Divides in Development: An Analysis on the Impact of Modern Western Colonialism on the Economies of Former Colonies

Faria Khandaker

04/15/2020

Abstract

The effects of Modern Western Colonialism are still visible in today's global political landscape. In this replication experiment, settler mortality is used as an instrumental variable to investigate the relationship between protection against expropriation and Gross Domestic Product per capita of former colonies. Former colonies with strong protection against expropriation laws tend to enjoy higher GDP per capita today. A high GDP contributes to the financial stability of a country and allows its citizens to enjoy a higher quality of life. Finding out the root causes of financial instability of a country can aid in the development of strategies that lead to more prosperous future.

Introduction

In 1492, Christopher Colombus with a crew of 90 men set sail from Spain on a mission to discover a sea route that would reach the spice ladened lands of India and China. He had no way of knowing that two giant landmasses stood in the way of his ambitions. His discovery of these new lands ignited the period of history dominated by the phenomenon of Western colonialism which saw 300 years of the slave trade, the decimation of entire Native populations through war and disease, and the emergence of European super powers whose wealth was attributed to amassing large amounts of precious metals and raw resources from their colonies.

The effects of colonization are deeply embedded into today's global political landscape with many countries still coping with the aftermath of having been governed by foreign powers, often with little to no understanding of local cultures, for hundreds of years. Of course, the impact of colonization is not felt equally across all former Colonies. This report aims to investigate one of the proposed explanations for why this is the case. In the paper by Daron Acemoglu, Simon Johnson, and James A Robinson titled "The Colonial Origins of Comparative Development: An Empirical Investigation", it is hypothesized that different strategies used by colonial powers impacted how the colonies were governed which in turn impacted the kinds of institutions that were set up early in the history of the colony. These early institutions then evolved into the institutions currently present in those countries today. They hypothesized that the rate of mortality for European settlers was key to the strategy of governance colonial powers chose for certain colonies. Areas where disease was prevalent and the environment was not ideal for settling down and raising a family, turned into extractive colonies whose main function was to gather raw materials for trade. Not many institutions needed to be established in colonies like these and the institutions that were established reinforced human exploitation as there was no need to invest in infrastructure or in the workers.

On the other hand, there were colonies termed in the paper as "Neo Europe" whose environments were relatively disease free and safe for settling, encouraging settlers to establish institutions similar to those that existed in their motherlands. Using the Settler Mortality rate as an instrumental variable, two stage least squares (2SLS) estimates were calculated to investigate the relationship between Expropriation Protection and GDP per capita. After running multiple 2SLS estimates using data from different colonies across three continents, it was found that the higher the protection against expropriation, the higher the GDP of a country appears to be.

The Dataset

The dataset was made available by Daron Acemoglu and his colleagues at MIT. The link is provided in the Reference section of this report. The article contains eight tables which provide different summary statistics for a diverse range of variables. The datasets are divided according to the tables they were used for. For this report, columns from datasets 4,5 and 6 were used for exploratory data analysis and for a deeper understanding of the data. The population of interest include 64 former colonies across North and South America, Africa, Asia and Australia. In total, eleven variables from three datasets were used to produce the master table on which two stage least square regression analysis was performed. An addition of two dummy variable columns were created to separate Latin American and Caribbean colonies and to differentiate colonies controlled by the British, French and other powers. Columns containing dummy variables for each continent allowed for exclusive analysis of specific colonies, which is important for the highlighting of outliers and exceptions that may be present.

The three key variables discussed in this report are Gross Domestic Product per capita (GDP), Expropriation Protection and Settler Mortality. GDP is an economic measure of wealth of a country; it is the total value of all goods and services produced in a country in given period of time. Expropriation as term refers to the acquiring of private property by a government against the wishes of the property owner. Therefore, expropriation protection as a measurement is used to gauge how accepting a country is to private and foreign investment. Settler mortality is the measure of deaths per 1000 people that occurred in a colony between the 1600s and the 1800s. Settlers in this case include soldiers, bishops and sailors.

Discussion and Analysis

The Instrumental Variables (IV) approach allows us to discuss causality between variables when randomization is difficult to do. It is useful when we have treatment and control groups, but the outcome of the experiment can be muddled due to variation between groups that are hard to control. In their paper, the authors wanted an answer for the question, "What is the effect of institutions on economic performance?". They considered the variables of Expropriation Protection and Gross Domestic Product per capita of a country to answer this question. Why is expropriation protection used? The literature suggests that founders of countries with strong institutions had capitalist inclinations which strongly advocated for private ownership of property. To improve their standards of living, countries like these tend to invest in infrastructure and human capital which aids in increasing a nation's wealth. A simple linear regression on the two variables would not produce accurate results as different countries have different factors affecting why they may or may not invest as much as others.

This is where, for the purposes of this specific research question, European settler mortality as an instrumental variable is used to explain the relationship between GDP and Expropriation Protection. The exclusion restriction for Settler mortality is that the mortality rates of settlers from more than 100 years ago should not impact the GDP of a country today. The hypothesis is that European Settlers established settlements in colonies which then led to the establishment of institutions to govern people and those past institutions developed over time to what exists today.

To test the strength of settler mortality as a good instrumental variable and to ensure that it is not being influenced by any observable variable, the authors calculate robust 2SLS estimate on a number variables that could be indirectly related to settler mortality and economic outcomes. They include latitude, temperature, soil quality, life expectancy, general health of the local population, etc. None of the results showed any significant impact on the Expropriation risk of the country or on settler mortality. The scope of my report, however, is only to replicate and add to the analysis for table number four of the Acemoglu paper which summarizes the two stage least squares estimates for Expropriation risk and Settler Mortality in relation to the GDP per capital of former colonies. Table 1 contains the summary statistics for this replication experiment.

Models

There are two models that can help us understand the relationships between the variables used for analysis. The first model explains the relationship between GDP per capita and Protection against Expropriation Risk. The second model explains the relationship between Protection against Expropriation risk and settler mortality.

Model 1: Linear Regression of Income per Capita and Expropriation Protection

$$log(y_i) = \mu + \alpha R_i + X_i(\gamma) + \epsilon_i$$

In the model above, the variable y_i represents income per capita, R_i represents Protection against Expropriation risk, X_i adjusts for variation from other variables and E_i is an error term. The effect of the alpha term is the variable of interest as it represents the effect of institutions on GDP per capita. This model represents the relationship between having strong institutions in place and the GDP of a country. If all other variables remain constant, a high R_i score (after taking the effects of alpha into consideration) will lead to a high Y_i score; the stronger the institutions are of a country, the higher their GDP will be.

Model 2: Linear Regression of Expropriation Protection and Settler Mortality

$$R_i = \zeta + \beta(log)M_i + X_i(\delta) + \upsilon_i$$

In the model above, R_i is the same as Model 1, the Protection against Expropriation Risk measure, M_i is the settler mortality rate per 1000 people. Settler mortality is not a variable used in Model 1 and therefore is the instrumental variable whose effect on expropriation risk will be tied back to GDP per capita. The assumption here is that settler mortality rate during colonial times is completely exogenous and unrelated to the current GDP per capita of a country that was a former colony.

Table 1 displays the effects of Expropriation on GDP across a variety of colonized regions worldwide by use of Settler Mortality as an Instrumental Variable. GDP per Capita is the dependent variable being analyzed. Table 1 is divided into 3 rows and should be read as follows: The first row contains the results of the 2SLS method, the second row contains the results of the first stage of the 2SLS method, and the third row contains the results of the ordinary least squares method. We ignore the intercept when interpreting the table and look at the effects next to the variables. The numbers in parenthesis are the standard errors. The five columns in the table represent the exclusion of different colonies. The columns should be interpreted as follows: Excolonies contain the entire sample of 64 colonies used for this study. Non_Neu excludes the 4 colonies who are considered "Neo European" or colonies whose institutions strongly resemble those found in Europe. Non_African and Non_Asian exclude African and Asian colonies from the results. Non_LAC exclude colonies from Latin America and the Caribbean.

The OLS regression scores demonstrate a strong positive connection between GDP and Expropriation Protection, the higher the expropriation protection, the higher the GDP of a country (See Figure 1). There is a strong negative relationship across all colonies between Settler Mortality and Expropriation protection; the higher the settle mortality, the lower the expropriation protection for the region (See Figure 3). The relationship between GDP and Settler Mortality is also negative (See figure 2). Figures 1, 2 and 3 clearly depict that all colonial powers owned colonies across the spectrum; no one colonial power had less settler deaths compared to others. This is to say that settler mortality is not impacted by the identity of the colonizer; thus adding to its strength as an instrumental variable.

The five former British colonies clustered around the top right hand corner of the plot (displaying high protection and high GDP) are the countries of Canada, America, Australia, New Zealand and Singapore. These colonies serve as outliers and mask the effect of settler mortality to some extent. The Non_Neu column of Table 1 (which excludes the outliers) shows the highest effect of Expropriation Protection on GDP. Figures 4, 5 and 6 located in the Appendix display the scatterplots after the removal of the outliers. The slope of the line is lower when the Neu is excluded which indicates that the outlier do indeed cause an upward bias of the results.

Table 1: Statistics of IV Regression of GDP per Capita by Regions

	Excolonies	Non Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	1.910 *	-0.141	4.554 ***	1.600	2.218 *
	(0.823)	(1.382)	(0.828)	(0.881)	(0.894)
Expropriation_Protection	0.944 ***	1.281 ***	0.578 ***	1.006 ***	0.850 ***
	(0.126)	(0.219)	(0.117)	(0.137)	(0.135)
N	64	60	37	55	40
R2	0.477	0.372	0.409	0.505	0.511
logLik	-72.268	-67.494	-38.703	-59.371	-47.851
AIC	150.536	140.988	83.406	124.743	101.702

*** p < 0.001; ** p < 0.01; * p < 0.05.

	Excolonies	Non_Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	9.341 ***	8.184 ***	11.844 ***	9.140 ***	9.140 ***
	(0.611)	(0.657)	(0.898)	(0.687)	(0.687)
settler_mortality	-0.607 ***	-0.391 **	-1.210 ***	-0.572 ***	-0.572 ***
	(0.127)	(0.133)	(0.219)	(0.139)	(0.139)
N	64	60	37	55	55
R2	0.270	0.130	0.466	0.243	0.243
logLik	-104.829	-93.976	-54.697	-90.923	-90.923
AIC	215.659	193.952	115.393	187.847	187.847

*** p < 0.001; ** p < 0.01; * p < 0.05.

	Excolonies	Non_Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	4.660 ***	4.866 ***	5.221 ***	4.751 ***	4.194 ***
	(0.409)	(0.489)	(0.464)	(0.425)	(0.462)
expro_pro	0.522 ***	0.487 ***	0.482 ***	0.514 ***	0.549 ***
	(0.061)	(0.076)	(0.065)	(0.065)	(0.068)
N	64	60	37	55	40
R2	0.540	0.414	0.611	0.543	0.629
logLik	-68.168	-65.419	-30.945	-57.171	-42.337
AIC	142.335	136.838	67.890	120.342	90.674

*** p < 0.001; ** p < 0.01; * p < 0.05.

Figure 1: Expropriation Protection VS GDP

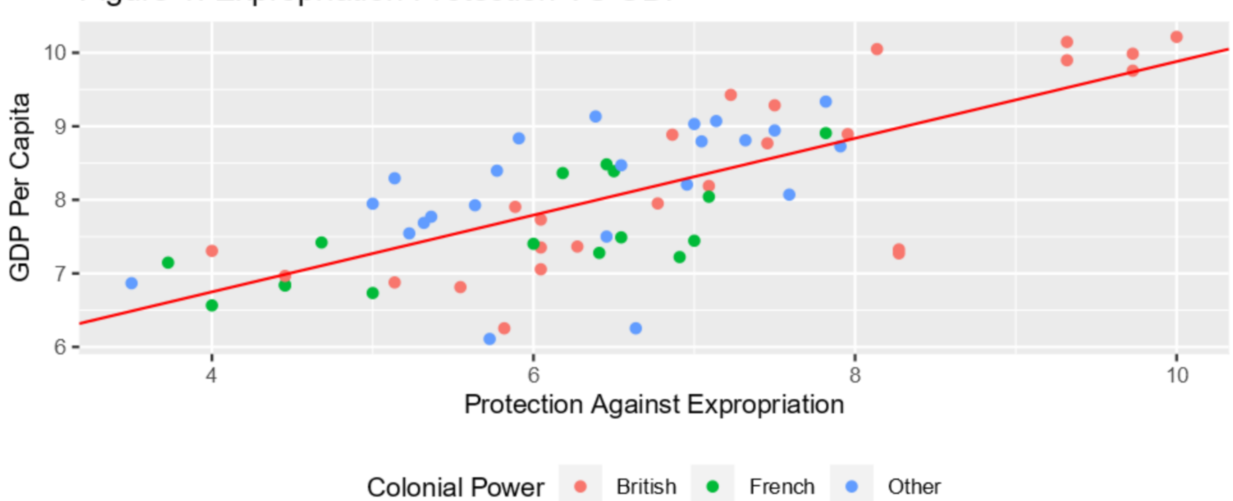


Figure 1 displays that colonies with higher expropriation protection have a higher GDP per Capita.

Figure 2: Relationship between settler mortality and GDP

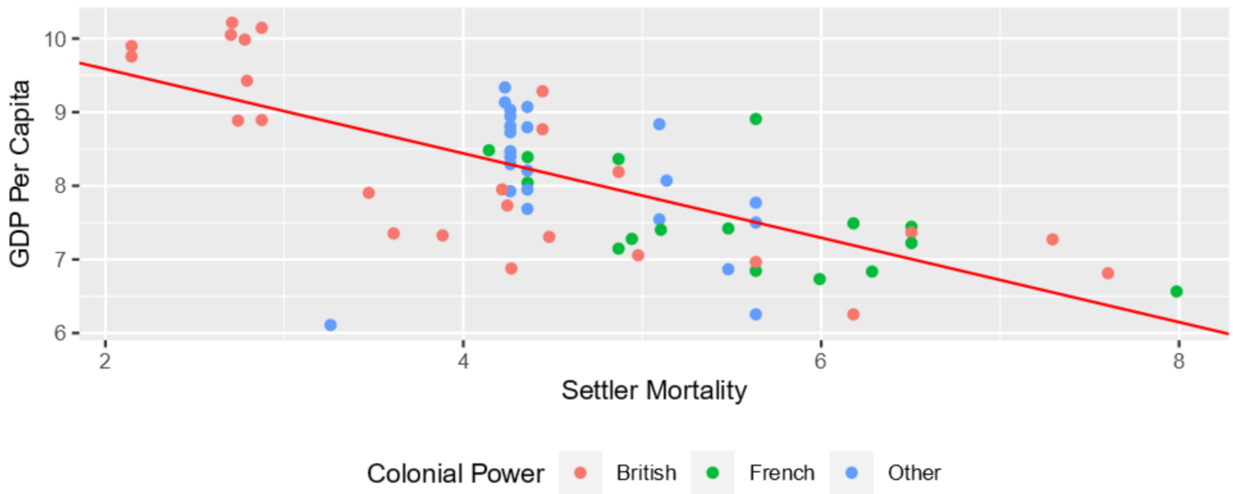


Figure 2 displays that colonies with higher settler mortality(in the past) tend to be less wealthy

Figure 3: Relationship between settler mortality and Expropriation Protection

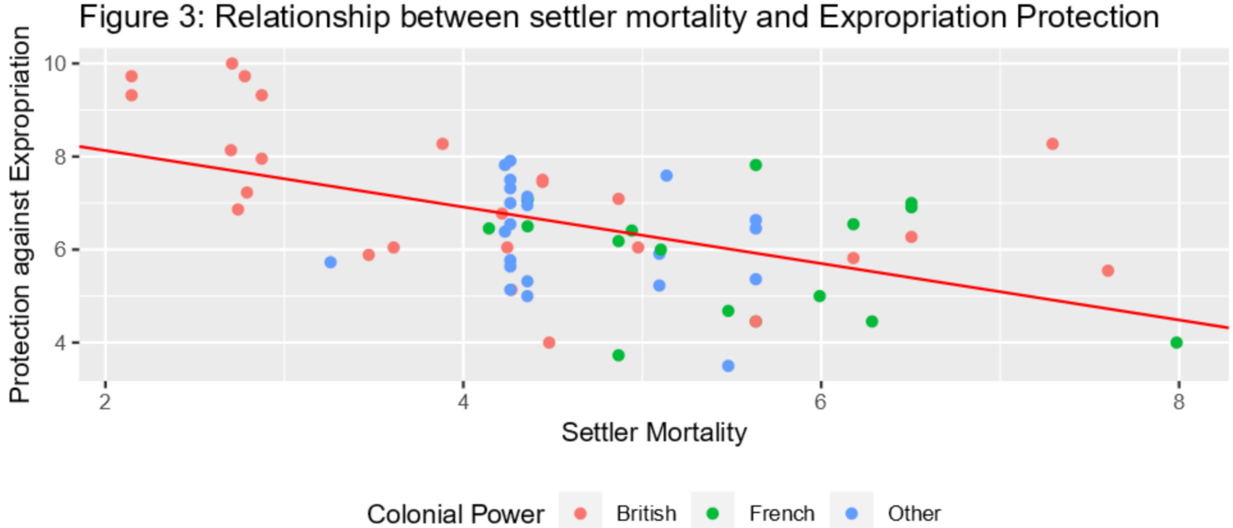


Figure 3 displays that colonies with high settler mortality rates(in the past) have lower expropriation protection (weaker institutions to protect private interests)

It should be noted that the columns for Non_Asian and Non_LAC are not included in the original paper which is a curious matter as expropriation values in these columns show very strong and significant effects on GDP. Furthermore, the original paper does not present data for specific individual continents. By assigning a few continent dummies of my own, I was able to separate colonies according to continent and region (as in the case of Latin America and the Caribbean). Table 4 in the appendix displays the effects of Expropriation and Settler mortality on GDP but for Neo European Colonies and colonies in Africa, Asia and Latin America and the Caribbean, individually. The effects of Expropriation protection on GDP are high for all four regions, however, the effect is only significant for African and Asian colonies (p-values below 0.05 for Neu and Latin/Carib columns). In fact, settler mortality is positively correlated for Neo European Colonies! (See Table 4 and Figure 10 in the appendix) This may be indicative of differences in colonization strategies used by the colonial powers. It could be that for certain colonies, the intention from the beginning was to settle and create a new country, not increase the wealth of the colonizing country. Perhaps positively correlated settler deaths are the result of battles and wars that took place to displace the original inhabitants of the land in order to establish a Neo-European colony; patriots willing to die for their land. After all, Acemoglu's dataset on settler mortality wasn't filtered according the causes of settler deaths. Another interesting finding is that the OLS regression for GDP and Expropriation protection is positive for Neu colonies but isn't statistically significant. It could be that the sample size is allowing for the effects of other variables, not considered by the authors, to show and produce this result. Further investigation of these findings is needed.

The results for the Latin American and Caribbean colonies are also quite surprising. Unlike the Neo European colonies, settler mortality shares a negative relationship with GDP and (similar to all colonies) higher Expropriation Protection shows higher GDP. But none of these values are significant in the statistical sense (p-values below 0.05). This can indicate the influence of other confounding variables on the results. Unlike Asia, Africa, who were colonized mainly by the British and French, Latin America and the Caribbean were mostly colonized by other Imperial Powers like Spain and Portugal (See figure 7 in the appendix). Latin American and the Caribbean were colonized with the intention of forming extractive colonies by their colonizers. It is quite possible that expropriation protection is not a significant measure of GDP for these colonies because for many countries, the exploitative and absolutist institutions established during colonial times were kept after independence. The wealth of these countries may have increased through trade and political monopolies, without the need to improve institutions. The differences in the culture and colony management strategies appear to be quite influential on the results as they can potentially introduce variables that are not apparent to us. These discrepancies are difficult to explain but it does encourage further analysis and investigation and highlights the importance of experimental replication in Data Science.

Shortcomings and Limitations

A limitation of the Instrumental Variables method is that the appropriate variables can be hard to encounter. It is difficult to find a variable that is related to the treatment group but is not a confounding aspect of the outcome.

While the tests and analyses done on various variables were very rigorous, the dataset used in the paper has potential biases within it. The authors were limited by the information available on settler mortality in creating their testing dataset. Not all former colonies are used in this experiment; much of the Middle East came under British rule but those none of those countries are a part of the test dataset. The sample size for Asia is nine colonies compared to 24 for Latin America and the Caribbean and 27 for Africa. Having merely nine colonies represent the vastness of Asia surely produces biased results.

The authors of the original article state that the general use of the term "institutions" is a shortcoming of their research as there is nothing specifying which institutions can reduce the risk of expropriation or even how to establish them. There are many aspects to consider for the establishment of a "strong institution"; Which political system is better, what does a strong legal system looks like, socialism or capitalism, etc. In addition to investigating other variables that may influence the relationship between expropriation protection and GDP, future studies should point out patterns in the structures of institutions of high GDP countries.

A major limitation of this paper is that the prevalence of disease is classified as the major factor between an extractive colony and a Neo European colony. It doesn't take into consideration the role that technological advancement played in the colonizing of nations. In fact it was after the first industrial revolution in Europe that colonial powers had been able to travel from the exterior to the interior of Africa. There may be a correlation between advancements in locomotive technology and a decreasing settler mortality rate that was not taken into consideration by this paper. The paper makes mention of advancements in medicine that helped reduce death rates in the latter half of the 19th century but does not explain further that the decrease in death rates didn't necessarily make those colonies an option for settling down. There are other confounding variables at play that should be investigated in future studies. The amount of time a country and its people had to heal from colonial exploitation can also create noise in the data and the results. "The Scramble for Africa" occurred during the later stages of Colonialism (Modern Democratic Republic of the Congo was a Belgian Colony from 1885 to 1960). Countries that became independent in more recent history face political and economical instability. For those former colonies, settler mortality would not be a strong measure of institutional strength. Another limitation of this paper is that the establishment of arbitrary borders by colonial powers is not taken into consideration. As mentioned previously, the colonial powers established colonies for their own personal gains, regardless of whether the colony was extractive or Neo-European. They did not consider the social tensions between natives of the land. In some cases, they created tensions between cohesive groups by favouring one group over the other (for example, the Hutus and Tutsis of Rwanda). Establishing borders where people of a common culture are separated, or where people of different cultures are forced to live together will create opportunities for conflict no matter how many well-established institutions may exist in a country. The culture of a country may not impact GDP (as was tested and discussed in the paper), but cultural tensions certainly can. While is this is understandably difficult to quantify, it is still a factor to be considered when interpreting and presenting findings on groups of people.

Ethics

This article does not mention any ethical considerations made when analyzing this dataset. Acemoglu and colleagues do a fantastic job of highlighting many external variables that may have an impact on their results. However, they don't discuss those variables outside the scope of their results; there is no explanation on the ways in which the variables may have impacted the populations of those colonies in a greater context. Their study on the impacts of colonialism on former colonies, while very insightful, simplifies centuries of exploitation down to a few numbers. Factors like whether a colony was a French or British colony did not impact settler mortality rates, however the colonies that are grouped as Neo-European are composed of former British Colonies. Without greater discussion of colonial history, one may be inclined to compare certain cultures and traditions as better or more successful than others. This was a key reason behind many of the colonial conquests; that the locals were thought to be lower than the Europeans settlers. An academic paper discussing the effects of colonialism should have addressed to some extent.

Another thing to consider is how well GDP can represent economic performance. Are majority of the citizens benefiting from the wealth that is being created? What industries are fueling this growth? How can a rich country have millions of people go to sleep hungry? Then what is really the difference between a rich and a poor country? Numbers and statistics make it easy to establish a dichotomy; to label something as good or bad, right or wrong, strong or weak, but there is a lot of grey area that gets overlooked by error values. Failing to consider the bigger picture can lead to the rise of issues around discrimination and bias reinforcement that is already present in society.

The sources of information used by this paper introduces bias into the dataset as the authors (with the exception of the South American Settler Mortality dataset) of these sources are white males. Instrumental Variables aid researchers when randomization isn't possible but if the variety of variables available in the dataset are already biased, how accurate can the results actually be? However, this is not to say that the authors did not understand the gravity of their claims. One of the key issues in data science is transparency and reproducibility and this is addressed by the fact that the code and dataset are freely available on internet for anyone to test and probe into. This shows that the authors are open to discussions about their findings and do not have personal gains to make from their results. Scientists who use data on humans for their research should always look to connect their finding to the bigger picture in order to limit the imagination and the inherent bias we are all born with.

Conclusion

The use of Instrumental Variables on treatment and control groups and the interpretation of the results using the two stage least squares method can speak to the causality of two variables, granted that the instrumental variable itself does not influence the outcome of the treatment. In this report, the relationship between GDP and Expropriation Protection was investigated through the use of settler mortality rates across 64 colonies worldwide. The exclusion restriction for settler mortality is that settler deaths from over 100 years ago are not affecting the current GDP of a country. Colonies with high settler mortality were developed into extractive states whose main purpose was to produce raw materials for trade. Colonies with low settler mortality were developed into countries modelling the European way of life with European institutions.

Colonies that became extractive states had higher risk of expropriation leading to a lower Gross Domestic Product per capita; They are the poorer countries in the world. Countries with lower settler mortality rates enjoyed higher GDP as expropriation risk was low for those countries. Expropriation Protection is correlated with having strong "institutions" by which a country is governed. This allows for more private and foreign investments which turns the wheels of the economy and allows for an increase in GDP. The positive relationship between Expropriation protection and GDP shows how settler mortality, as an instrumental variable, meets the criteria of relevance. A country's colonial history has a big impact on their current state. It is difficult to quantify such effects for many reasons: the length of time that has passed, the data needed may not be available or was not collected, too many confounding variables (because of the complex nature of human beings), etc. It is, however, important to ask these questions and investigate such matters as the lessons learned can change the direction of policies enacted by certain institutions, in the future.

References

libraries used

David Hugh-Jones (2020). huxtable: Easily Create and Style Tables for LaTeX, HTML and Other Formats. R package version 4.7.1. https://CRAN.R-project.org/package=huxtable

Graeme Blair, Jasper Cooper, Alexander Coppock, Macartan Humphreys and Luke Sonnet (2020). estimatr: Fast Estimators for Design-Based Inference. R package version 0.22.0. https://CRAN.R-project.org/package=estimatr

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020). dplyr: A Grammar of Data Manipulation. R package version 0.8.4. https://CRAN.R-project.org/package=dplyr

Hadley Wickham and Evan Miller (2019). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.2.0. https://CRAN.R-project.org/package=haven

Sam Firke (2020). janitor: Simple Tools for Examining and Cleaning Dirty Data. R package version 1.2.1. https://CRAN.R-project.org/package=janitor

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Datasets used can be found here

http://economics.mit.edu/faculty/acemoglu/data/ajr2001

Code for data cleaning,graphs and tables can be found here

https://github.com/fariiaakh/Experimental-Design-in-Data-Science/blob/master/problemset5.Rmd

Research sources

Encyclopedia Britannica. (2019, January 23). History of Latin America. Retrieved From https://www.britannica.com/place/Latin-America/The-independence-of-Latin-America

Gertler, Paul, Sebastian Martinez, Patrick Premand, Laura Rawlings, and Christel Vermeersch. (2016). Instrumental Variables. (2nd.Ed), Impact Evaluation in Practice (pp. 89-110). World Bank Group.

Historic UK. (2019). Timeline of the Industrial Revolution. Retrieved From https://www.historic-uk.com/HistoryUK/HistoryofBritain/Timeline-Industrial-Revolution/

Martin, D. L., & Goodman, A. H. (2002). Health conditions before Columbus: paleopathology of native North Americans. The Western journal of medicine, 176(1), 65–68. https://doi.org/10.1136/ewjm.176.1.65

Telling Stories with Data.(2020 March). Chapter 35 Instrumental variables. Retrieved From. https://www.tellingstorieswithdata.com/instrumental-variables.html

Towards Data Science. (2018, May 10). Benefits and Ethical Challenges in Data Science — COMPAS and Smart Meters. Retrieved From https://towardsdatascience.com/benefits-and-ethical-challenges-in-data-science-compas-and-smart-meters-da549dacd7cd

VOX. (2014, February 24). Map: European colonialism conquered every country in the world but these five. Retrieved From https://www.vox.com/2014/6/24/5835320/map-in-the-whole-world-only-these-five-countries-escaped-european

Wikipedia. (2020, April 14). Belgian Congo. Retrieved From https://en.wikipedia.org/wiki/Belgian_Congo

Wikipedia. (2020, April 14). Rwandan genocide. Retrieved From https://en.wikipedia.org/wiki/Rwandan_genocide

Youtube.(2012, September 12). Imperialism: Crash Course World History #35. Retrieved From https://www.youtube.com/watch?v=alJaltUmrGo

Appendix

Code and extra tables and figures are here

creating specific data tables from exisiting datasets

```
countries<-tibble(country=data4$shortnam,
                  dummy_africa=data4$africa,
                  dummy_asia=data4$asia,
                  dummy_neu=data4$rich4,
                  britcol=data5$f_brit,
                  frencol=data5$f_french,
                  sample=data4$baseco
                  )


excolonies <- tibble(country=data4$shortnam,
                     samples=data4$baseco,
                     dummy_africa=data4$africa,
                     dummy_asia=data4$asia,
                     dummy_neu=data4$rich4,
                     britcol=data5$f_brit,
                     frencol=data5$f_french,
                     landlocked=data6$landlock,
                     settler_mortality=data4$logem4,
                     expro_pro=data4$avexpr,
                     gdppc=data4$logpgp95

                     )%>% subset(samples==1)



excolonies<-excolonies %>% mutate(counttrycol= case_when(
                                  britcol==1~"British",
                                  frencol==1~"French",
                                  britcol==0~"Other",
                                  frencol==0~"Other"))


#latin american / caribbean countries

excolonies<-
  excolonies %>%
  mutate(nonasaf= dummy_africa+dummy_asia+dummy_neu,
```

```r
        nonasaf=if_else(nonasaf==0,1,0))

#only latin american and caribbean
excolnocon <- excolonies %>% subset(dummy_africa==0)%>%
  subset(dummy_asia==0)%>%
  subset(dummy_neu==0)

#excolonies is master table with everything

#excolonies2 master table without dummy_neu, to take away outliers
excolonies2<-excolonies[!(excolonies$dummy_neu==1),]


#only africa
african<-
excolonies %>%
  filter(dummy_africa == "1")

# all colonies without africa
non_african <-
excolonies %>%
  filter(dummy_africa == "0")

#only asia
asian<-excolonies %>%
  filter(dummy_asia == "1")
#all colonies without asia

non_asian<-excolonies %>%
  filter(dummy_asia == "0")
#all colonies without latin america/caribbean
nonlacar<- excolonies%>%
  filter(nonasaf=="0")

#only neo-europe
neu <- excolonies %>%
  filter(dummy_neu == "1")
```

Excolonies contains all colonies including first world countries

```r
exmortalgdp<- lm(gdppc~settler_mortality, data=excolonies)
exmortalexpro<-lm(expro_pro~settler_mortality, data=excolonies)
coef(exmortalgdp)["settler_mortality"]/ coef(exmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##         0.9442794
```

```r
first_stage <- lm(expro_pro~settler_mortality, data=excolonies)
Expropriation_Protection <- first_stage$fitted.values
second_stage <- lm(gdppc~Expropriation_Protection, data=excolonies)
```

Excolonies2 is includes all excolonies except the neo european first world countries

```
ex2mortalgdp <- lm(gdppc~settler_mortality, data=excolonies2)

ex2mortalexpro <- lm(expro_pro~settler_mortality, data=excolonies2)
```

```
coef(ex2mortalgdp)["settler_mortality"]/ coef(ex2mortalexpro)["settler_mortality"]
```

```
## settler_mortality
##           1.28124
```

```
first_stage2<-ex2mortalexpro
Expropriation_Protection<-first_stage2$fitted.values
second_stage2<-lm(gdppc~Expropriation_Protection,data=excolonies2)
```

Lacar are the colonies from latin america and the caribbean

```
# relation b/w settler mortality,gdp and expropriation for latin american and caribbean countries

lacmortalgdp <- lm(gdppc~settler_mortality, data=lacar)
lacmortalexpro <- lm(expro_pro~settler_mortality, data=lacar)

coef(lacmortalgdp)["settler_mortality"]/coef(lacmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##         0.7365772
```

```
first_stage1<-lacmortalexpro
Expropriation_Protection<-first_stage1$fitted.values
second_stage1<-lm(gdppc~Expropriation_Protection,data=lacar)
```

Non Latin American and Caribbean colony 2SLS

```
nonlacmortalgdp <- lm(gdppc~settler_mortality, data=nonlacar)
```

```
nonlacmortalexpro <- lm(expro_pro~settler_mortality, data=nonlacar)
```

```
coef(nonlacmortalgdp)["settler_mortality"]/ coef(nonlacmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##         0.8501676
```

```
first_stage7<-nonlacmortalexpro
Expropriation_Protection<-first_stage7$fitted.values
second_stage7<-lm(gdppc~Expropriation_Protection,data=nonlacar)
```

african contains only colonies in Africa 2SLS Expropriation

```
afmortalgdp <- lm(gdppc~settler_mortality, data=african)
```

```
afmortalexpro <- lm(expro_pro~settler_mortality, data=african)
```

```
coef(afmortalgdp)["settler_mortality"]/ coef(afmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##          2.400495
```

```
#your gdp will increase by 2.4 points if you have strong expropriation laws

first_stage3<-afmortalexpro
Expropriation_Protection<-first_stage3$fitted.values
second_stage3<-lm(gdppc~Expropriation_Protection,data=african)
```

non_african contains all colonies outside Africa 2SLS

```
nonafmortalgdp <- lm(gdppc~settler_mortality, data=non_african)
```

```
nonafmortalexpro <- lm(expro_pro~settler_mortality, data=non_african)
```

```
coef(nonafmortalgdp)["settler_mortality"]/ coef(nonafmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##         0.5779968
```

```
first_stage4<-nonafmortalexpro
Expropriation_Protection<-first_stage4$fitted.values
second_stage4<-lm(gdppc~Expropriation_Protection,data=non_african)
```

Asian contains only colonies within Asia 2SLS

```
asmortalgdp <- lm(gdppc~settler_mortality, data=asian)
```

```
asmortalexpro <- lm(expro_pro~settler_mortality, data=asian)
```

```
coef(asmortalgdp)["settler_mortality"]/ coef(asmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##          1.227493
```

```
first_stage5<-asmortalexpro
Expropriation_Protection<-first_stage5$fitted.values
second_stage5<-lm(gdppc~Expropriation_Protection,data=asian)
```

Non__Asian contains all colonies outside of Asia 2SLS

```
nonasmortalgdp <- lm(gdppc~settler_mortality, data=non_asian)
```

```
nonasmortalexpro <- lm(expro_pro~settler_mortality, data=non_asian)
```

```
coef(nonasmortalgdp)["settler_mortality"]/ coef(nonasmortalexpro)["settler_mortality"]
```

```
## settler_mortality
##          1.006109
```

```
first_stage6<-nonasmortalexpro
Expropriation_Protection<-first_stage6$fitted.values
second_stage6<-lm(gdppc~Expropriation_Protection,data=non_asian)
```

Neu contains only the 4 Neo European colonies 2SLS

```
neumortalgdp <- lm(gdppc~settler_mortality, data=neu)
```

```
neumortalexpro <- lm(expro_pro~settler_mortality, data=neu)
```

```
coef(neumortalgdp)["settler_mortality"]/ coef(neumortalexpro)["settler_mortality"]
```

```
## settler_mortality
##         0.8020011
```

```
first_stage8<-neumortalexpro
Expropriation_Protection<-first_stage8$fitted.values
second_stage8<-lm(gdppc~Expropriation_Protection,data=neu)
```

OLS regression of all colonies

```
#excolonies
exprogdp<-lm(gdppc~expro_pro, data=excolonies)

#excolonies2
exprogdp2<-lm(gdppc~expro_pro, data=excolonies2)

#non_african
exprogdpnonaf<-lm(gdppc~expro_pro, data=non_african)
```

```r
#Neo European
exprogdpneu<-lm(gdppc~expro_pro, data=neu)
#African
exprogdpaf<-lm(gdppc~expro_pro, data=african)
#Asian
exprogdpas<-lm(gdppc~expro_pro, data=asian)
#Non Asian
exprogdpnonas<-lm(gdppc~expro_pro, data=non_asian)
#Latin American and Caribbean
exprogdplacar<-lm(gdppc~expro_pro, data=lacar)
#Non Latin American and Caribbean
exprogdpnonlacar<-lm(gdppc~expro_pro, data=nonlacar)
```

Robust IV Regression of Colonies

Table 2 contains the IV robust statistics of the relationship between GDP and Expropriation via Settler Mortality. For technical reasons, the estimate of effect was not showing for the "Excolonies" column. The value 0.944 under table 3 are is the effect of expropriation on GDP estimate for Excolonies.

```r
#excolonies
Excoloniesiv<- iv_robust(gdppc~expro_pro | settler_mortality,data=excolonies)


#excolonies2
Excolonies2iv<- iv_robust(gdppc~expro_pro | settler_mortality, data=excolonies2)


#non_african
Non_Africaniv<- iv_robust(gdppc~expro_pro | settler_mortality, data=non_african)


#Neo European
Neuiv<- iv_robust(gdppc~expro_pro | settler_mortality, data=neu)
#African
Africaiv<-iv_robust(gdppc~expro_pro | settler_mortality, data=african)
#Asian
Asiaiv<-iv_robust(gdppc~expro_pro | settler_mortality, data=asian)
#Non Asian
Non_Asianiv<-iv_robust(gdppc~expro_pro | settler_mortality, data=non_asian)
#Latin American and Caribbean
LaCiv<-iv_robust(gdppc~expro_pro | settler_mortality, data=lacar)
#Non Latin American and Caribbean
Non_LACiv<-iv_robust(gdppc~expro_pro | settler_mortality, data=nonlacar)

IV_in<- huxreg("Excolonies"=first_stage, "Non_Neu"=Excolonies2iv, "Non_African"=Non_Africaniv,"Non_Asia



#IV_ex<- huxreg("Neu"=Neuiv,"Africa"=Africaiv,"Asia"=Asiaiv,"Latin/Carib"= LaCiv)

caption(IV_in) <- "IV robust statistics of Table 1"
position(IV_in) <- "left"
```

```
IV_in
```

Table 2: IV robust statistics of Table 1

	Excolonies	Non_Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	9.341 ***	-0.141	4.554 ***	1.600	2.218 *
	(0.611)	(2.699)	(0.617)	(1.564)	(0.993)
settler_mortality	-0.607 ***				
	(0.127)				
expro_pro		1.281 **	0.578 ***	1.006 ***	0.850 ***
		(0.420)	(0.086)	(0.240)	(0.144)
N	64	60	37	55	40
R2	0.270	-0.688	0.587	0.045	0.439
logLik	-104.829				
AIC	215.659				

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

```
Excoloniesiv
```

```
##               Estimate Std. Error  t value     Pr(>|t|)   CI Lower CI Upper DF
## (Intercept) 1.9096665  1.2174380 1.568595 1.218325e-01 -0.5239573 4.343290 62
## expro_pro   0.9442794  0.1825866 5.171679 2.638652e-06  0.5792939 1.309265 62
```

Huxtable code for tables

```
exclusive_second_stage<-huxreg("Neu"=second_stage8,
                               "Africa"=second_stage3,
                               "Asia"=second_stage5,
                               "Latin/Carib"= second_stage1)

exclusive_first_stage<-huxreg("Neu"=first_stage8,
                              "Africa"=first_stage3,
                              "Asia"=first_stage5,
                              "Latin/Carib"= first_stage1)
exclusive_ols<-huxreg("Neu"=exprogdpneu,
                      "Africa"=exprogdpaf,
                      "Asia"=exprogdpas,
                      "Latin/Carib"= exprogdplacar)

exclusive_summary<-rbind(exclusive_second_stage,
                         exclusive_first_stage,
                         exclusive_ols)
caption(exclusive_summary)<-"Statistics of IV Regression of
GDP per Capita by Specific Regions"
position(exclusive_summary) <- "left"


inclusive_first_stage<-huxreg("Excolonies"=first_stage,
                              "Non_Neu"=first_stage2,
```

```
                                "Non_African"=first_stage4,
                                "Non_Asian"= first_stage6,
                                "Non_LAC"=first_stage6)
caption(inclusive_first_stage)<-"First Stage for
Average Protection against Expropriation"

inclusive_second_stage<-huxreg("Excolonies"=second_stage,"Non Neu"=second_stage2,"Non_African"=second_s
                                "Non_Asian"=second_stage6,
                                "Non_LAC"=second_stage7)


inclusive_ols<-huxreg("Excolonies"=exprogdp,
                      "Non_Neu"=exprogdp2,
                      "Non_African"=exprogdpnonaf,
                      "Non_Asian"= exprogdpnonas,
                      "Non_LAC"=exprogdpnonlacar)

caption(inclusive_ols)<-"Ordinary Least Squares"
allinclusive_summary<-rbind(inclusive_second_stage,
                            inclusive_first_stage,
                            inclusive_ols)
caption(allinclusive_summary)<-" Statistics of IV Regression
of GDP per Capita by Regions"
position(allinclusive_summary) <- "left"


#inclusive_first_stage
allinclusive_summary
```

```
exclusive_summary
```

Table 3: Statistics of IV Regression of GDP per
Capita by Regions

	Excolonies	Non Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	1.910 *	-0.141	4.554 ***	1.600	2.218 *
	(0.823)	(1.382)	(0.828)	(0.881)	(0.894)
Expropriation_Protection	0.944 ***	1.281 ***	0.578 ***	1.006 ***	0.850 ***
	(0.126)	(0.219)	(0.117)	(0.137)	(0.135)
N	64	60	37	55	40
R2	0.477	0.372	0.409	0.505	0.511
logLik	-72.268	-67.494	-38.703	-59.371	-47.851
AIC	150.536	140.988	83.406	124.743	101.702

*** p < 0.001; ** p < 0.01; * p < 0.05.

	Excolonies	Non_Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	9.341 ***	8.184 ***	11.844 ***	9.140 ***	9.140 ***
	(0.611)	(0.657)	(0.898)	(0.687)	(0.687)
settler_mortality	-0.607 ***	-0.391 **	-1.210 ***	-0.572 ***	-0.572 ***
	(0.127)	(0.133)	(0.219)	(0.139)	(0.139)
N	64	60	37	55	55
R2	0.270	0.130	0.466	0.243	0.243
logLik	-104.829	-93.976	-54.697	-90.923	-90.923
AIC	215.659	193.952	115.393	187.847	187.847

*** p < 0.001; ** p < 0.01; * p < 0.05.

	Excolonies	Non_Neu	Non_African	Non_Asian	Non_LAC
(Intercept)	4.660 ***	4.866 ***	5.221 ***	4.751 ***	4.194 ***
	(0.409)	(0.489)	(0.464)	(0.425)	(0.462)
expro_pro	0.522 ***	0.487 ***	0.482 ***	0.514 ***	0.549 ***
	(0.061)	(0.076)	(0.065)	(0.065)	(0.068)
N	64	60	37	55	40
R2	0.540	0.414	0.611	0.543	0.629
logLik	-68.168	-65.419	-30.945	-57.171	-42.337
AIC	142.335	136.838	67.890	120.342	90.674

*** p < 0.001; ** p < 0.01; * p < 0.05.

Graphs and Tables

Figures 1-10 are below

```
exprogdp<-lm(gdppc~expro_pro, data=excolonies)

ggplot(data=excolonies)+
  geom_point(mapping=aes(x=expro_pro,
                         y=gdppc, color=counttrycol))+
  geom_abline(intercept = 4.660,
              slope=0.522, color="red")+theme(legend.position = "bottom",
        text=element_text(size=10))+labs(title="Figure 1: Expropriation Protection VS GDP",x="Protectio
```

Table 4: Statistics of IV Regression of GDP per Capita by Specific Regions

	Neu	Africa	Asia	Latin/Carib
(Intercept)	2.190	-6.765	-0.661	3.763
	(4.447)	(5.981)	(3.135)	(2.556)
Expropriation_Protection	0.802	2.400 *	1.227 *	0.737
	(0.459)	(1.018)	(0.433)	(0.396)
N	4	27	9	24
R2	0.604	0.182	0.535	0.136
logLik	3.341	-27.170	-10.623	-19.773
AIC	-0.681	60.339	27.246	45.545

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

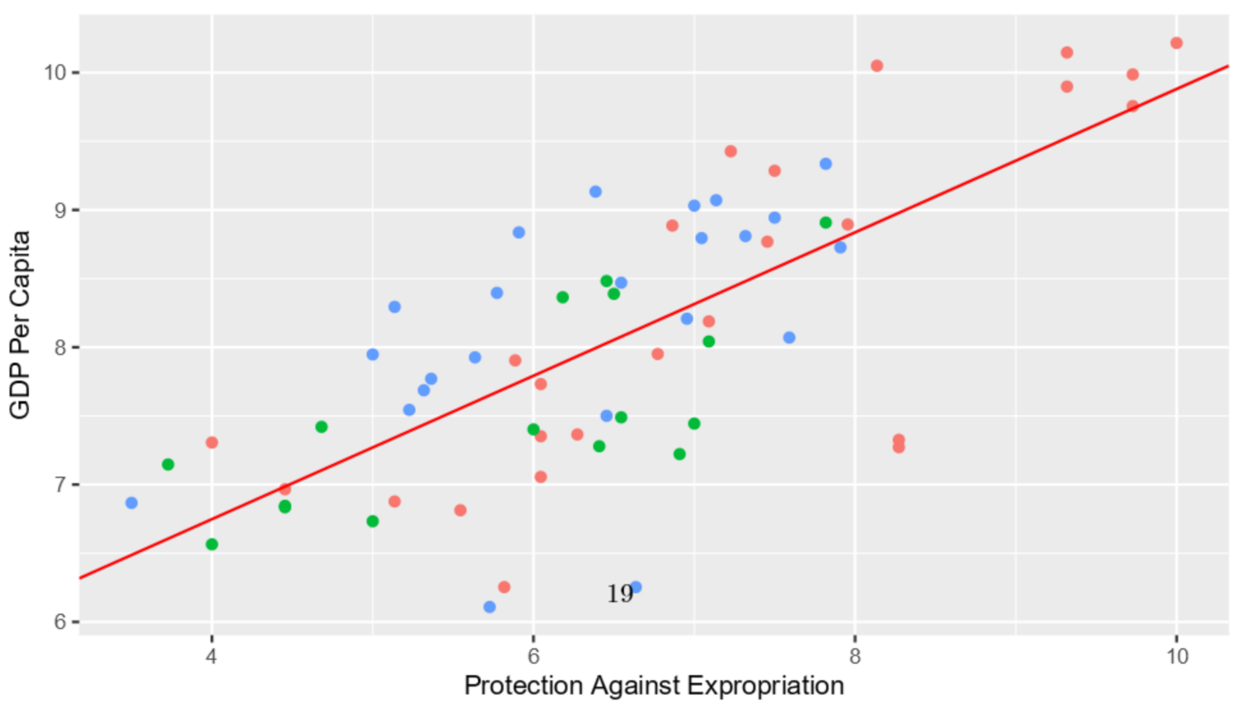
	Neu	Africa	Asia	Latin/Carib
(Intercept)	8.374 *	6.475 ***	10.329 ***	9.210 ***
	(1.057)	(1.123)	(1.896)	(2.001)
settler_mortality	0.540	-0.108	-0.812	-0.636
	(0.429)	(0.198)	(0.482)	(0.458)
N	4	27	9	24
R2	0.442	0.012	0.288	0.081
logLik	1.141	-43.020	-13.474	-34.097
AIC	3.717	92.040	32.948	74.193

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

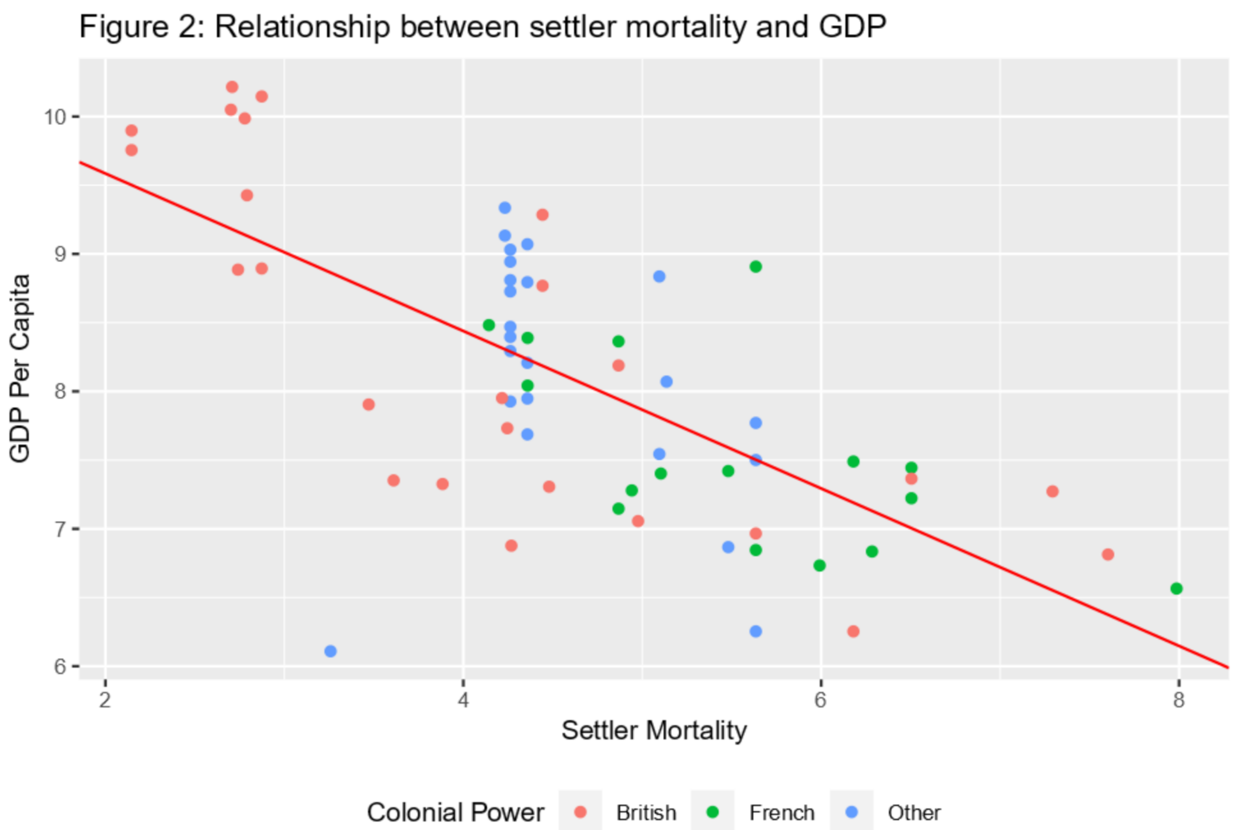
	Neu	Africa	Asia	Latin/Carib
(Intercept)	6.052	5.566 ***	3.134	5.541 ***
	(3.800)	(0.637)	(1.567)	(0.462)
expro_pro	0.404	0.302 **	0.701 *	0.461 ***
	(0.392)	(0.106)	(0.214)	(0.071)
N	4	27	9	24
R2	0.346	0.244	0.606	0.658
logLik	2.336	-26.107	-9.881	-8.648
AIC	1.328	58.215	25.762	23.297

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

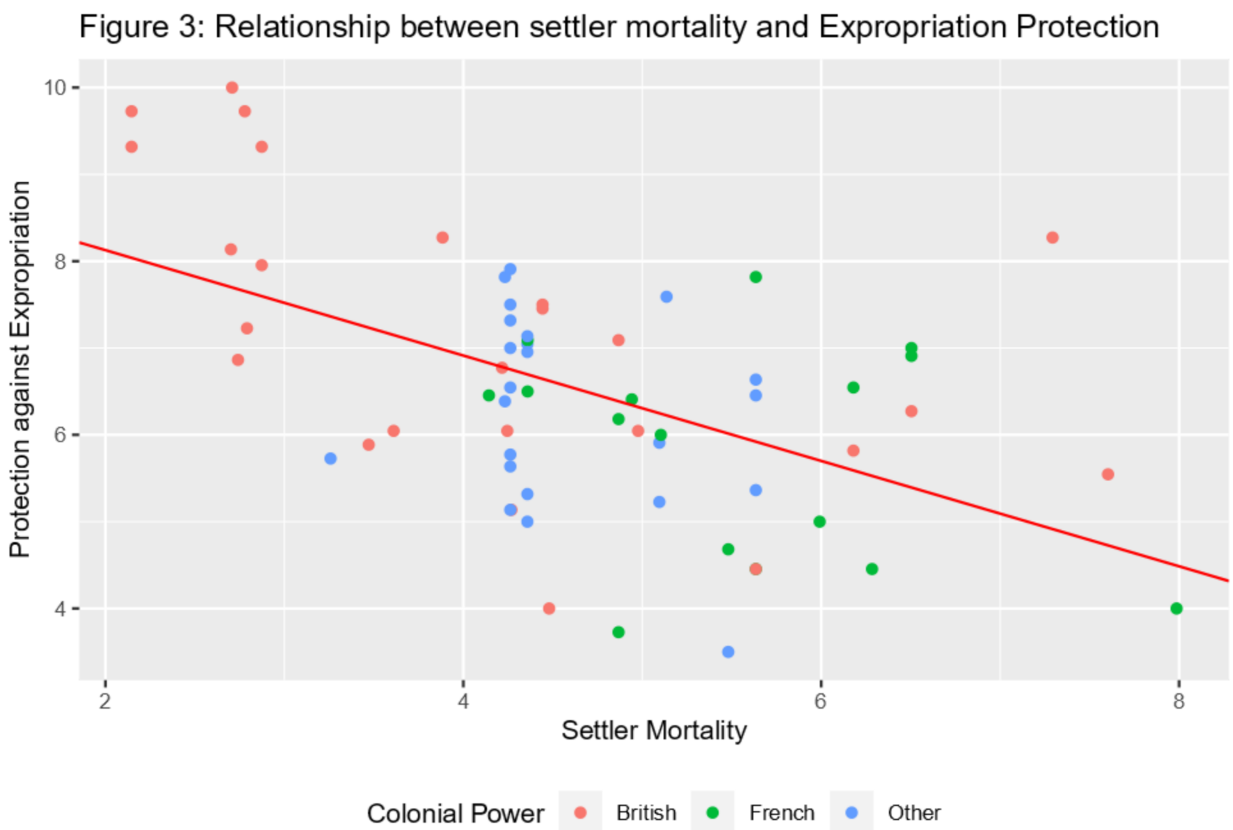
Figure 1: Expropriation Protection VS GDP



```r
mortaldppccol<-lm(gdppc~settler_mortality,data=excolonies)
ggplot(data=excolonies)+
  geom_point(mapping=aes(x=settler_mortality,
                         y=gdppc, color=counttrycol))+
  geom_abline(intercept = 10.731,
              slope=-0.573, color="red")+theme(legend.position = "bottom",
        text=element_text(size=10))+labs(title="Figure 2: Relationship between settler mortality and GDI
```



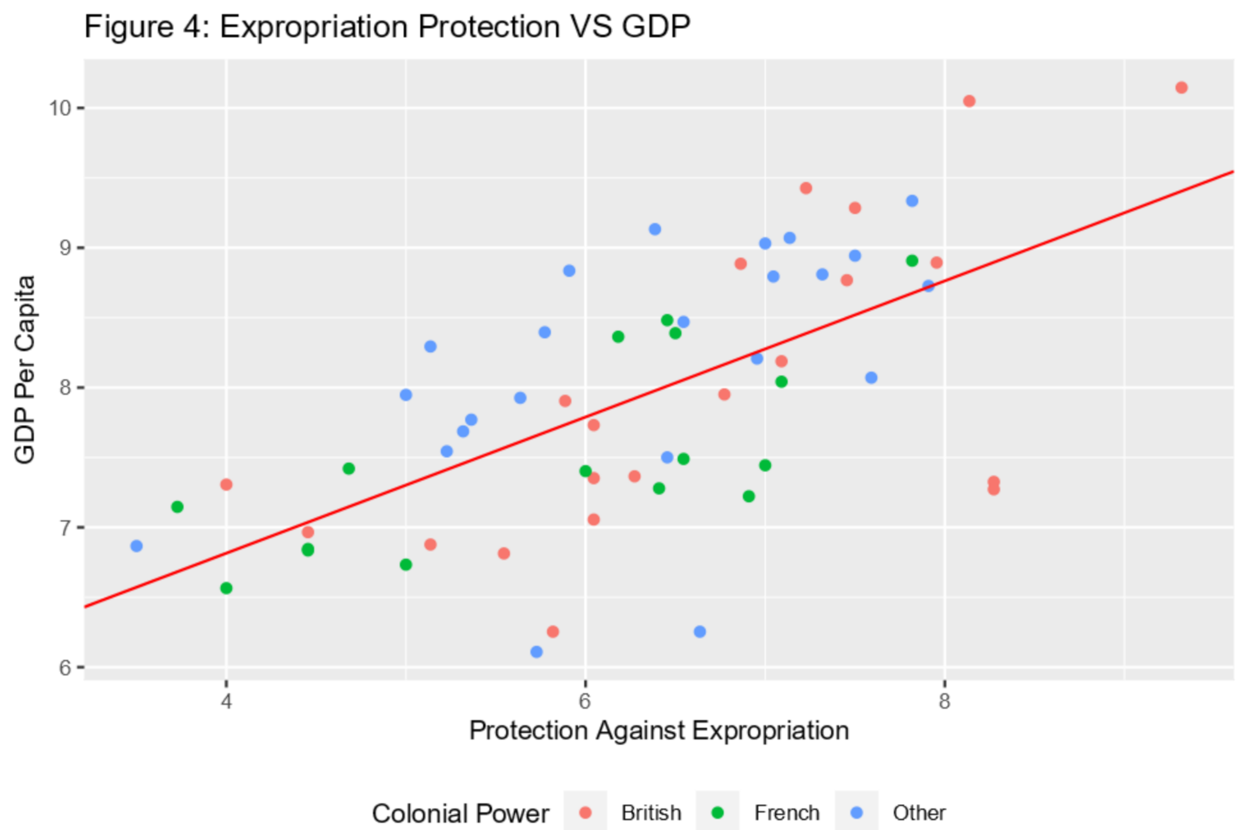

```r
mortalexprocol<-lm(expro_pro~settler_mortality,data=excolonies)
ggplot(data=excolonies)+
  geom_point(mapping=aes(x=settler_mortality,
                         y=expro_pro,colour=counttrycol))+
  geom_abline(intercept = 9.341,
              slope=-0.607, color="red")+theme(legend.position = "bottom",
        text=element_text(size=10))+labs(title="Figure 3: Relationship between settler mortality and Exp
```



Expropriation vs GDP for Non Neo European countries

```
exprogdp2<-lm(gdppc~expro_pro, data=excolonies2)

ggplot(data=excolonies2)+
  geom_point(mapping=aes(x=expro_pro, y=gdppc, color=counttrycol))+
  geom_abline(intercept = 4.867,
              slope=0.487, color="red")+
  theme(legend.position = "bottom",
        text=element_text(size=10))+
  labs(title="Figure 4: Expropriation Protection VS GDP",
       x="Protection Against Expropriation", y="GDP Per Capita",
       colour="Colonial Power")
```

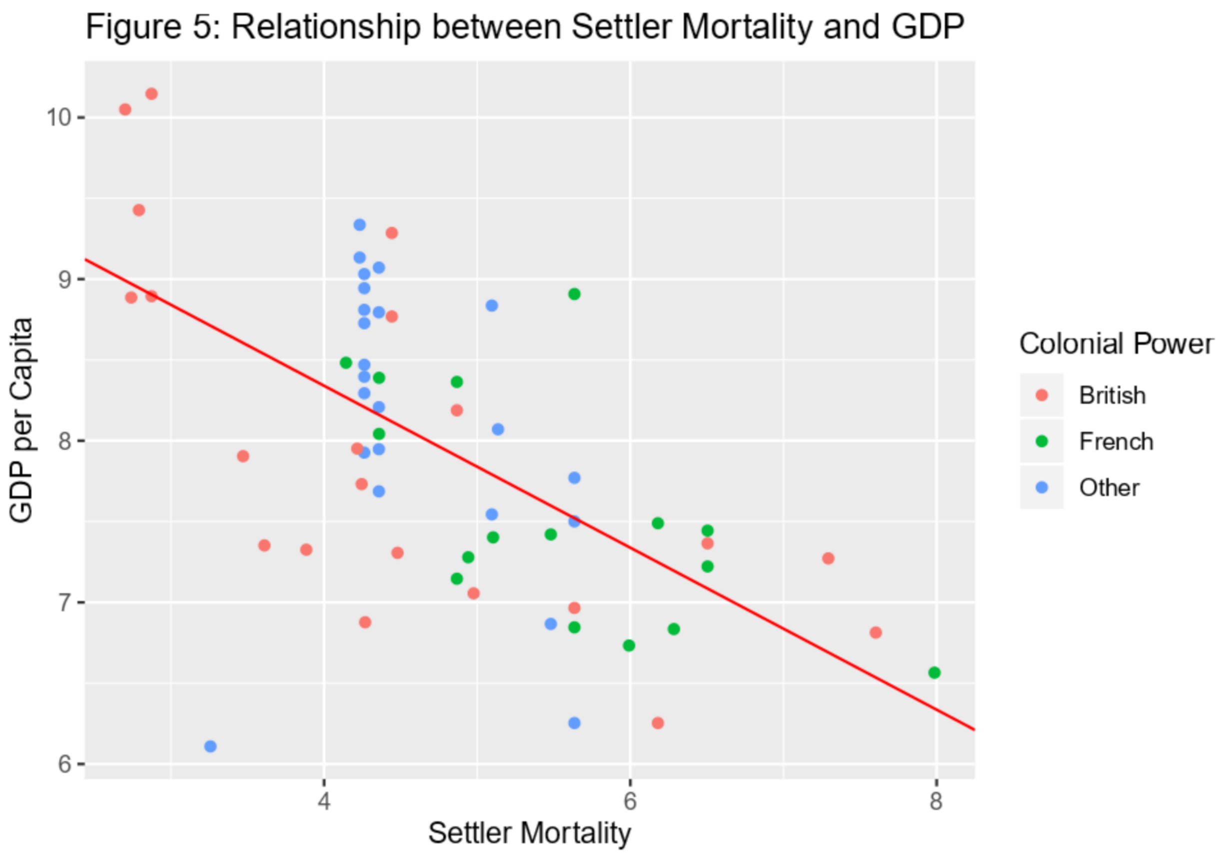


GDP vs Settler mortality without Neu European Countries

```
mortalgdp<-lm(gdppc~settler_mortality,data=excolonies2)
summary(mortalgdp)
```

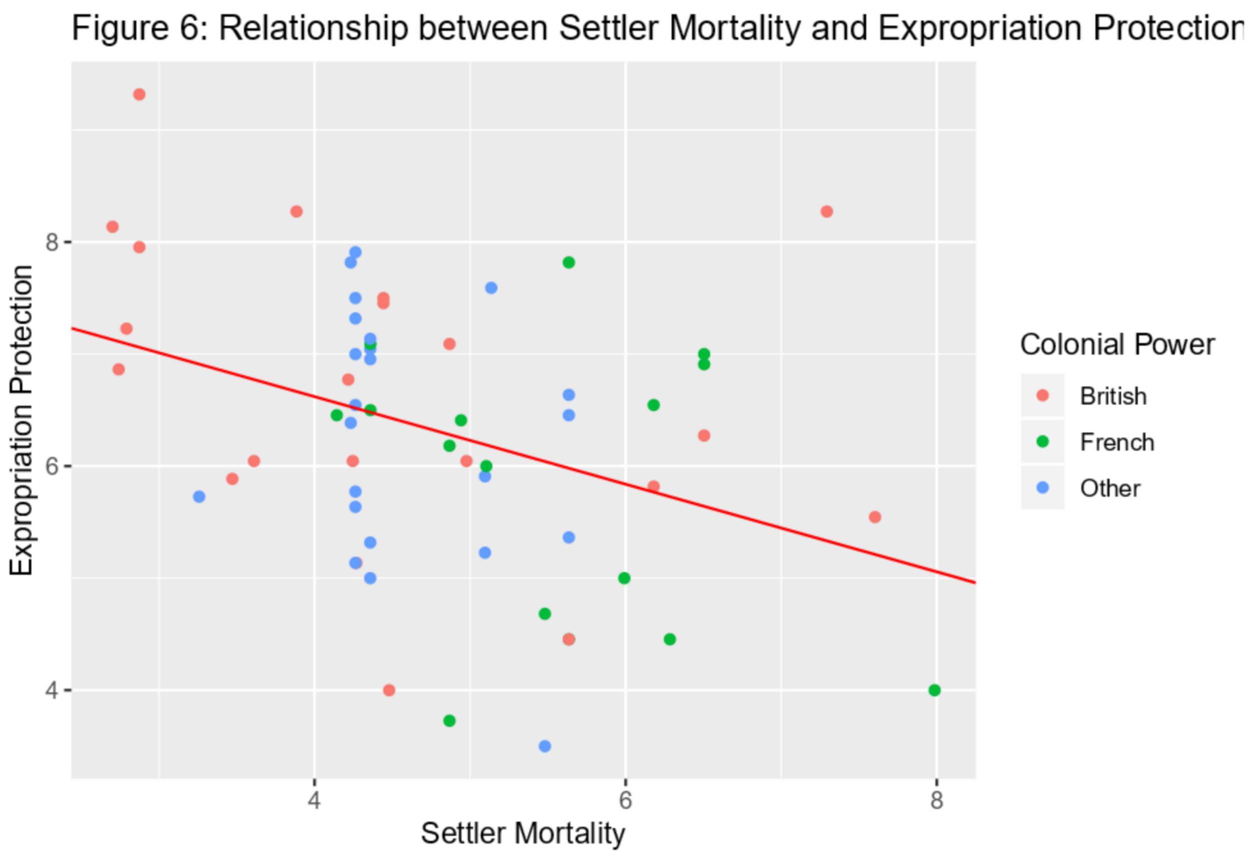
```
##
## Call:
## lm(formula = gdppc ~ settler_mortality, data = excolonies2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6014 -0.5011  0.1131  0.4919  1.3886
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       10.34381    0.42263  24.475  < 2e-16 ***
## settler_mortality -0.50127    0.08558  -5.858 2.34e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.758 on 58 degrees of freedom
## Multiple R-squared:  0.3717, Adjusted R-squared:  0.3609
## F-statistic: 34.31 on 1 and 58 DF,  p-value: 2.337e-07
```

```r
ggplot(data=excolonies2)+
  geom_point(mapping=aes(x=settler_mortality,
                         y=gdppc, color=counttrycol))+
  geom_abline(intercept = 10.344,
              slope=-0.501, color="red")+
  labs(title="Figure 5: Relationship between Settler Mortality and GDP",
       x="Settler Mortality", y="GDP per Capita",colour="Colonial Power")
```



Expropriation Protection vs Settler mortality without Neu European Countries

```r
mortalexpro<-lm(expro_pro~settler_mortality,data=excolonies2)
ggplot(data=excolonies2)+
  geom_point(mapping=aes(x=settler_mortality,y=expro_pro,colour=counttrycol))+
  geom_abline(intercept = 8.184,
              slope=-0.391, color="red")+
  labs(title="Figure 6: Relationship between Settler Mortality and Expropriation Protection",
       x="Settler Mortality", y="Expropriation Protection",colour="Colonial Power")
```



Latin America and Caribbean colonies breakdown by colonial power

```
ggplot(data=lacar)+
  geom_point(mapping=aes(x=settler_mortality, y=gdppc, color=counttrycol))+
  geom_abline(intercept = 10.55, slope=-0.4687, color="red")+
  labs(title="Figure 7: Relationship between Settler Mortality and GDP",
       x="Settler Mortality", y="GDP per Capita",colour="Colonial Power")
```

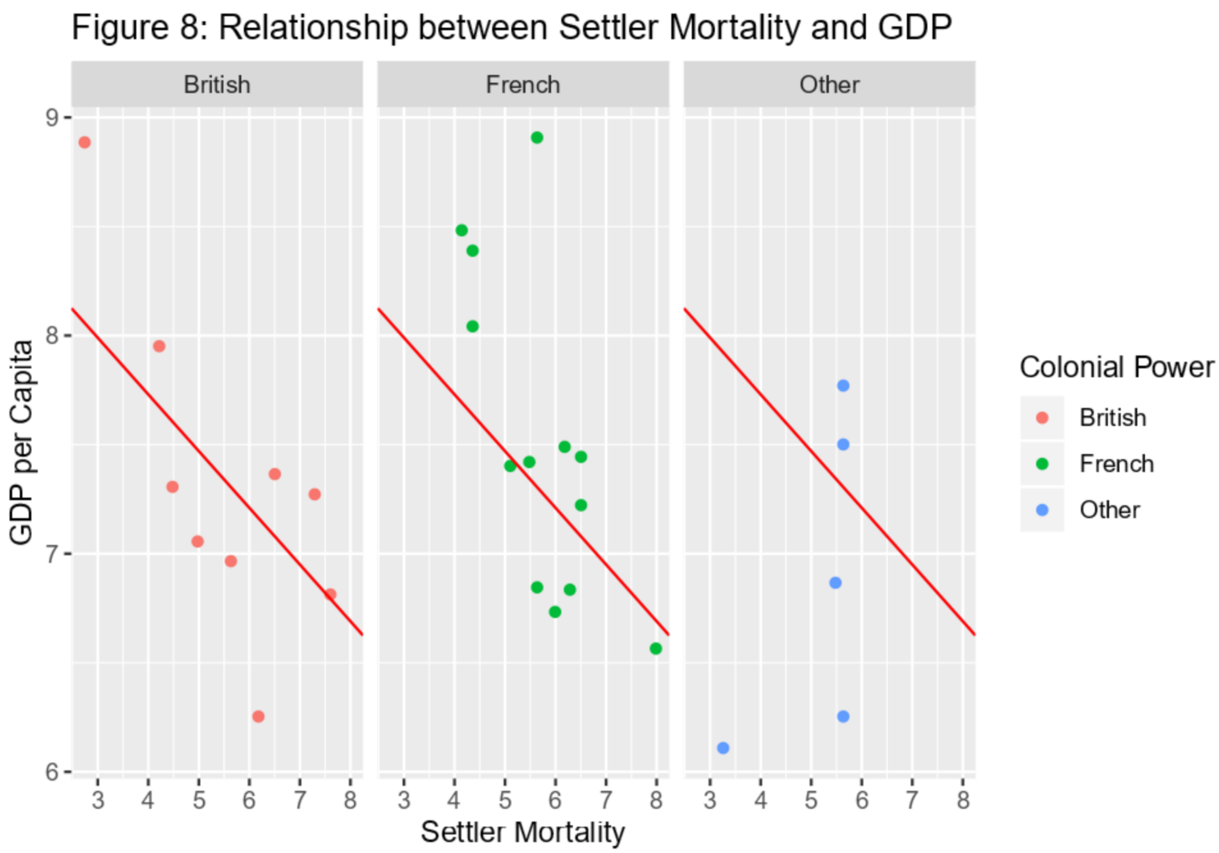

```r
#facet_wrap(vars(counttrycol))
```

African colonies breakdown by colonial power

```r
summary(afmortalgdp)
```

```
##
## Call:
## lm(formula = gdppc ~ settler_mortality, data = african)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.82197 -0.38752  0.06745  0.37402  1.59484
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.7787     0.6246  14.056 2.26e-13 ***
## settler_mortality  -0.2601     0.1103  -2.359   0.0265 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6879 on 25 degrees of freedom
## Multiple R-squared:  0.182,  Adjusted R-squared:  0.1493
## F-statistic: 5.563 on 1 and 25 DF,  p-value: 0.02647
```

```r
ggplot(data=african)+
  geom_point(mapping=aes(x=settler_mortality,y=gdppc, color=counttrycol))+
  geom_abline(intercept = 8.77, slope=-0.260, color="red")+
  labs(title="Figure 8: Relationship between Settler Mortality and GDP ",
       x="Settler Mortality", y="GDP per Capita",colour="Colonial Power")+
  facet_wrap(vars(counttrycol))
```



Asian colonies breakdown by colonial power

```r
summary(asmortalgdp)
```

```
##
## Call:
## lm(formula = gdppc ~ settler_mortality, data = asian)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.06549 -0.81914 -0.05288  0.72520  1.17365
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        12.0180     1.3810   8.702 5.31e-05 ***
## settler_mortality  -0.9971     0.3514  -2.837   0.0251 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8932 on 7 degrees of freedom
## Multiple R-squared:  0.5349, Adjusted R-squared:  0.4684
## F-statistic: 8.049 on 1 and 7 DF,  p-value: 0.02515
```

```r
ggplot(data=asian)+
  geom_point(mapping=aes(x=settler_mortality,y=gdppc, color=counttrycol))+
  geom_abline(intercept = 12.55, slope=-0.9971, color="red")+
  labs(title="Figure 9: Relationship between Settler Mortality and GDP",
       x="Settler Mortality", y="GDP per Capita",colour="Colonial Power")
```


```r
#facet_wrap(vars(counttrycol))
```

Neo European colonies breakdown by colonial power

```r
summary(neumortalgdp)
```

```
##
## Call:
## lm(formula = gdppc ~ settler_mortality, data = neu)
##
```

```
## Residuals:
##        1        2        3        4
##  0.06322 -0.12227 -0.07861  0.13766
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.9058     0.6098  14.605  0.00466 **
## settler_mortality   0.4329     0.2476   1.748  0.22251
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1484 on 2 degrees of freedom
## Multiple R-squared:  0.6045, Adjusted R-squared:  0.4067
## F-statistic: 3.057 on 1 and 2 DF,  p-value: 0.2225
```

```r
ggplot(data=neu)+
  geom_point(mapping=aes(x=settler_mortality,
                         y=gdppc, color=counttrycol))+
  geom_abline(intercept = 8.91, slope=0.433, color="red")+
  labs(title="Figure 10: Relationship between Settler Mortality and GDP",
       x="Settler Mortality", y="GDP per Capita",colour="Colonial Power")
```