Effect of a Country's GDP on the Availability of its Agricultural Land

Christian Kim

University of California, San Diego

ECON 5 / POLI 5D: Data Analytics

2020 Spring Quarter

Abstract

The following report observes the effect of a country's GDP on the amount of arable agricultural land in 2006 and 2016. GDP is expressed in USD, while the amount of arable land is shown in percentage of the total amount of land within that country. In exploring this relation, this report takes data from over 200 countries and examines whether a rise in GDP does correlate to a decrease in the percentage of agricultural land. This report also examines a confounding variable, a country's total human population, and examines its effects on a country's total human population. The findings disconfirm the hypothesis that a country with a higher GDP correlates to a corresponding lower agricultural land percentage

I. Introduction

In Macroeconomics, Gross Domestic Product (GDP) is one of the most important measurements of assessing a country's economic health since it takes into account a nation's production, the types of incomes, and how the incomes are utilized (Bureau of Economic Analysis). As a statistical form of analysis, GDP is utilized to show how much consumers are purchasing products and how much people, overall, are getting paid through wages. For instance, a higher GDP figure indicates that incomes are increasing, consumers are purchasing more goods, and, above all, the economic health of the country is stellar. Oftentimes, elected officials and the Federal Reserve would utilize the figure as a percentage of rate of change, to plan or examine spending, monetary, and tax policies, while business leaders and corporate executives would use the statistics as a way to gauge future expansions, investments, the possibility of inflation, and so much more (Bureau of Economic Analysis). Thus, we can see how a country's GDP becomes a sort of trickle-down concept, where the prosperity of a country based on this figure can ultimately be advantageous for working-class people who benefit from either higher wages or more job opportunities.

One of the most prominent contributors to a country's overall sustainability and GDP stability is the availability of arable agricultural land utilized for crops, pastures, and any other necessities that contribute to the production of goods and services to the economy. As 60% of the world's population depends on agricultural production, an increase from 7 billion to 9 billion people would require at least 12% of available land to be used for agricultural production, such as crops and farming labor (Zapata, Perrone, & Figus). While the production of essential goods, such as food and water, and necessities, such as labor, income, and shelter, all derive from the contribution of agriculture to the economy, an increase in life expectancy is one of the most important benefits of all. Much research within economics, especially those that focus on comparative studies between first-world and third-world countries, see a direct correlation between higher levels of life expectancy would be associated with higher economic development; additionally, such an increase continues as a repetitive cycle where a higher economic development leads to higher life expectancy, which then again leads to more worker productivity and ultimately even more economic development (De Sormeaux & Pemberton). Therefore, there is a cycle where the amount of arable agricultural land contributes to the production of everyday essentials, which then increases life expectancy, following an increase in workers, and ultimately an increase in GDP due to the surge in production and consumption.

Unfortunately, while there is a connection between the availability of arable agricultural land and GDP, there is much uncertainty regarding whether or not the amount of agricultural land is affected by a country's GDP. In other words, scholarly reports done in the past have not been able to consistently explore the idea of whether or not a high GDP indicates either more, equal, or less land availability. For instance, D. Gale Johnson, a Professor of Economics at UChicago, has affirmed that a more developing economic country, such as higher incomes for families, means a decrease on food production; for instance, a family's 100% income increase may only indicate a 60% increase in food, thus leading less dependence upon land area and human labor since animal power and machinery begin to take place around less land (Johnson, 2019). On the other hand, University of the West Indies's Afiya De Sormeaux and Carlisle Pemberton's report on agricultural economics concludes that a higher GDP can lead to much greater expansion of rural population since pressures of land expansion of agricultural producers, such as crop fields and farmland, are extended to other areas to account for a growing GDP rate and growing worker productivity (De Sormeaux & Pemberton). We can see how there is no overall established ground on whether or not a country's GDP increases, remains constant, or decreases the amount of arable agricultural land.

For these reasons, the question of this research is as followed: What is the relationship between a country's GDP level and its percentage of available arable agricultural land in the years 2006 and 2016? For this project, I will be focusing on a country's GDP level as an absolute figure to examine that data in comparison to the percentage of land that is dedicated to arable agricultural production. I hypothesize that a country with a higher GDP will have a lower percentage of available land since a higher GDP mixed with a society that is constantly innovating technologically would mean that there is less reliance on more arable agricultural land. I suspect that the creation of more labor through innovation and an increasing job market through an increase in a country's GDP could potentially cause urban expansion and possibly lead to the decimation of arable agricultural areas due to less reliance on more land for more food and to accommodate for population expansion.

II. Data

Four different data sets have been used for the investigative process, which include the GDP levels of each country in 2006, the GDP levels of each country in 2016, the amount of arable agricultural land by percentage in 2006, and the amount of arable agricultural land by percentage in 2016, all of which have been accumulated by The World Bank. The World Bank is a well-known initiative that has the goal of helping to end extreme poverty in developing nations while also promoting shared prosperity and sustainable development (World Development Indicators). They are also a premier resource for attaining comparable statistics about global development across over 200 different economies in areas such as economics, wealth inequality, and population dynamics. Furthermore, this source is widely utilized by many economists and social scientists to assess First World, Second World, and Third World countries in areas such as agriculture, poverty, and business, making this website to be extremely valuable and suitable for this report. The only setback to the compiled data is that certain countries do not have a recorded value for some of the variables; for instance, one country may have its GDP listed for 2006 but not for 2016, which could potentially skew the statistical process in analyzing the data if certain crucial outliers are not recorded.

Table 1: Data Dictionary

Variable Name	Variable Description
GDP_2006	GDP at Purchaser's Price in US Dollars in 2006
GDP_2016	GDP at Purchaser's Price in US Dollars in 2016
Land_2006	Percentage of Arable Agricultural Land in 2006
Land_2016	Percentage of Arable Agricultural Land in 2016

Table 1 above illustrates the descriptions of each variable name within the data frames and graphs provided in the report. The GDP levels for 2006 and 2016 are measured in US Dollars based on the sum of gross value added by all producers in the economy and taxes without subsidies taken into account (World Development Indicators). It was decided to use a ten-year gap between 2006 and 2016 since GDP is generally measured on a long-term basis rather than a short-term since newly implemented policies, regulations, and emerging consumer products from different countries can come into effect during this gap. The percentages of available agricultural land for 2006 and 2016 are measured by the share of land area that is defined by the Food and Agricultural Organization as suitable for mowing, pasture, permanent or temporary crops, kitchen gardens, land under market, and land providing trees for timber and wood (World Development Indicators). Again, since the change in land percentage will be much greater when assessing a longer time gap, it has been decided, similar to the GDP levels, that using a 2006 to 2016 gap is best when assessing differences in the amount of arable agricultural land. For this analytical report, the independent variable is a country's level of GDP, while the dependent variable is the amount of agricultural land that a country has within the same year. This study will be examining the independent and dependent variables across 2006 and 2016 to see if there is a consistent linear regression pattern since a consistent overall statistical analysis between both years will create a stronger, more compelling, revelation as to what kind of relationship exists between the two variables.

Table 2: Summary Statistics of GDP and Land in 2006 and 2016

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
GDP_2006	252	1,578,949,701,459	5,869,810,912,712	22,902,861	4,563,241,250	24,214,123,293	344,956,827,410	51,448,204,042,311
GDP_2016	250	2,508,489,106,244	8,413,909,575,036	36,547,420	8,172,515,985	51,417,989,632	513,971,245,712	76,164,616,246,718
Land_2006	254	37.790	20.528	0.467	21.549	37.799	51.959	83.962
Land_2016	254	37.668	20.583	0.558	22.427	38.338	49.680	82.560

Table 2 above examines the basic statistical outputs of the number of cases, mean, standard deviation, minimum, 25th percentile, 75th percentile, and maximum outputs of GDP levels in 2006 and 2016 as well as the amount of land in 2006 and 2016 across all recorded nations. While there is a change in the number of cases under the "N" column across all variables, it should be noted that some countries were not able to be recorded by The World Bank; for instance, while Bermuda's GDP was recorded in 2006 at $5,414,299,000, there is no recorded GDP for 2016. We can see that of the 252 countries in 2006, the average mean GDP is $1,578,949,701,459, whereas in 2016 the average mean with 250 countries taken into account was $2,508,489,106,266; therefore, the overall GDP in all countries has for the most part risen during the ten years. On the other hand, we see that the average land percentage in 254 recorded countries was 37.790 in 2006, where there is a slight decrease in 2016 with the average land percentage of the same number of recorded countries being 37.668.

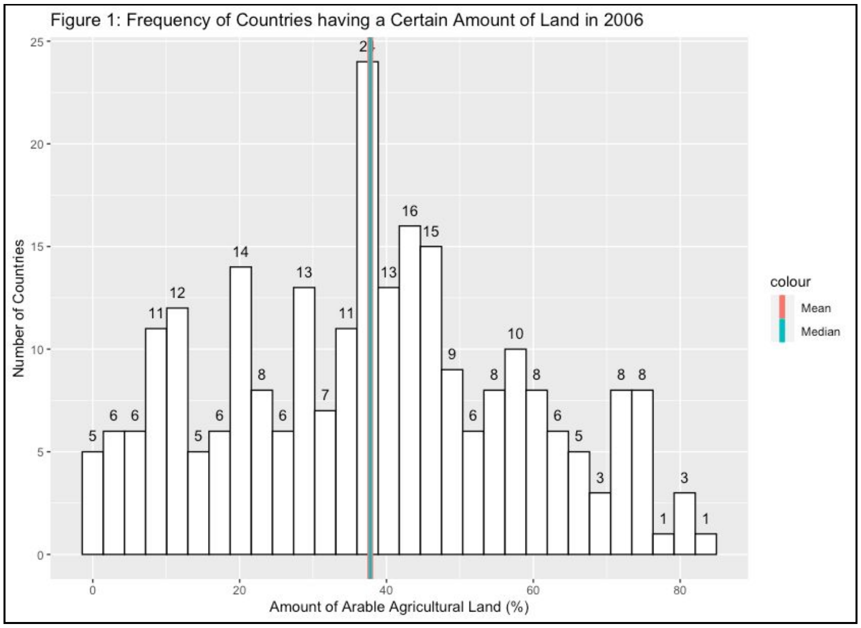
Additionally, it is vital to note that the skew that exists for GDP and Land during both years in Table 2. When looking at the mean and median for the GDP levels of both years, the mean for 2006 is larger than the median by $1,554,735,578,166, while the mean for 2016 is larger than the median by $2,529,071,116,612. Generally, the main difference between the median and mean is that if there is a huge outlier within the data, which in this case could be an outlying country having a colossal GDP level compared to the others, then the median would stay the same but then mean would be changed. Here, we see that the mean average GDP is exceptionally larger than the median for both 2006 and 2016, meaning that both data are positively skewed.

This report will also analyze a confounding variable, total population by country, to assess the possibility that a spurious relationship exists, where the connection between the independent variable, a country's GDP level in a certain year, and the dependent variable, the amount of arable agricultural land may be caused by a third property.

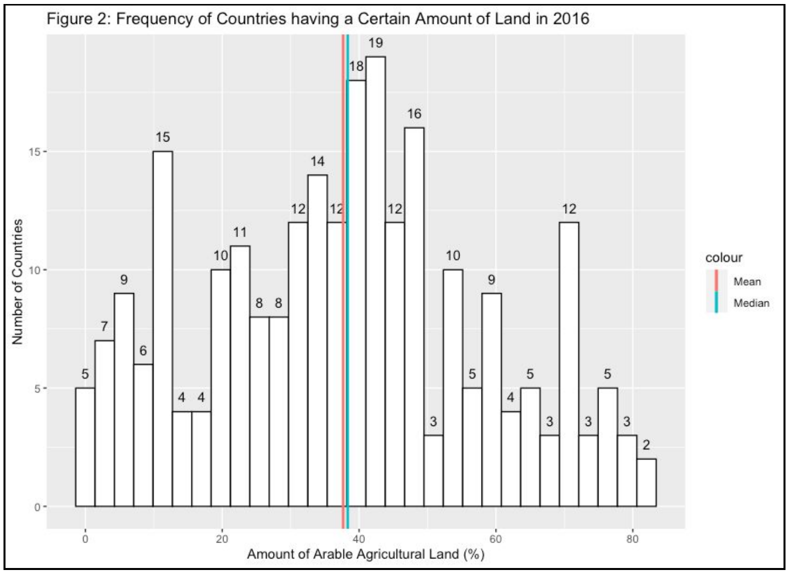
Table 3: Summary Statistics of a Country's Total Population

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
2006	263	263,740,858	840,467,421	9,828	1,324,686	8,484,550	48,895,254	6,593,623,202
2016	262	300,492,788	948,523,505	11,225	1,514,952	10,124,268	55,415,438	7,426,103,221

Table 3 above outlines the statistical analyses of all countries based on the confounding variable, the total population of a country in 2006 and 2016. Overall, there is an observable increase in world population between the 10-year gap; for instance, despite one less country having their population recorded in 2016 (262 countries) compared to 2006 (263 countries), the average mean population across all nations increased by 36,751,930, the country with the smallest population increased by 1,397 as shown by the minimum values, and the country with the largest population increased by 832,480,019 as shown by the maximum values. A possible explanation for this significance can be attributed to how the exploitation of agricultural land can sometimes be attributed to massive urbanization, usually in the form of rapid expansion of cities and metropolitan areas, where housing, commercial development, and roads are exponentially expanded with a growing population (Atlas of Urban Expansion).



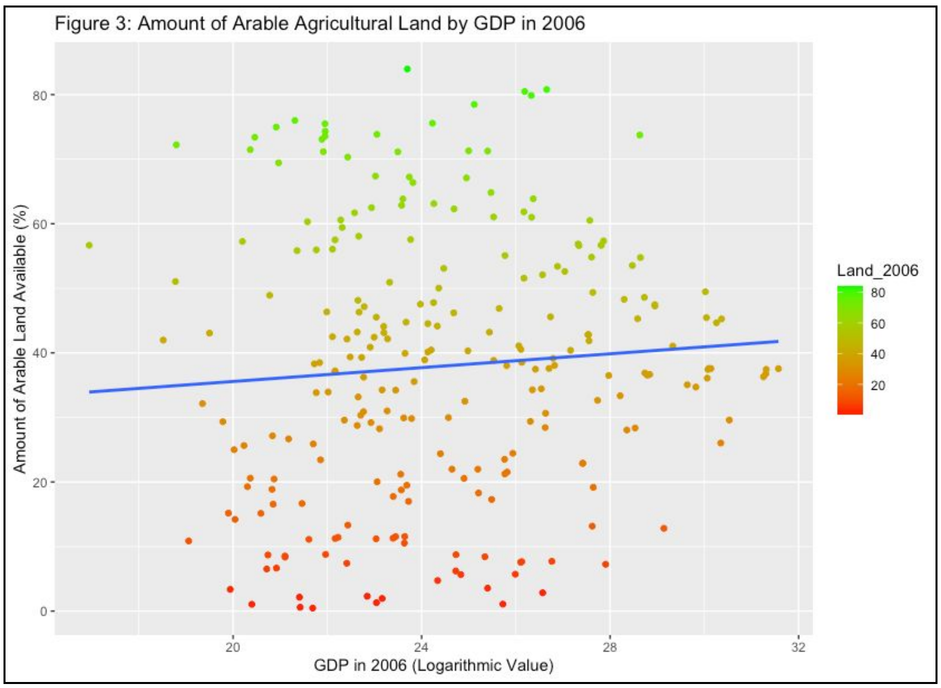
Figures 1 and 2 illustrate the frequency distribution of all countries with a certain percentage of availability of arable agricultural land for 2006 and 2016. While not entirely visible since the median and mean line overlap in Figure 1, we can see that the median (37.799) and mean (37.790) in 2006 lines appear to overlap one another since the difference between the median and mean is a mere 0.009 percentage point. On the other hand, the mean percentage of agricultural land in Figure 2 for 2016 (37.668) is to the right of the median (38.338); the difference here between the median and mean is 0.67 percentage point, meaning that there is more of a negative skew in the data for 2016 than 2006 when it comes to the percentage of available agricultural land across all countries. Furthermore, while the median amount of agricultural land in 2016 is larger than the median in 2006 by 0.539 percentage point, the mean is smaller in 2016 than 2006 by 0.122, which indicates that there is an overall decrease in data of agricultural land availability across all countries ten years later. A smaller average mean figure compared to the median figure only means that the middle figure is larger, which, unlike the mean, does not speak for the data when it comes to increases and decreases in all data.



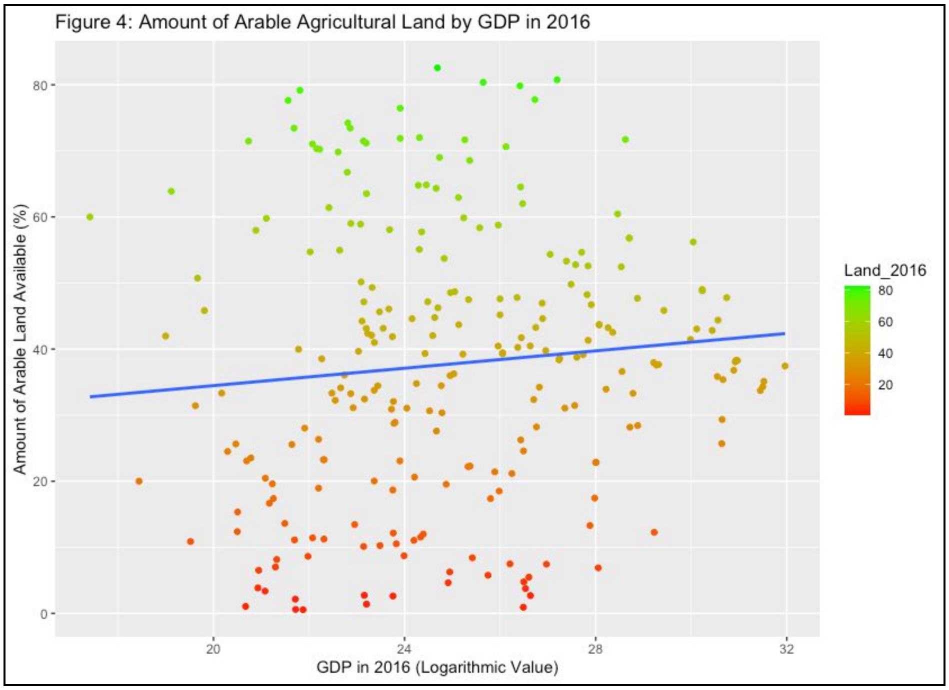
While these findings from the histogram and statistical summary diagrams so far do provide some preliminary results to my hypothesis that a country with a higher GDP will have a lower percentage of available land, more statistical analyses such as linear regression are needed to further confirm this notion.

III. Methods and Results

The first statistical analysis for this report is the linear regression, where a country's GDP in 2006 and 2016 is measured along with its corresponding percentage of availability of agricultural land. However, for this analysis, logarithmic regression was utilized for all GDP values to maintain similar comparative values; some GDP levels were in the millions while others were in the trillions, which would have resulted in an extremely skewed graph, so a logarithmic expression of GDP values for both years was utilized to counteract this problem.



Figures 3 and 4 both illustrate a scatter plot comparing a country's GDP in logarithmic value to its corresponding amount of arable agricultural land by percentage for all observations, where Figure 3 illustrates the observations in 2006 and Figure 4 illustrates the observations for 2016. Each point represents a country, and the color of a point shifts from red to yellow to green as the country's percentage of agricultural land increases as shown by the legend; for instance, whereas countries with available arable land 70% and above will be more green, countries with percentages 20% and below will be redder. For both Figures, the vast majority of points appear to be between 20 Logarithmic GDP and 28 Logarithmic GDP. Further, both Figures have a line of best fit modeled as a logarithmic function.



The lines of best fit for Figures 3 and 4 are shown in the regression tables in Figures 5 and 6 below, where we can see the linear regression of land availability by percentage of a country in 2006 and 2016 as a logarithmic function of that country's corresponding GDP level in 2006 and 2016.

Figure 5

	Dependent variable:
	Land_2006
Logarithmic_GDP_2006	0.536
	(0.440)
Constant	24.850**
	(10.838)
Observations	248
R^2	0.006
Adjusted R^2	0.002
Residual Std. Error	20.458 (df = 246)
F Statistic	1.486 (df = 1; 246)
Note:	*p<0.1; **p<0.05; ***p<0.01

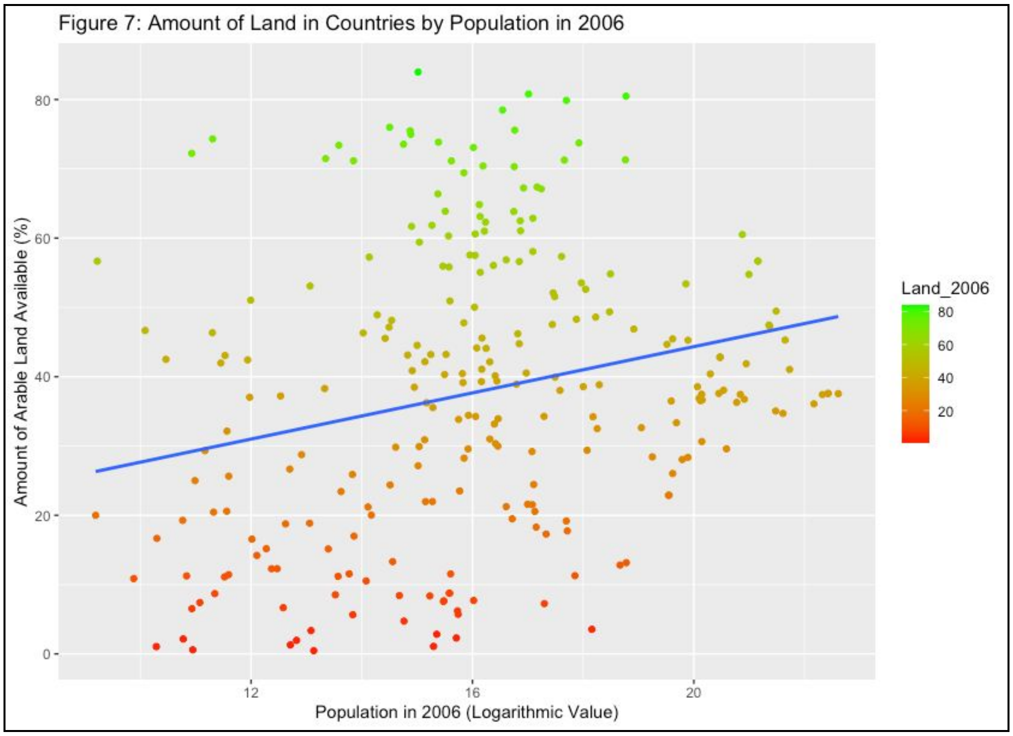
Figure 6

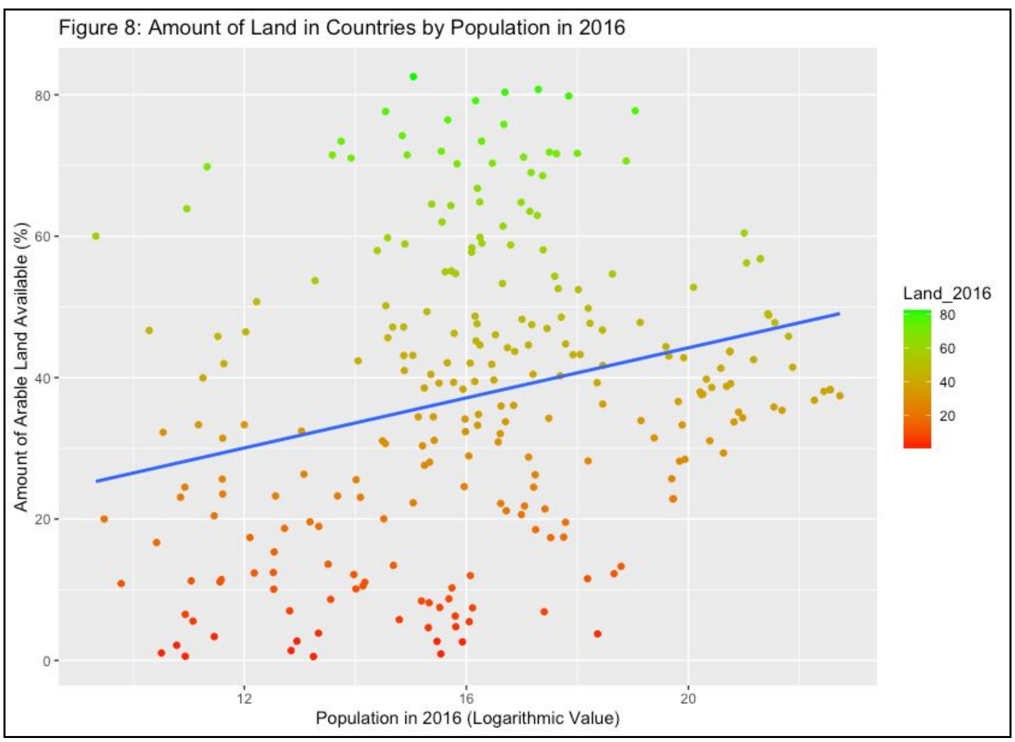
	Dependent variable:
	Land_2016
Logarithmic_GDP_2016	0.658
	(0.433)
Constant	21.310*
	(10.910)
Observations	245
R^2	0.009
Adjusted R^2	0.005
Residual Std. Error	20.322 (df = 243)
F Statistic	2.304 (df = 1; 243)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 5 outlines the linear regression-based logarithmic function for 2006 of all countries. We're able to see that a 1% increase in the logarithmic GDP for 2006 entails a 0.536% increase in the amount of land of that corresponding country. When a country's logarithmic GDP for 2006 is zero, then the percentage of available land for that corresponding country is 24.850% on average. The Residual Standard Error illustrates that the average percentage deviates from the true regression by 20.458%. The R-squared statistic, which measures how well the model fits the actual data by explaining if the regression explains the variance the dependent variable, is 0.006 or 0.6%, meaning that roughly 0.6% of the variance in the dependent variable (land) can be attributed to the independent variable (GDP). In other words, a 0.6% R-square outlines that GDP would likely not be the most important one that helps explain the variance of the availability of land if we were concerned with finding the most suitable property in assessing changes in agricultural land.

A similar analytical approach can be applied for Figure 6 for the 2016 year. For every 1% increase in logarithmic GDP entails a 0.658% increase in the amount of land within that corresponding country. A GDP logarithmic value of 0 would mean that the corresponding country's percentage of available land would be 21.310% on average. The Residual Standard Error, 20.322, tells us that the actual land availability value to a country's GDP would deviate from the true regression line for 2016 by 20.322%. The R-squared statistic is 0.009 or 0.9%, meaning that roughly 0.9% of the variance in the dependent variable (land) can be explained by the independent variable (GDP) in 2016; similar to the 0.6% R-square statistic, the 0.9% R-square outlines that GDP would likely not be the most important one that helps explain the variance of the availability of land if we were concerned with finding the most suitable property in assessing changes in agricultural land. Therefore, we can see that for both 2006 and 2016, there is an increase in the availability of agricultural land by 0.536% and 0.658% for every 1% increase in GDP.

Now, we turn to the confounding variable, the total population by country, and examine the linear relationship between a country's total human population and the amount of arable agricultural land. Again, the purpose of examining the human populations of each country is to see whether or not a spurious relationship exists, where GDP and agricultural land depend on this third property. Similar to GDP, the human population will be in a logarithmic form to maintain comparative values by numerical places. The charts below, Figures 7 and 8, illustrate a scatter plot comparing a country's population in 2006 and, separately, 2016 with its corresponding percentage of arable land. The colors of the plots each represent a separate country and shift from red to green as the amount of available agricultural land increases.





The lines of best fit as well as the logarithmic expressions for Figures 7 and 8 are outlined in detail on the regression tables in Figures 9 and 10 below for the years 2006 and 2016.

Figure 9

	Dependent variable:
	Land_2006
Logarithmic_Population_2006	1.667***
	(0.420)
Constant	10.992
	(6.859)
Observations	254
R^2	0.059
Adjusted R^2	0.055
Residual Std. Error	19.953 (df = 252)
F Statistic	15.789*** (df = 1; 252)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 10

	Dependent variable:
	Land_2016
Logarithmic_Population_2016	1.771***
	(0.415)
Constant	8.797
	(6.850)
Observations	253
R^2	0.068
Adjusted R^2	0.064
Residual Std. Error	19.823 (df = 251)
F Statistic	18.181*** (df = 1; 251)
Note:	*p<0.1; **p<0.05; ***p<0.01

In Figure 9, we see that for every 1% increase in human population for 2006 correlates with a 1.667% increase in a country's agricultural land availability, whereas, in Figure 10, every 1% increase in human population for 2016 correlates to a 1.771% increase. When a country's logarithmic population is zero, we see the percentage of available agricultural land being 10.992% in 2006, while in 2016 it would be 8.797. Thus, here we can clearly see that the change in percentage of land by population (1.667% and 1.771%) is greater than the change instigated by a country's GDP level (0.536% and 0.658%) as shown by Figures 5 and 6 between the ten-year period. Also, the total human population in each country as a variable has an R-squared statistic value of 5.9% in 2006 and 6.8% in 2016. In other words, 5.9% in 2006 of variances in land availability during 2006 can be attributed to total population, while 6.8% in 2016 of variances can be attributed to total population.

IV. Discussion

The findings of this report based the relationship between a country's GDP and availability of arable agricultural land disconfirms my original hypothesis that a country with a higher GDP will ultimately have a lower percentage of available agricultural land. We can see based on Figures 5 and 6 that in 2006, a 1% logarithmic GDP increase correlates with a 0.536% increase in land availability in 2006, and a 1% increase leads to a 0.658% increase in land availability in 2016. Thus, both years showcase an increase in agricultural land, where, in fact, we even see growth in percentage between the ten-year gap by 0.122%. Therefore, my hypothesis is further unsupported by the findings since, in addition to the fact that land availability increases with a higher GDP, we see that the percentage of land escalates between 2006 and 2016 by 0.122% as shown through the linear regressions.

While the report does illustrate that to some extent a higher GDP can lead to a higher percentage of land availability, it does not entirely confirm a unilateral direction between GDP and land. In other words, there is the possibility that the availability of land agriculture could be the independent variable, whereas the GDP level of a country could be the dependent variable. As mentioned earlier, the productions of supply and demand in agricultural economics partially rely on the amount of farming land that is available for producing agricultural goods (Johnson). If more land is available, then more crops, live animals, and pastures suitable for arable use by the Food and Agriculture Organization can bring about more economic production in a country and ultimately lead to a higher GDP. Overall, we can't conclude that GDP does unilaterally influence the amount of land when, in reality, there is the possibility that this direction can be reversed.

Furthermore, the findings from Figures 9 and 10 reveal that an increase in a country's total human population leads to greater growth in the percentage of land availability; in 2006 is a 1.667% increase and in 2016 is a 1.771% of land availability per every 1% increase in population, a 0.104% increase within the ten years. What is even more interesting is how the R-square statistics for the total human populations, 5.9% in 2006 and 6.8% in 2016, is higher than the R-squared statistics for GDP, 0.6% in 2006 and 0.9% in 2016. This difference in the R-squared statistics reveals that the confounding variable, which has the higher overall R-squared statistics, is influencing the variance of the agriculture more so than a country's GDP levels. If we wanted to see which variable has a higher effect on land availability, then a country's total human population would've been the variable with more influence on the dependent variable, but this report was focused on land availability through an economic perspective. Moreover, the total human population variable could be influencing the GDP and the availability of agricultural land variables, which would be considered a spurious relationship since a third property would be influencing both the independent and dependent variables. In other words, the variable of human population could have affected GDP of a country and agricultural land separately. The rationale for bringing up this confounding variable is to assess whether or not an alternative explanation exists in exploring how and why agricultural land either increases or decreases.

Within the scope of this study, there are some limitations to the findings. First, while we can see that a rise in GDP correlates to a rise in land availability, we can't state that a higher GDP causes higher land availability. In other words, there could potentially be other spurious relationships from natural phenomenons that lead to a change in land availability other than

GDP. For instance, outside factors such as land deforestation, urban expansion, and natural disaster in rural areas that lead to a decrease in land availability are not taken into account in this study. Another limitation is that this report also does not consider the implementation of new or improved public policies and legislation from countries. For example, in 2014, the Agricultural Act of 2014 was signed into law on February 7, which expands upon programs such as commodity support, nutrition assistance, and conservation by expanding agricultural programs in rural areas pertaining to specialty crops, bioenergy, and rural development (Agricultural Act of 2014). In other words, the expansion of rural land for arable use was one of the key priorities in expanding the agricultural economic framework of the United States. Such policies would not be taken into account, where an actual piece of legislation contributed to the growth in land availability rather than a country's GDP.

V. Conclusion

Throughout this report, we can see that a higher GDP level of a country does correlate to higher availability of agricultural land. It has also shown that examining the total human population and its effect on land has more influence in causing the variance in the availability of agricultural land. While this research does take into consideration how GDP can influence land availability, the statistical analyses do not explain why or how this phenomenon exists in the first place. Also, this research does not take into account a country's practices and legislation pertaining to agriculture since the amount of arable land is assessed by examining the country's corresponding GDP level. All in all, more extensive research that examines all sorts of independent variables in place of GDP should be done to see what factor contributes the most to a shift in land availability in a country. Since agriculture, especially within the scope of available arable land, is a complex framework, continuing to test our different independent variables, just as this report did with the total human population of each country, is necessary to examine which factor contributes the most to changes to agricultural land availability.

References

Agricultural Act of 2014: Highlights and Implications. (n.d.). Retrieved from https://www.ers.usda.gov/agricultural-act-of-2014-highlights-and-implications/

Atlas of Urban Expansion. (n.d.). Retrieved from https://www.lincolninst.edu/research-data/data/atlas-urban-expansion

De Sormeaux, A., & Pemberton, C. (2017, June 2). Factors Influencing Agriculture's Contribution to GDP: Latin America and the Caribbean

FocusEconomics. (2017, December 18). The World's Largest Economies (2019-2023). Retrieved from https://www.focus-economics.com/blog/the-largest-economies-in-the-world

Gross Domestic Product. (n.d.). Retrieved from https://www.bea.gov/resources/learning-center/what-to-know-gdp

Johnson, D. G. (2019, January 16). Agricultural economics. Retrieved from https://www.britannica.com/topic/agricultural-economics

World Development Indicators. (n.d.). Retrieved from http://datatopics.worldbank.org/world-development-indicators/

Zavatta, G., Perrone, T., & Figus, C. (2014, October 27). Agriculture Remains Central to the World Economy. 60% of the Population Depends on Agriculture for Survival. Retrieved from http://www.expo2015.org/magazine/en/economy/agriculture-remains-central-to-the-world-economy.html