression line, we find a p-value of 97% and very firmly support the null hypothesis in the notion that the slope of the regression line is 0. This is an important observation. A metric like legal rights essentially defines the ease to facilitate credit lending; if this shows no correlation with GDP growth, then that would be a direct contradiction to what economists put forward.

Continuing, the population growth model generally predicts that a nation with a higher growth rate of the population may also be experiencing greater forms of economic growth (Bucci, Eraydin, Muller, 2018). By finding the correlation between the population growth and GDP growth, one may be able to delineate whether this economics convention holds up in the observed data so far.

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```
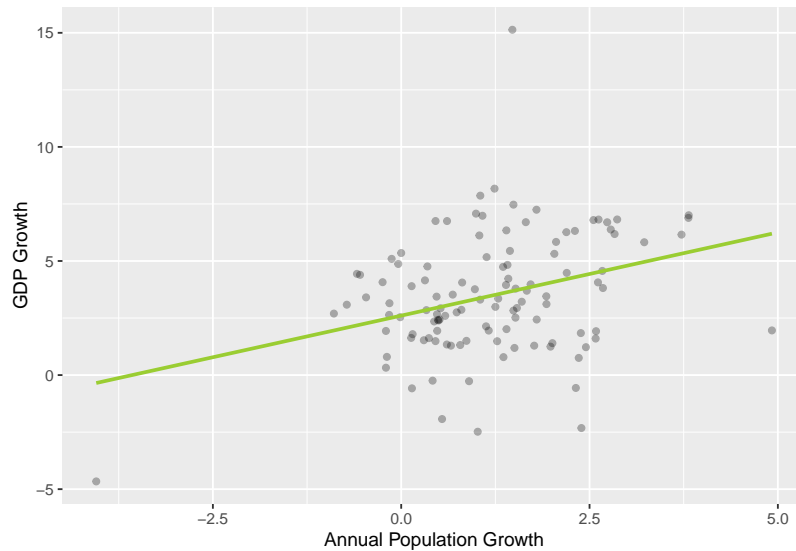

As shown by the model, annual population growth does have a relationship to GDP growth. We are able to reject the null hypothesis at the 99.9% level, rendering it extremely difficult to argue that the slope of the regression line is equal to 0. This supports the notion perpetuated by Bucci, Eraydin, and Muller in their assertion that the population growth of a nation is related to its GDP growth.

The ease of doing business metric refers to the regulatory environment that would determine the advantages and efficacy associated with starting and maintaining a firm. Through the growth accounting formula, economists would generally predict greater GDP growth with a greater growth rate of Total Factor Productivity (TFP). In this paper's case, business ease would be reflected within TFP. Beyond capital and labor, TFP would contribute towards growth in output. A classical economist would generally argue that the ease associated with building a business would have a positive correlation with GDP growth.

```
## `geom_smooth()` using formula 'y ~ x'
```
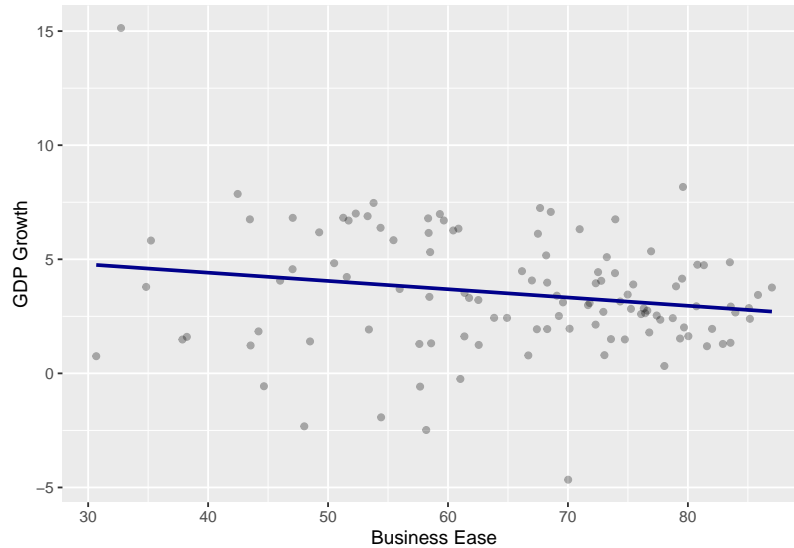
In correlating the log of business ease and log of GDP, we get a negative correlation of -.18. With a p-value of .04, we reject the null hypothesis at the 95% level. This is a direct contradiction to what we've come to know about business ease in that it should contribute to higher GDP growth.

Furthermore, the Romer model attempts to demonstrate how technology– non-rival and non-excludable goods that are used to produce goods or services– contributes to per-capita income– found by dividing the GDP of a nation by it's population– by asserting that the growth of technology contributes to GDP per capita (Hawkins, 2020). An increased use of telephones would indicate a greater growth rate of efficiency. Therefore, Economists may argue that the accessibility to a non-rival, non-excludable, productive good like telephones would contain a correlation with per-capita income in an ideal world.

## `geom_smooth()` using formula 'y ~ x'
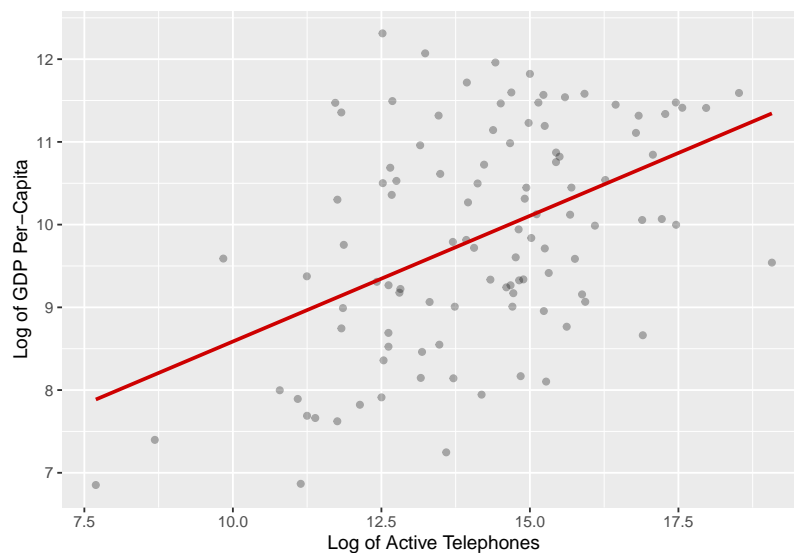
Through the observed data, one finds that the accessibility to a non-rival, non-excludable, productive good like telephones has a correlation with per-capita income. We reject the null hypothesis; the slope of the regression line is absolutely not equal to 0, so there is a relation when predicting the log of GDP per-capita

from the log of active telephones in a nation. Using the number of active telephones as a predictor of GDP per-capita would show efficiency.

The Cobb-Douglas production function is paramount to an understanding of how output can be defined in terms of capital, labor, and total-factor productivity. Labor being positively correlated with GDP is a broadly accepted macroeconomic principle, but does this correlation apply in real-life data?

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```
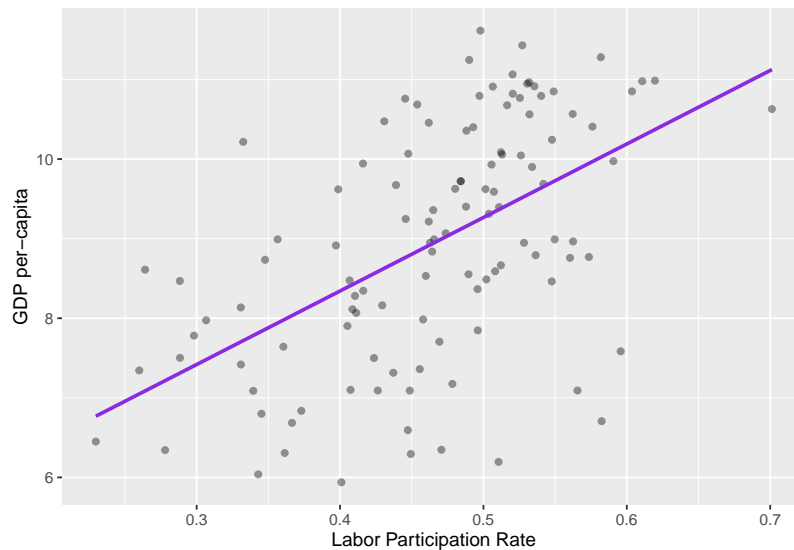

Here we see that Labor Percent and GDP by Population have a significant positive correlation. Countries that have a higher percentage of people in their workforce tend to have increased GDP as well. The football-shaped plot implies a linear association. We strongly reject the null hypothesis; the slope of the regression line is not equal to 0, supporting a positive correlation. This coincides with our previous understanding of how labor and GDP should be related.

Finally, it seems like an intuitive idea that greater land would correlate with higher GDP growth. Land use can enhance the welfare of citizens and promote local economic growth in a multitude of ways — it provides the space for human activity, labor, agriculture, while also allowing markets to expand and flourish (Kim, 2011). In fact, economic superpowers around the globe have, in some way, taken advantage of their land to increase their GDPs. Given that extremely large countries are often recognized as successful nations (United States, China, Russia, Canada), we assume that greater land associates with more growth. Thus, when comparing land to GDP growth, one would expect a strong correlation between the two variables.

```
## `geom_smooth()` using formula 'y ~ x'
```
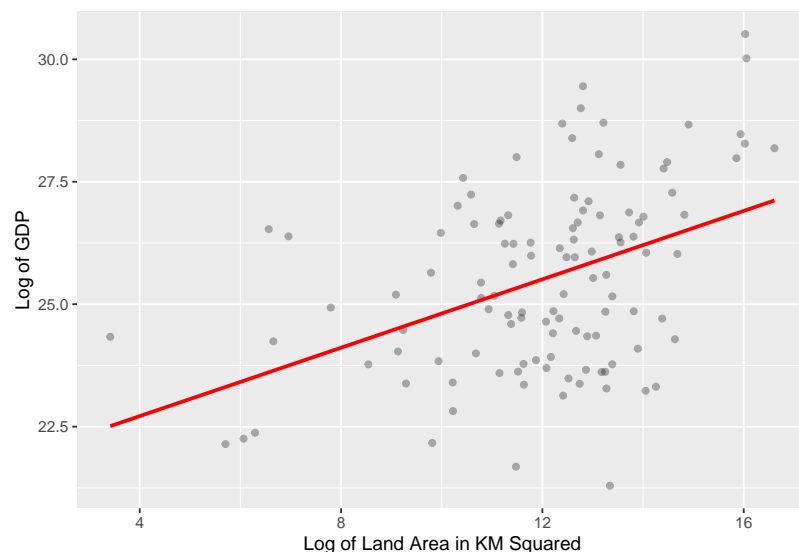
    The graph shows that there exists a correlation between land and GDP. We find that the positive correlation is r = 0.537, and when attending to the logs of both variables, r = .425. We graphed the natural logs of both variables in order to account for outliers, showing a more complete image of the data. Nevertheless, we reject the null hypothesis as the slope of the regression line is clearly a positive correlation, not 0. This corroborates our previous idea of how land and GDP are related: generally, countries with greater areas also have high GDPs.

IV. Conclusion

    One of the central limitations encountered during our apotheosis of efforts included the scattered "N/A" entries in the data frame. By filtering these out in each summary and scatterplot, we are able to run statistics on the parameters; however, this would leave an omitting of entire nations' entries. Filling the holes in the data would yield more enveloping and inclusive data, providing conclusions of marginally greater significance. As a result of our reflection, we conclude that some variables are better predictors of their respective output-statistic than others.

    There may be confounding variables that could account for the discrepancies in our assumed economics knowledge and the relationships shown in the real-life data. For example, domestic instability in foreign exchange markets and speculative attacks on a nation's currency may lead to an inadequate measure for GDP; also, consumer culture varying from nation to nation may affect spending habits. However, one would assume that economics principles at least hold in their relationship to GDP. Though factors like labor, access to telephones (in other words, technology), the population growth rate, and land follow their expected relationship with their respective output-variable, business ease and the metric for legal rights draw contrast with their presumed relationships to GDP growth.

V. Cited References

Bucci, A., Eraydin, L., Müller, M. (2018). Dilution effects, population growth and economic growth under

    human capital accumulation and endogenous technological change. *Working Paper Series in*

    Economics, 113, pp. 1-5.

Hawkins, R. J. (2020). *Growth and the Solow-Swan Model III.* Personal Collection of Raymond J. Hawkins,

    UC Berkeley, Berkeley, CA.

Hawkins, R. J. (2020). *The financial system and economic growth.* Personal Collection of Raymond J.

Hawkins, UC Berkeley, Berkeley, CA.

International Crisis Group. (2020). A glut of arms: Curbing the threat to Venezuela from violent groups. *Latin America Report, 78.*

Jae Hong Kim, (2011). Linking land use planning and regulation to economic development: A literature review. *Journal of Planning Literature, 26*(1), 35–47. doi:10.1177/0885412210382985

Niwa, S. (2018). Effects of a blocking patent on R&D with endogenous survival activities. *Journal of Economics, 125*(3), pp. 263-277. DOI: 10.1007/s00712-018-0617-1