

# The Impact of Colonization on Home Country Wealth

**Robert P. Rogers**  
**Ashland University**

*In this paper, a model is developed to estimate the impact of conquering and colonizing other countries on the per capita income of the mother or metropole country. The estimated impacts of the model are positive and quite large, the increases running between 14.2 and 78.3 percent of the total average income. While the omission of certain variables could have biased the results, the employment of Bayesian techniques indicates that the results are insensitive to the use of a large number of model specifications.*

*Keywords: Economic Growth, Bayesian Analysis*

## INTRODUCTION

The history of the world between 1700 and 1970 was characterized by a great increase in total economic product. This growth has been especially large in certain European, North American, and Oceanic countries. In this period, a number of countries mainly European annexed many distant countries -- usually on different continents. For this paper, the policy is called "colonialism," and the conquered countries are called colonies. This paper raises the question of whether the inhabitants of the colonizing nations as a whole benefitted from the colonies.

A large literature has arisen on colonialism which attributes its rise to many causes, but there has been disagreement on its effects. Some scholars have attributed the economic growth of many countries to the colonizing policy. Certainly, wealth could very well have been distributed from the colonies to the colonizing or metropole country. The American sugar, cotton, and gold colonies possibly enriched the mother countries (Pomeranz, 1999). In contrast, some assert that the colonial policies have had little or no impact of the growth on the colonizing nations.

Still others have argued that colonial policies could have attenuated the colonizing nation's economic growth. In the last three hundred years, the two earliest European colonial countries, Spain and Portugal, were not noted from economic growth. The argument behind this claim is that groups controlling the governments of the metropole countries used the resources of the home country to set up colonies that were beneficial to them and/or certain small groups but not necessarily to the home country citizens as a whole. Certainly, some of the resources absorbed in the colonizing effort could have been diverted from not only home country consumption but also from domestic investment. The question addressed here is, then, did these annexations add to the total home or mother country income. This paper, thus, raises the question of whether the inhabitants of the colonizing nations as a whole could have benefitted from the colonies.

To accomplish this task, I do the following. After examining the history of modern colonialism, I outline the major theories on its causes. The hypothesis that the mother country did not benefit from colonialism is not inconsistent with many of them – often even ones that conflict with each other. Next, I develop a reduced form model for the level of mother country income per capita, **GDPC**, for a sample consisting of certain countries for selected years running from 1600 to 1965, the major criterion for selection being data availability. I use the Maddison data on GDP (Maddison, 2001, and Bolt and Van Zanden, 2014) and other economic variables to test the hypothesis that the colonial endeavor led to changes in the total wealth (either positive or negative) as measured by the **GDPC** of the home country. A Bayesian methodology is used which is atheoretical in that it makes no stringent assumption about the underlying model of colonizing behavior. It estimates a large number of aggregate GDP models to test the hypothesis of colonizing activity lowering (or raising) GDP under a large number of different assumptions. This method answers this question --- are the results robust.

In what follows, the paper, first, summarizes of the history of modern colonialism. The third section, then, gives a brief overview of the most important literature on the subject. Then, the next two sections develop a reduced form econometric model for the impact of colonialism on GDP. The sixth section describes the results, and finally the last section concludes.

## **HISTORY OF THE COLONIAL METROPOLES**

This section consists of a short description of the major colonial empires in the last 300 years. The countries heading these empires constitute the colonialist part of the sample. With three exceptions, the mother countries or metropolises were European. The height of this colonial development came in the early 20th century. In all, there were twelve important colonial empires: the mother or metropole countries being Portugal, Spain, France, Great Britain, the Netherlands, Russia, Belgium, Germany, Italy, Turkey, the United States, and Japan. These country's empires rose and fell at different times between 1500 and 2000. To put the empirical model in context, -- especially the definition of the variables -- I now outline the history of these empires.

The earliest of these empires to develop was that of Portugal. In the 1400s, the Portuguese started with explorations in the Atlantic and Indian oceans; they were trying to find all-water routes to eastern Asia and the East Indies where valuable spices and other products could be procured. Due to this search, the Portuguese annexed territories in not only eastern Asia such as Goa, Macau, and East Timor but also in areas in other parts of the world like Brazil in South America and Angola and Mozambique in Africa. Brazil separated from Portugal in the early 19th century, but the Asian and African colonies separated from Portugal only in the late 20th century. Thus, for every observation in the sample period (selected years being between 1600 and 1965) Portugal is classified as a colonial metropole or mother country.

The second modern colonial empire to develop was that of Spain. It arose out of the voyages of Christopher Columbus. Columbus oversaw the establishment of colonies in the West Indies. Others established Spanish colonies in Central and South America where some found wealthy civilizations and precious metals (mainly in Mexico and Peru) from which metropole Spain could draw great wealth. Additionally, the Spanish expanded into North America, the Philippines Islands in the Far East, and northern Africa. Through a series of events, this empire was lost by 1900 except for some fragments in Northwest coast of Africa. These areas were taken over by Morocco, a neighboring African nation in the late 20<sup>th</sup> century. Thus, Spain is classified as a colonial nation for the entirety of the sample.

The third empire in the sample is the French Empire. It started with the colonization of Canada, parts of India, and some of the West Indies in the 17th Century. With military defeats in the 1700s and early 1800s, France lost most of these colonies. In the last 70 years of the 19th century, however, France gained another set of colonies in Africa and southeastern Asia. Also, as a result of World War I, France gained the additional colonies of Lebanon and Syria in the Middle East and the Cameroons in Africa. After World War II, independence movements led to the liberation from France of most but not all of its colonies by the 1970s. Thus, for my model France will be treated as a colonial metropole from 1700 to 1965.

The fourth empire in this sample was the largest, the British Empire. In the 17th and 18th centuries, Britain acquired colonies in North America, the West Indies, and India. After Britain lost its most populous American colonies, the United States, it enlarged its holdings in India and established other colonies in Burma, Malaysia, Sri Lanka, Australia, New Zealand, Latin America, and Africa. After World War II, most of its colonies became independent. Since it continued to hold some colonies, Britain is classified as a colonial power throughout the sample with the exception of 1600 which was before its empire started.

The Russian Empire differed from the other European empires in that all its conquests were contiguous to either its original territory or lands that had been previously annexed. By 1850 it had acquired not only most of the northern part of the Eurasian landmass but also Alaska in North America. The latter it sold to the United States in 1868. In most of these lands, the bulk of the inhabitants were not Russian. Since most of the Russian conquests were still held in the 1960s, it is classified as a metropole throughout the sample.

In the 1600s, the era of its commercial and technological dominance, the Netherlands acquired colonies in many parts of the world. Among these lands were South Africa, Sri Lanka, parts of South America, and most of Indonesia. While later it lost the lands in Brazil and Africa and Sri Lanka, it retained Indonesia until 1949 and part of the island of New Guinea until 1969. Thus, the Netherlands is classified as colonial for all the years from 1700 to the end of sample, 1965.

Germany was united into one country in 1871. Having a large population and large modern economy, it soon acquired colonies in Africa and Oceania, one colony being part of New Guinea. At the end of World War I, it lost its colonies, and therefore, Germany is classified as a colonial power for only 1900 and 1913.

Belgium only became independent in 1830, but it acquired a large African colony, the Belgian Congo, in the 1880s, and it controlled that land until 1962. Thus, for the period from 1880 until 1960, Belgium is treated as a colonialist country.

Italy was only united into one country in the 1860s, but it soon acquired three colonies in Africa: Libya, part of Somalia, and Eritrea. After World War II, these lands were alienated from Italy. Libya and Somalia became independent, while Eritrea was merged into Ethiopia. Thus, from 1945 on, Italy is not treated as a mother country in the model. It is classified as a colonial metropole from 1900 to 1940.

The three metropole countries that were not European are Turkey, the United States, and Japan. Turkey arose from the Byzantine Empire in the 14th and 15th century totally conquering it in 1453. It acquired lands in the Balkan part of Europe, the Middle East, and northern Africa. While the annexation of the lands in the Balkans and Middle East was effected by land arms, the colonies in North Africa were acquired partly by naval operations, and they were at least partly motivated by trade. This is consistent with the methods of the other colonial metropolises. In the late 19th century, the bulk of the Balkan and African lands were taken away from Turkey, and later in the early 20th century, the Middle East lands were lost to Turkey except for the parts of Anatolia where the bulk of the population is Turkish. Consequently, in the model, Turkey is treated a colonial power only before 1920.

The United States gained about half of its present continental territory on independence. Through either purchase, direct annexation, or conquest, it acquired the rest of its North America territory. Today, most of the inhabitants of these areas are U.S. citizens. In the late 19th century, the United States acquired territories that were never incorporated into its union. One, the Philippines, was made independent in 1945; the others, Puerto Rico and numerous islands in the Pacific, are still held by the United States. Consequently, the United States is treated as a colonial power from 1899 until the end of the sample.

The last country to be classified as colonialist is Japan. It acquired Taiwan, Korea, Manchuria, other parts of China, and a number of Pacific islands in the early 20th century. They were taken away at the end of World War II. Consequently, Japan is classified as a colonial metropole from 1900 to 1945 and as non-colonial for the remainder of the sample.

## LITERATURE ON COLONIALISM

In the last 150 years, a literature has developed on the interconnection between colonization and national wealth. The major scholars on this relationship were John A. Hobson, Vladimir Lenin, and Joseph Schumpeter. (See Wright, 1961.) Hobson argued that certain elements in the capitalist class would persuade the government to acquire colonies. These colonies would benefit certain individuals in the class but not necessarily the home economy as a whole. Hobson posits that the skewed distribution of incomes prevented the poorer classes from purchasing the goods that a growing capitalist system could provide. This condition led to a need for an outlet for the surplus capital that a market economy could produce. Thus, capitalists wanted to find a place to invest. In a later revision of his views, Hobson came to see the benefits to the home country of international (including colonial) finance, and trade with an undeveloped country.

Vladimir Lenin adopted much of Hobson's theory to the grand view of the ultimate development of the capitalist system as envisioned by Karl Marx. Lenin saw the capitalist system as a structure that would ultimately destroy itself. As investment opportunities dried up in home countries, capitalists would seek opportunities in other countries. Often, they needed the help of government to obtain access to these investment opportunities. This led the capitalist ruling classes to pressure the government to conquer and set up colonies. These colonies would provide investment opportunities for the metropole country capitalists.

In contrast to Hobson and Lenin, Joseph Schumpeter believed that imperialism or 19th colonialism was not endemic to the capitalist system. Colonialism itself and the administrative, military, and naval activities and skills needed for colonialism did not necessarily come from the capitalist class. Rather these skills and propensities came from the old European ruling classes who controlled society when it was based on a rural and feudal economy.

Interestingly, only Schumpeter of these three people would make a prediction on the impact of colonialism of the metropole country. He would assert that the resources diverted to colonialism would attenuate the development of the industries supplying the home market. The increase in taxes and government spending from colonialism would be taken out of the nation's domestic economic product. Hobson and Lenin seem ambivalent on the impact of colonial investment on home country wealth. On the one hand, the income distribution is so skewed that even if the resources expended on colonies remained at home they would not necessarily have gone for wealth-increasing activities. On the other hand, the increases in the wealth of the ruling class from the diverting of investment to colonies could have added to total home country income. Thus, the predictions of the first two writers on the impact of colonialism on total home country GDP are ambiguous.

Later writers have found much to criticize in the theories of Hobson, Lenin, and Schumpeter. The major problem that many of these writers have had is that the bulk of the foreign investment of Britain, France, the United States, and some of the other colonial powers went to places other than their colonies. (See William Langer in Wright, 1961.) This seems to support the Schumpeter hypothesis that colonialism did not arise from the capitalist classes but from the other elements of society that still exercised control over the metropole country governments. This fits well with Great Britain where many of the people in its government did come from the old aristocracy, examples being Lord Salisbury and Randolph and Winston Churchill. Thus, the metropole governments were often influenced on colonial policy by people other than capitalists and financiers. Colonialism lowering home country income is consistent with much of the information provided in recent years.

Most of the theory and history of the colonial movement can be put in context by seeing the metropole governments as arenas where different interest groups compete for the use of government coercive power to further their ends. A capitalist individual or group might have an interest in a given country, but without the backing of the metropole government they could not earn a good return. Examples of these situations can be found among all the colonial nations. For instance, a French group wanted to build a railroad into China using the port of Hanoi and the Red River which would be secure only if the French government controlled the area. Almost all the activity of Cecil Rhodes in Africa was

predicated on British control of the areas in which he operated. Thus, capitalist groups could prevail on governments to annex areas where they operated.

Additionally, this interplay between interest groups and governments can account for the actions of other metropole country groups desiring colonies. This is consistent with Schumpeter's idea of how non-capitalist classes or groups would influence governments. Interest groups arise from not only financial incentives but also from the customs of different sorts of people. Governments can also be influenced by people merely practicing their professions such as army and navy officers, explorers and scientists.

Lastly, a recent theory puts colonialization by certain European countries in a different light. Kenneth Pomeranz (2000) posits that the resources from the British (and to a lesser extent the French and Dutch colonies) in North America was the impetus for the European industrial revolution. With the resources from these colonies, mainly cotton and lumber, and the coal available in Britain, the European countries especially England had the wherewithal to improve and expand their manufacturing industries. Thus, this availability of the cotton and coal allowed for the investments in advancing technology thereby giving certain northern European countries the ability to improve their production system.

In all these theories, the control of governments by interest groups pursuing their own ends be it financial success, professional advancement or some other individual goals is consistent with any income impact of colonialism on the mother country, negative or positive.

## THE MODEL

In this section, I develop a model to estimate the impact of colonialism on the GDP per capita of the metropole or mother countries under a wide range of different colonialism theories. The basic theoretical construct I use comes from the Solow (1957) production function model with population and capital on the right hand side of the equation. A reduced-form model of this aggregate production function is used to test my hypotheses on the impact of the colonialism on the GDP per capita (**GDPC**) of the metropole countries. While the Solow model includes variables exogenous to the production function, some of its variables are endogenous, and thus, they would not appear in my reduced form model. A reduced form production function model, then, is assumed to explain **GDPC**. This **GDPC** and other variables pertain to the countries as they were geographically located in 2000.

Angus Maddison and his associates have developed the relevant income variables for the past for the areas occupied by the sample countries in 2000. In this data set are available 887 observations for 73 countries. (An Appendix showing the sample countries and years can be obtained from the author.) The sample consists of observations from the Maddison data (Maddison, 2001 and Bolt and Van Zanden, 2014) for selected years between 1600 and 1965: these years being 1600, 1700, 1820, 1850, 1870, 1900, 1913, 1920, 1925, 1930, 1935, 1940, 1950, 1955, 1960, and 1965, when available. If the **GDPC** observation is not available for the given year, then, either it is not included in the estimation model or a nearby available year is used. Thus, I use GDP per capita, **GDPC**, on the left hand-side. On the right-hand side, I use a set of explanatory variables. They are divided into the three classes described below.

With this model, the impact of colonialism on GDP per capita is analyzed. To do this, I use a combination of Extreme Bound Analysis (**EBA**) and Bayesian Averaging of Classical Estimates (**BACE**). (See Granger and Uhlig, 1990, Hoeting, Madigan, Raftery and Volinsky, 1999, Levine and Renelt, 1992, Madigan and Raftery, 1994, Raftery, Madigan and Hoeting, 1997, Sala-i-Martin, 1997, and Sala-i-Martin, Doppelhofer, and Miller, 2004.) Before describing the details of these methods, I clarify matters by laying out the econometric model.

The dependent variable, **Y**, is the GDP per capita or **GDPC**. The independent variables are divided into three categories. The first consists of the vector, **Z**, which according to theory is included in every **GDP** model, and in the second category, **X**, are the variables that may or may influence GDP in the sample. The third category consists of a dummy variable, **COL**, equaling one for the colonial metropole countries, and zero, otherwise. The model, then, can be displayed as follows:

$$Y = \alpha_0 + \alpha_1 Z + \beta X + \rho COL + \mu. \quad (1)$$

Here  $Y$  equals GDP per capita, and  $\alpha_0$ ,  $\alpha_1$ ,  $\beta$ , and  $\rho$  are parameters, and  $\mu$  is the residual. Before discussing the particular variables in the model, an estimation problem should be addressed. While the act of adding colonies may impact on the **GDPC** of a country, the country's income could affect its ability to set up colonies making the incidence of colonization a function of the country's **GDPC**. Consequently, the above model may have an element of simultaneity, and it may be necessary to use an instrumental variables estimation technique. To do this, one can develop a model of the act of colonization. Thus, I hypothesize a model for a country's incidence of colonization, and develop the variables for such a model. One idea implicit in my theory is that the **GDPC** impacts on colonization. Since I am using a reduced form model for **GDPC**, it seems reasonable to use a reduced form model for the colonization equation.

To identify an estimable instrumental variables model, I need to include some variables affecting **COL** but not **GDPC**. Then, one can hypothesize this model for the incidence of colonization.

$$\text{COL} = \delta_0 + \delta_1 \mathbf{Z} + \delta_2 \mathbf{X} + \delta_3 \mathbf{W} + \varepsilon. \quad (2)$$

Here the  $\delta$ 's are the parameters, while  $\mathbf{X}$  and  $\mathbf{Z}$  are discussed above, and  $\mathbf{W}$  consists of the variables in the **COL** model not included in the **GDPC** equation. The estimating equation for **GDPC** will include predicted value of the above **COL** equation in place of the raw **COL**. This predicted value for **COL** that can be used is as the instrument in the estimating equation. (See Madalla, 1978 and Heckman, 1978.)

To assess the impact of colonization on **GDPC**, I use a combination of Extreme Bounds Analysis (**EBA**) and Bayesian Averaging of Classical Estimates (**BACE**). A model is estimated for every combination of the  $\mathbf{X}$  variables. With these estimates, an analysis is performed for each rendition of the model. For **EBA**, the coefficients for **COL** for each of the models are examined to see if the largest and smallest of the coefficients are significantly different from zero in the direction predicted by theory. If they are, there is strong evidence for the hypothesis. If they are not, or if they have different signs from the theory prediction, it mitigates against the hypothesis.

The second part of the analysis is Bayesian Averaging of Classical Estimates (**BACE**). (See Raftery, Madigan and Hoeting, 1997, Sala-i-Martin, 1997, and Sala-i-Martin, Doppelhofer, and Miller, 2004.) With **BACE**, one calculates weighted averages of **COL** and its standard error and ascertains whether the average is significantly greater than zero in the direction hypothesized. If there are  $k$  variables in vector,  $\mathbf{X}$ , there are a total of  $2^k$  models including every configuration of the  $k$  vector variable plus the model with no  $\mathbf{X}$  variables included. Essentially, with **EBA** and **BACE**, the researcher runs  $2^k$  variations on the model, each of which consists of the base variables,  $\mathbf{Z}$  and **COL**, and a given combination on the  $\mathbf{X}$  variables (including a model with no  $\mathbf{X}$  variables).

The goal is to ascertain whether **COL** has statistically significant impacts on **GDPC** for the large number of combinations of  $\mathbf{X}$ . Thus, the exercise consists of estimating models with variations on which of the  $\mathbf{X}$  variables are included to see whether **COL** has a consistent impact in large variety of models. For **BACE**, then, several statistics are computed for the model; the most important is the weighted average of the **COL** coefficients and the standard errors, the weights being derived from the square of the sum of the standard errors for each equation. (See Sala-i-Martin, et al, 2004, p. 815-819.) The other statistics used are discussed in the results section.

## THE VARIABLES

Here, I discuss the variables used in the model. Scholars have examined the addition of many political and social variables to the Solow production function to determine the causes of economic growth. (See McGuire and Olson, 1996, Owen, Videras, and Davis, 2009, Johnson and Koyama, 2017, Ramon Moreno, and Trehan, 1997, Gassebner, Gutmann, and Voigt, 2016, and Auer, 2013.)

The dependent variable is **GDPC**, the Gross Domestic Product per capita in 2011 United States dollars for each observation nation and year. This statistic has been adjusted for purchasing power parity.

Next, I describe the variables that will be included in all renditions of the econometric model called here the **Z** vector. One can safely assume that the physical aspects of a country impact on per capita GDP. The first representing the physical situation of a nation is its area, **AREA**. More land area means more physical resources to produce its GDP. Since the GDP data collected by Maddison et al (2001) refer to the present areas encompassed by the sample country, the issue of simultaneity between physical size and GDP is not relevant. Climate can also affect a country's productivity. **LAT** is the absolute value of latitude which represents the distance from the equator, and thus a reasonable proxy for climate.

Another variable representing the physical environment, **ELEV**, is the elevation of the central point of the country. This affects the capacity of a country to produce output at given latitude and given level of technology. In the far north or south, higher elevation could lower productivity, while for countries near the equator it might increase productivity.

Since the most efficient way to move large quantities of goods is by sea, landlocked countries have less access to trade. This was especially true in the 17<sup>th</sup>, 18<sup>th</sup> and early 19<sup>th</sup> centuries before the advent of railroads and large mechanized ships. Therefore, a dummy variable, **LANDL**, equaling one for landlocked countries and zero otherwise is included.

Because the sample consists of a panel data set, dummy year variables, **YRDUM** for the years, 1600, 1700, 1820, 1850, 1870, 1900, 1913, 1920, 1925, 1930, 1935, 1940, 1950, 1955, and 1960 are included in the model. For some countries, certain of these years are not available, but years close to them are used. Thus, when available, I use those close years as substitutes, and define them the same as the unavailable year.

The variable used to represent a country being a colonial metropole is **COL**. It is a dichotomous variable equaling one for the twelve above described metropolises and zero, otherwise. Again, the mother metropole countries are Portugal, Spain, France, Great Britain, the Netherlands, Russia, Germany, Belgium, Italy, Turkey, the United States, and Japan for sometimes the whole and sometimes only part of the sample period as discussed in section II. Some of these countries did not have colonies in some years in the sample.

Now I examine the variables in the **X** vector. Most of these variables have been proposed by economists and historians, but there may not be general agreement on their relevance. Thus, as mentioned above, models are estimated for every combination of these variables. Since there are eight **X** variables, 2<sup>8</sup> models (256) are estimated for each general specification (discussed below). These **X** variables and the rationale for their use are now discussed.

One variable in **X** is **COLE**, a dummy indicating whether the observation country is a colony or not. It may or may not be included in the model. There are several theories on what might happen to a country on becoming a colony. Of course, there may very well be wealth transfers from the colonies to the metropolises, but other things can happen. Some colonies were attractive to the inhabitants of the metropole country, and thus they were settled by colonists from those countries. Often, these colonies would have a higher income than the other sample countries.

Other countries that became colonies were unattractive to the citizens of the metropole countries, and arguably they were exploited for their human and natural resources. Thus, many of the natives were enslaved either at home or elsewhere. Thus, it is likely that becoming a colony could lower a country's income. In contrast, even with the exploitation of the people and resources the metropole country colonization may raise income because they introduce superior technology. Thus, the coefficient of **COLE**, the impact of colonization on the colony's income, could be negative or positive.

Another possibility is that colonization could lead to a skewed income distribution. In some cases, the overall wealth of the country might increase on being colonized, but the wealth and income could be distributed in a skewed manner with a small very wealthy ruling class usually consisting of metropole country born citizens ruling over a very large poor proletarian or slave class. Among the overall richest countries in the world in the 18<sup>th</sup> century were Haiti and other Caribbean Sea sugar exporters where the bulk of the population was enslaved. Thus, models are estimated both with **COLE** included and excluded.

Among the **X** variables would be those depicting the position of the state. For many countries, the most relevant restraint on the state power is the ability of subjects to migrate to other countries or

jurisdictions. The small size of some countries and the large number of countries in relatively small space mean that there are a large number of different governments in a relative short distance from each other. This situation gives the subject opportunities to move, if they were not satisfied with the situation in their own country. (See Jones, 2007, Karayalcin, 2008, and Chaudry and Garner, 2006.)

In this paper, I depict the situation with a variable representing the ability of the subjects and citizen to leave a given jurisdiction: **NEIGHNO**. It equals the number of different independent governments within 1000 miles of the center of the observation country. It roughly indicates the number of countries to which it is feasible for a subject to migrate if they are dissatisfied with their present situation. These alternatives not only restrain the ability of a given government to tax its subjects but also increase the incentives of a government to produce valuable public goods.

Therefore, the higher **NEIGHNO** the greater the immigration choice of subjects or citizens and the lower the cost of immigration. It portrays the ability of subjects to migrate, and thus, it depicts the limits on the ability of states to tax their subjects. Furthermore, it also gives some indication of the choices they face. A larger choice of places to which subjects can migrate will attenuate the ability of governments to overtax and encourage it to produce more public goods. Thus, a higher **NEIGHNO** should increase country's **GDPC**.

Biological research has suggested another hypothesis. It is that the genetic composition of populations has affected economic growth in the last 10,000 years. Ashraf and Galor (2013) find a connection between the genetic diversity of a nation's population and its GDP per capita. Essentially, as population becomes more genetically diverse, GDP per capita rises up to a point. At a certain threshold, however, the relationship becomes negative. Thus, genetic diversity has a curvilinear relationship to economic success. It has been found that this genetic diversity is negatively correlated with the distance the population and its ancestors have traveled from the eastern African origin of the human species. The sub-Saharan African populations have the most diversity, and the Pacific and western hemisphere populations have the least.

The wealthiest countries tend to be in Eurasia with middling amount of genetic diversity. Ashraf and Galor (2013) exploit that this relationship to posit a curvilinear relationship between economic success and genetic diversity. Therefore, I use the relationship to analyze the impact of genetics on economic performance.

To obtain this statistic, I measure the distance of each observation country from Addis Ababa, Ethiopia by the circuitous route that their ancestors hypothetically took from East Africa. Essentially all the populations in the sample except for those in sub-Saharan Africa are hypothesized to have gone from Addis Ababa to Cairo, Egypt and then spread out over the globe by a set of different routes. There are only two Sub Sahara countries in the sample, Ethiopia and South Africa; their situation is discussed below.

The populations in North Africa and the Middle East and South Asia all the way through India to Thailand are assumed to have migrated from Ethiopia to Cairo and then from Cairo to each of their population centers. The populations of European countries are assumed after coming from Ethiopia to have migrated from Cairo to Istanbul and then their eventual countries. The populations of East Asia and Oceania are assumed (after arriving in Cairo) to have migrated from Egypt to Phnom Penh, Cambodia and then to their eventual population center. The populations (pre-Columbian) of the western hemisphere are hypothesized to have gone from Cairo to Anadyr, Siberia and then to Prince Rupert, Canada and hence to their final destination population center.

To depict this situation a variable, called **GENETIC**, is included in the model. It equals the total distance in miles between Addis Ababa and the population center of each observation country by the routes described above. For France, **GENETIC** would be the distance from Addis Ababa to Cairo plus the distance from Cairo to Istanbul plus the distance from Istanbul to Paris. For China, **GENETIC** would be the distance from Addis Ababa to Cairo plus the distance from Cairo to Phnom Penh plus the distance from the latter city to Beijing.



The two sub-Saharan countries in the sample, Ethiopia and South Africa, are treated as follows. Since Addis Ababa is the starting point of the hypothetical sample, Ethiopia's **GENETIC** would be zero. For South Africa, this variable would be the distance from Addis Ababa to its central point.

Ashraf and others, however, have hypothesized a nonlinear relationship between, **GDPC** and **GENETIC**. This would call for a quadratic form. Since distance and genetic diversity are negatively correlated, the coefficient for **GENETIC** would be negative, but its square, **GENETIC2**, is added to the model. Its coefficient would, then, be positive. Thus, the curvilinear nature of this relationship can be depicted.

A dummy variable, **EURA**, equaling one for countries on the Eurasian east-west axis and zero, otherwise is included in the model. It is based on Jared Diamond's geographic theory (1997). He posits that the shape of Eurasia has affected the dispersal of technology to the benefit of countries in the area between Portugal on the west and Japan and the Siberian part of Russia on the east. In that part of the world, neighboring countries are apt to have similar climate conditions, and therefore, they can more readily adopt the technology developed in nearby countries. This facilitates the transfer of technology. In the western hemisphere and the part of Africa south of the Sahara Desert, the geographic axis is north-south and often neighboring countries have radically different climatic conditions. Consequently, they may not be able to adopt new ideas from their neighbors.

My sample has seven countries in Africa, Egypt, Libya, Tunisia, Algeria, Morocco, Ethiopia, and South Africa. Since the first five face the Mediterranean which gives them access to Asian and European countries often with advancing technology, they are included with Eurasian countries. The other two, Ethiopia and South Africa, are not.

Another theory has been suggested by the literature (Jones, 2007 and Karayalcin, 2008). It posits a bargaining relationship between each country's government and its population. The rulers provide the services from government in exchange for taxes. Accordingly, the subjects pay taxes for the desired services provided by the government rulers. In most all situations the government is a monopoly provider giving it considerable if but not necessarily complete market power. The power of the ruler is constrained by a number of factors. First some states have constitutions that limit their power. These states have less market power than others. Second the managers of the states often recognize that a too high price for state services can result in less revenue; thus, some situations call for restraint in taxes (Van Zanden, 2009). These types of states seemed to have developed in Europe for most observations. The populations of four former European colonies, Canada, Australia, New Zealand, and the United States, were made up mostly of emigrants from northern Europe, and they have evolved into European type societies. Consequently, they are classified as European societies for this variable.

Thus, scholars argue that a dummy variable, **Europe**, being one for European countries and countries originally settled by northern Europeans and zero, otherwise, should be added to the **GDPC** model.

Finally, a theory has been developed that posits a curvilinear relationship between **GDPC** and the distance of country from the technological leading nation in the world. If a nation was close to the technological leader, it could better use the leader's knowledge than could more distant countries. But the countries farthest away became more self-reliant, and consequently they may have advantages over some countries closer to the leader (Ozak, 2018). The idea behind this theory is countries closer to the technological leader can appropriate its technology sooner than other nations. The countries farthest away are the most self-sufficient and they may be able to learn and invent better methods than some of the nearer countries.

This implies a curvilinear relationship between the **GDPC** of a country and its distance from the world technological leader. Therefore, the variables, **TECHD**, equaling the distance between the observation country and the technological leader and **TECHD2**, the square of **TECHD** can be added to the model. The identity of the technological leader changed over the sample. For the years, 1600 and 1700, the technological leader was the Netherlands, for 1800, 1820, 1850, and 1870, it was Great Britain, and for the rest of the sample it was the United States.

The below equation, then, summarizes the variables in the full **GDPC** equation:

$$\text{GDPC} = f(\text{AREA, LAT, ELEV, LANDL, YEAR Dummies, COL, COLE, NEIGHNO, GENETIC, GENETIC2, EURA, EUROPE, TECHD, TECHD2}). \quad (3)$$

Since I am estimating some **GDPC** models using the instrumental variables, I need to define the elements of **W**, the variables in the **COL** equation not in the **GDPC** model. One variable in this equation represents the opportunity of a country to set up colonies overseas is the location of the country on the Northern Atlantic Ocean, **BOCEAN**. The nations on this ocean developed ships capable of crossing large bodies of water. This made faraway lands available for their conquest or predation. (In the 14<sup>th</sup> and 15<sup>th</sup> centuries, the Chinese developed large ocean-going ships, but the effort was abandoned in the 1430s due to political considerations.) Furthermore, Atlantic Ocean mariners developed methods of using the winds and currents to facilitate the movements across the oceans. (See Crosby, 1977, p. 104-131.) Of the colonizing nations, Portugal, Spain, France, Great Britain, France, and the United States border directly on this ocean. Notice that still six of the colonizing nations, the Netherlands, Russia, Germany, Belgium, Italy, Turkey, and Japan do not border directly on the North Atlantic Ocean. Thus, being on the North Atlantic is not perfectly collinear with being a colonial metropole.

The second variable in **W**, **NAVY**, is the importance and standing of a nation's navy. Of the colonizing nations, eleven had strong navies for at least part of the sample, Portugal, Spain, France, Great Britain, the Netherlands, Russia, Germany, Italy, Turkey, the United States, and Japan. Not all of these nations had large navies for the whole period; in fact, only two, Great Britain and France, had important navies for the entirety of the sample period. Thus, the variable, **NAVY**, equaling one for a country with a large navy and zero, otherwise is used in the IV estimating equations.

## THE RESULTS

As discussed above, two approaches to Bayesian analysis in regression are employed, Extreme Bounds Analysis (**EBA**) and Bayesian Analysis of Classical Estimation (**BACE**). For this paper, these approaches complement each other. **EBA** gives information on these ranges of the results, and **BACE** provides information on the distribution of the results within these ranges.

I present below the **EBA** and **BACE** analysis, but first to convey a sense of the models, I examine the two models with all the posited variables in vector **X** present. One model includes the **COL** variable, and the other does not. **Table 1** gives us some indication of the influence of the other variables on the **GDPC** models. In both models with and without **COL**, several of the other variables are statistically significant in the predicted direction. In addition, all of the variables representing the different years in the sample are also statistically significant. Furthermore, the model without the **COL** variable has an overall significance as indicated by its F value. The  $R^2$  is over 0.53 close to that of the model with the **COL** variable. Thus, variables other than **COL** do contribute to the model.

**Table 2** presents the Extreme Bounds Analysis for **COL**. This analysis gives one an idea of the range of the results that come from estimating a large number of models. As discussed above, my model finds the impact on the income as represented by **GDPC** of being a colonial metropole. It uses a large number of specifications. Given that the **X** vector contains eight variables, there are 256 model specifications counting the models with all the variables in **X** and with no variables in **X**. Due to the possibility of different specifications, the number of models is still larger. Both linear and non-linear (log-log) models are reasonable approximation of the situations faced by the sample countries. Furthermore, since simultaneity is possible, I use some instrumental variable equations in addition to the ordinary least squares (OLS) models. This approach may double the number of potential models.

However, the patterns of variable correlations limit the models that can be estimated. It turns out that there are 256 specifications each for the OLS linear model and the OLS log-log models. For the two instrumental variable (IV) models, only 64 specifications for the linear model and only 128 specifications for the log-log model can be estimated. Essentially certain matrices used in the Instrumental Variables estimation process are singular, and they cannot be inverted. This situation leads to a total of 704 specifications.

The analysis in **Table 2** gives one an idea of the range of the results that come out of estimating a large number of specifications. For each of the four models, this table displays the minimum and maximum impacts of **COL** on **GDP** along with the t values and the parameters from which these impacts are derived. For the linear model, the impacts are the actual regression coefficients for **COL**, while for the log-log specifications, the impacts are derived from those coefficients. The largest impact is \$2584; it come from the linear IV model; it is more than 78 percent of the average **GDP** for the sample, \$3297. The smallest impact is \$468 or 14.22 per cent of the total from the log-log OLS model.

Thus, for every specification, the impact of colonialization is positive. Strengthening this result is the range of the t values for the various models -- remember 704 in all. All the **COL** coefficients are statistically significant as well as positive. The results give strong support to the assertion that colonialization increased the income of the metropole countries.

Given their nature, the strong results raise two questions. First what, if anything, is the impact on a country of being a colony? This analysis can be answered by examining the impact of **COLE**, the dummy variable for being a colony. As with the **COL** variable, the **Extreme Bounds Analysis (EBA)** can contribute to the analysis. Second, what else can be said about the size of the **COL** impact or the **COLE** impact for that matter? To answer the first question, I apply **EBA** to the **COLE** variable. To examine the second issue, I do the Bayesian Analysis of Classical Estimates (**BACE**) for both the **COL** and **COLE** variables. From these exercises, I may be able to give an indication of the likely **COL** and **COLE** impact.

Here, I start with the **EBA** analysis of the **COLE** results. **Table 3** presents the **EBA** result for **COLE**. Since **COLE** is part of the **X** vector, it is not included in half of the models. (This is reasonable because it very well could be that **COLE** has no impact.) Using estimation models without **COLE** enlarges the number of models that can be utilized for determining the likely impact of **COL**. For the four specifications, the linear OLS, linear IV, log-log OLS, and log-log IV, the highest impact of **COLE** on **GDP** is \$828.4 (25.13 per cent of the sample average **GDP**). The minimum impact is -900.84 implying a negative impact on **GDP** of being a colony. For all four models, the minimum **COLE** impact is less than zero. For the two log-log models, this minimum impact is statistically significant. The log-log OLS model implies that that on the average being a colony lowers **GDP** by \$900.84 or 27.32 per cent of the average sample **GDP**. Thus, applying **EBA** to the **COLE** variable indicates that its impact could be either negative or positive.

The ambiguous result for **COLE**, being a colony merits some further comment. These results may reflect the varying nature of the sample colonies. Some colonies became places of European settlement where mother country population replaced the original native population. Some of these countries have become the richest in the world. Further, there were cases like Haiti where the overall wealth of a colonized country rose but the bulk of the population were enslaved. The latter derived little --if any --benefit. A third possibility is that the colonizers may have applied advanced technology to their colonies raising the average **GDP**, but with most of the wealth accruing to the metropole country.

Having established the bounds of the **COL** and **COLE** impacts, I now use the **BACE** analysis to convey a sense of the distribution of the colonialization impacts on the wealth of both colonial metropolises and the colonies themselves. **Table 4** conveys the key information on the **BACE** results. There are four general specifications: OLS linear, IV linear, OLS log-log, and IV log-log. For all models, the weighted average for **COL** impact is significantly greater than zero. The high t values mean that it is highly unlikely that the **COL** coefficient and thus the effect on the dependent variable, **GDP**, is less than zero. The implied impacts are derived from the weighted average of the **COL** coefficients for the 704 specifications. The weights are derived from the sum of the squared errors (SSE) for all these equations. (See Sala-i-Martin, et al, 2004, p. 815-819.) This weighting scheme gives a stronger influence to the better fitting equations. As shown in **Table 4**, the weighted percentage impact of **COL** on **GDP** ranges from 25 to 63 percent of the average **GDP**. For the linear model, the impact is the coefficient of **COL**, while for the log-log equations it is derived from this coefficient.

Applying the **BACE** analysis to the **COLE**, I find ambiguous result as I did with the **EBA** analysis. For two specifications, linear OLS and log-log OLS, **BACE** indicates that being a colony increased

wealth, but for the other two, linear IV and log-log IV, being a colony means a lower **GDPC**. For the two models with negative impacts, the results are not statistically significant. Furthermore, unlike the results for the **COL**, mother country impact, where for all models imply that positive increases are statistically significant, only 29.0 per cent of the models imply a statistically significant change in **COLE** positive or negative. Thus, in general, the results for **COLE**, the impacts of being a colony, are ambiguous; some specifications showing a positive impact and some showing a negative impact.

All in all, the analysis from both **EBA** and **BACE** indicates that a policy of colonization at least coincided with increased national income for the metropole countries. Consequently, being a metropole country seems to have increased national income. Thus, colonizing other lands could benefit given metropole countries. In contrast, the impact of being a colony on **GDPC** implied by this model is ambiguous. Some specifications imply that being a colony increased wealth while others indicate that it decreased wealth.

## CONCLUSION

The estimation of the **GDPC** models discussed above indicates that having colonies result in a quite large increase in national income, other things equal. Over the different specifications, the estimated mean increases range from 14 to over 78 per cent, other things equal. While these results are consistent with many theories on the institution of colonies, they are higher than one might expect.

Consequently, there may be something about the model that leads to biased results. One reasonable surmise is that there may be some omitted variables. The omission of relevant variables could bias the results.

Even with these problems, the analysis of this paper does include the impact of colonization on the colonies themselves, **COLE**. With this inclusion, therefore, the significant positive impact of being a colonial metropole country still remains. Thus, notwithstanding the arguments of Schumpeter and others, this paper indicates that the total income of the metropole countries rises as a result of colonization under a wide range of model specifications.

**TABLE 1**  
**SAMPLE REGRESSIONS FOR THE COLONIAL IMPACT MODEL**

Independent Variable	Coefficients for Model with COL	t Value	Coefficients for without COL	t Value
(Intercept)	4138.0		4392.0	
AREA	-0.00002	-0.302	0.00005	0.907
LATI	77.480	9.308 ***	72.440	8.586***
ELEV	0.0001	0.003	-0.0298	-0.755
LANDL	54.520	0.205	-203.10	-0.759
YR1600	-6691.0	-13.294***	-6900.0	-13.487***
YR1700	-6348.0	-13.380***	-6430.0	-13.303***
YR1820	-5574.0	-13.834***	-5589.0	-13.610***
YR1850	-6086.0	-13.350***	-6080.0	-13.086***
YR1870	-5025.0	-13.125***	-4998.0	-12.811***
YR1900	-4755.0	-11.349 ***	-4597.0	-10.788***
YR1913	-3925.0	-10.504***	-3776.0	-9.939***
YR1920	-4158.0	-10.605***	-4019.0	-10.077***
YR1925	-3828.0	-9.925***	-3698.0	-9.423***
YR1930	-3709.0	-9.825***	-3607.0	-9.386***
YR1935	-3678.0	-9.601***	-3565.0	-9.144***
YR1940	-3218.0	-8.231***	-3097.0	-7.783***
YR1950	-2545.0	-7.172***	-2505.0	-6.929***
YR1955	-1839.0	-5.173***	-1801.0	-4.973***
YR1960	-1178.0	-3.330***	-1157.0	-3.210 **
COL	1417.0	5.836 ***		
COLE	415.9	1.805	176.20	0.763
NEIGH1	18.740	0.079	5.9910	0.341
GENETIC	0.0270	0.410	0.04608	0.686
GENETIC2	0.0000008	0.352	0.00000008	0.034
EURA	-1739.0	-4.946***	-1411.0	-3.990***
EUROPE	1532.0	5.136***	1880.0	6.311***
TECHI	-0.2669	-3.333***	-0.27460	-3.366***
TECHI2	0.00002	4.571***	0.00002	4.174***
Adjusted R-squared:	0.5503		0.533	
F-statistic:	39.68		38.41	
p-value:	< 2.2e-16		<2.2e-16	

\* significant at the 95 per cent level.  
 \*\* significant at the 99 per cent level.  
 \*\*\* significant at the 99.9 per cent level.

**TABLE 2**  
**THE EXTREME BOUNDS ANALYSIS FOR THE IMPACT OF COLONIZATION ON THE**  
**GDPG OF COLONIAL METROPOLE COUNTRIES**

<b>Type of Specification and Estimation Technique*</b>		<b>Maximum Impact</b>	<b>Minimum Impact</b>	<b>Number of Models</b>
Linear OLS	Coefficient	1871.0	561.5	256
	t values	(7.7234)	(2.243)	
	Impact	1871.0	561.5	
	Per Cent of Total	56.75	17.03	
Linear IV	Coefficient	2584.0	1011.0	64
	t values	(7.083)	(4.051)	
	Impact	2584	1011.0	
	Per Cent of Total	78.38	30.66	
Log-Log OLS	Coefficient	0.5743	0.1534	256
	t values	(9.698)	(2.552)	
	Impact	1440.5	468.8	
	Per Cent of Total	43.69	14.22	
Log-Log IV	Coefficient	0.8412	0.2572	128
	t values	(10.23)	(3.003)	
	Impact	1875.3	747.77	
	Per Cent of Total	56.88	22.68	

\* The average GDPG for the sample is \$3297.

**TABLE 3**  
**THE EXTREME BOUNDS ANALYSIS FOR THE IMPACT OF COLONIZATION ON THE**  
**GDPG OF COUNTRIES THAT WERE COLONIZED**

Type of Specification and Estimation Technique*		Maximum Impact	Minimum Impact	Number of Models
Linear OLS	Coefficient	708.8	-381.6	128
	t values	(3.124)	(-1.609)	
	Impact	708.8	-381.6	
	Per Cent of Total	21.50	-11.57	
Linear IV	Coefficient	828.4	-152.6	32
	t values	(3.551)	(-0.629)	
	Impact	828.4	-152.6	
	Per Cent of Total	25.13	-4.63	
Log-Log OLS	Coefficient	0.0172	-0.242	128
	t values	(0.318)	(-3.958)	
	Impact	56.10	-900.84	
	Per Cent of Total	1.701	-27.32	
Log-Log IV	Coefficient	0.0364	-0.1984	64
	t values	(0.662)	(-3.179)	
	Impact	117.873	-723.46	
	Per Cent of Total	3.575	-21.943	

\* The average GDPG for the sample is \$3297.

**TABLE 4**  
**THE RESULTS OF THE BACE ANALYSIS FOR THE IMPACT OF COLONIZATION MODEL**  
**FOR THE PERIOD BETWEEN 1600 AND 1965**

Specification/ Estimation Method*	Variable	Number of Observations	Weighted Mean Coefficient	Standard Deviation	Mean Implied t Value	Weighted Average Percentage Impact	Percentage with Significant Sign
Linear OLS	COL	256	1416.13	242.82	5.832	42.95	100.00
Linear IV	COL	64	2088.59	366.42	5.700	63.35	100.00
Log-Log OLS	COL	256	0.2962	0.059	5.059	25.64	100.00
Log-Log IV	COL	128	0.4750	0.087	5.489	37.81	100.00
Linear OLS	COLE	128	415.66	230.34	1.805	12.61	21.88
Linear IV	COLE	32	667.48	236.20	2.826	20.50	78.13
Log-Log OLS	COLE	128	-0.069	0.056	-1.218	-7.11	29.69
Log-Log IV	COLE	64	-0.037	0.057	-0.638	-3.73	17.19

\* The average GDPC for the sample is \$3297.

## REFERENCES

- Ashraf, Q., & Galor, O. (2013). The “out of Africa” Hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review*, 103(1), 1-46.
- Auer, R. (2013). Geography, institutions, and the making of comparative development. *Journal of Economic Growth*, 18, 179–215.
- Bolt, J., & Van Zanden, J. L. (2014). The Maddison project: collaborative research on historical national accounts. *Economic History Review*, 67(3), 627–651.
- Chaudhry, A., & Garner, P. (2006). Political competition between countries and economic growth. *Review of Development Economics*, 10(4), 666–682.
- Crosby, A. W. (1986). *Ecological imperialism: the biological expansion of Europe, 900-1900*. Cambridge University Press.
- Diamond, J. M. (1997). *Guns, germs, and steel: the fates of human societies*. New York: W.W. Norton & Co.
- Gassebner, M., Gutmann, J., & Voigt, S. (2016). When to expect a coup d'état? an extreme bounds analysis of coup determinants. *Public Choice*, 169, 293–313.
- Granger, C. W. J., & Uhlig H. (1990). Reasonable extreme-bounds analysis. *Journal of Econometrics*, 44, 159-170.
- Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46, 931-59.
- Hoeting, J. A., Madigan, D., Raftery A. E., & Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical Science*, 14(4), 382-401.
- Johnson, N., & Koyama, M. (2017). States and economic growth: capacity and constraints. *Explorations in Economic History*, 64, 1-20.



- Jones, E. L. (2007). *Growth recurring: economic change in world history*. Ann Arbor: University of Michigan Press.
- Karayalcin, C. (2008). Divided we stand, united we fall: The Hume-North-Jones mechanism for the rise of Europe. *International Economic Review*, 49(3), 973-997.
- Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, 82(4), 942-963.
- Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.
- Maddison, A. (2001). *The world economy: a millennial perspective*. Paris: Development Centre of the Organization for Economic Cooperation and Development.
- Madigan, D., & Raftery, A. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, 89(428), 1535-1546.
- McGuire, M., & Olson, M. (1996). The economics of autocracy and majority rule: the invisible hand and the use of force. *Journal of Economic Literature*, 34(1), 72-96.
- Owen, A. L., Videras, J., & Davis, L. (2009). Do all countries follow the same growth process? *Journal of Economic Growth*, 14, 265–286.
- Ozak, O. (2018). Distance to the pre-industrial technological frontier and economic development. *Journal of Economic Growth*, 23, 175–221.
- Pomeranz, K. (2000). *The great divergence: China, Europe, and the making of the modern world economy*. Princeton, N.J.: Princeton University Press.
- Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437), 179-191.
- Ramon Moreno, R., & Trehan, B. (1997). Location and the growth of nations. *Journal of Economic Growth*, 2, 399–418.
- Sala-i-Martin, X. (1997). I just ran two million regressions. *The American Economic Review*, 87(2). Papers and Proceedings, 178-183.
- Sala-i-Martin, X., Doppelhofer, G. & R. Miller (2004). Determinants of long-term growth: a Bayesian averaging of classical estimates (EBA) approach. *The American Economic Review*, 94(4), 813-835.
- Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3), 312-320.
- Van Zanden, J. L. (2009). *The long road to the industrial revolution: the European economy in a global perspective, 1000-1800*. Leiden: Brill.
- Wright, H. M. (1961). *The new imperialism: analysis of late nineteenth-century expansion*. Boston: D. C. Heath.