

Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Colonial Origin and Human Capital

A cross-country analysis of European colonial Settlement

By Luuk Grandia (S1048241)

Nijmegen, 16 November 2022

Program: Master's Program in Economics
Specialisation: International Economics & Development
Supervisor: Prof. Natasha Wagner

Radboud Universiteit



Page of Content

1	Introduction.....	3
2	Theoretical Background	5
2.1	<i>Colonial origin and economic development</i>	<i>5</i>
2.2	<i>The role of Institutions.....</i>	<i>7</i>
2.3	<i>The role of Human Capital.....</i>	<i>8</i>
2.4	<i>Gap in the literature</i>	<i>9</i>
2.5	<i>Hypotheses</i>	<i>11</i>
3	Data & Methodology.....	11
3.1	<i>Data.....</i>	<i>12</i>
3.2	<i>Econometric Specification</i>	<i>12</i>
3.3	<i>Empirical Approach</i>	<i>13</i>
3.4	<i>Summary Statistics & Variable Definitions</i>	<i>14</i>
3.5	<i>Dependent Variable: Human Capital.....</i>	<i>17</i>
3.6	<i>Explanatory Variable: Euro Share.....</i>	<i>18</i>
3.7	<i>Control Variables.....</i>	<i>19</i>
3.8	<i>Endogeneity Issues</i>	<i>19</i>
3.9	<i>Instrumental Variables</i>	<i>20</i>
4	Results	21
4.1	<i>Ordinary Least Squares regressions (OLS)</i>	<i>22</i>
4.2	<i>Two-Stage Least Squares regressions (2SLS).....</i>	<i>24</i>
4.3	<i>Robustness Checks: Repeating regressions with an alternative Human Capital Measure</i>	<i>27</i>
4.4	<i>Robustness Checks: Repeating regressions for 1995, 2000 and 2005</i>	<i>28</i>
5	Conclusion	31
6	Discussion/Research Limitations	33
7	Bibliography.....	34

Abstract

This research revisits the relation between colonial origin and economic development by focusing directly on the role between European colonial settlement and human capital. By conducting OLS regressions and 2SLS regressions utilizing recent data on European colonial settlement this research finds evidence suggesting a significant cross-country relation between European colonial settlement and human capital. The findings of this research suggest that besides the persistent negative effects of extractive institutions established by prior research, colonial origin also shows positive effects on human capital. These positive effects hold when only including countries with a low European colonial settlement. The findings of this research also suggest that the positive effects on human capital were far greater per settler amongst countries with a low share of European colonial settlement. This suggests that even within extractive colonies European colonizers implemented processes that helped gain long-term development. Further within-country research on these processes could help improve policy directed at improving human capital accumulation in individual formerly colonized countries.

1 Introduction

Colonialism is widely considered to be one of the most shameful events in Europe's recent history. It is difficult to consider the great differences in paths of development across the world without noticing the fact that some of the most underdeveloped countries in the world have a dark colonial history. The idea that European colonialism has played a role in the current underdevelopment of some of its former colonies has been discussed in various historical works like Walter Rodney's *How Europe underdeveloped Africa* (1972).

In the field of economics, the relation between colonial origin and economic development has also been widely studied. Starting in the late 90's a surge of empirical studies emerged looking at the relation between colonialism and economic development using instrumental variable techniques (Engerman et al., 1994; La porta et al., 1997; Acemoglu et al., 2001).

For a long time, a widely accepted theory has been that conization focused on extracting resources have led to the instalment of weak extractive institutions and that this process has had persistent negative effects on economic development. The theory continues to suggest that in settler colonies such as the United States and Canada colonizers installed more inclusive institutions that were aimed at protecting property rights of European settlers (Acemoglu et al., 2001)¹.

An alternative view on the relation between colonial origin and economic development is that European Settlers brought human capital and technologies, that have had persistent positive effects on economic development, regardless of whether a colony was focused primarily on resource extraction or settlement (Glaeser et al., 2004; Easterly & Levine, 2016).

The debate around colonial origin and economic development has largely been about the question whether institutions or human capital play a bigger role in economic development. On the one hand there are proponents of the institutions view which has its roots in New Institutional economic theory (North & Thomas, 1973; AJR, 2001; Acemoglu et al., 2014) while on the other hand there are proponents of the human capital view which has its roots in Lipset's

¹ This 2001 research conducted by Daron Acemoglu, Simon Johnson, and James A. Robinson is often referred to as the AJR article and will be referred to in this manner throughout this document.

(1960) modernization hypothesis and is more in line with the neoclassical view that human capital is one of the main drivers behind economic development (Peck, 2016).

Ever since the AJR (2001) study, the institutionalist view has been confirmed by various studies looking at specific countries (e.g. Banerjee & Iyer., 2005; Dell, 2010). More recently however, within-country evidence has also emerged supporting the human capital view (e.g. Wantchekon, 2015; Dell & Olken, 2020).

This research focusses on the effect of colonial origin and human capital without specifically focussing on the relation between human capital and economic development. The idea that economic development measured by GDP per capita is the ultimate end goal in terms of economic development has long been contradicted by economists like Amartya Sen (2009) and organizations like the United Nations (UNDP, 1990). This research aims to contribute to providing more convincing evidence towards the relation between colonial origin and human capital, treating human capital not as a channel towards economic development, but as a development goal on its own. This research therefore thrives to answer the following research question:

Has European colonial settlement had long-term positive effects on human capital?

The most recent cross-country study suggesting a positive relationship between European colonial settlement and human capital is conducted by Easterly & Levine (2016). Acemoglu et al., (2014)² have criticized their findings by suggesting that their measurement for colonial settlement is inaccurate and that they fail to treat for endogeneity issues in their OLS regressions. To provide more evidence on the relation between colonial origin and human capital this research uses Easterly & Levine's (2016) dataset in combination with data from Barro-Lee (2013) and from the World Development Indicators database (World Bank, 2022) to directly regress colonial origin against human capital. To account for Acemoglu et al.'s (2014) critique on earlier research supporting the human capital view the empirical strategy of this research incorporates both OLS regressions as well as 2SLS regressions using instrumental variables to treat for endogeneity. The

² Easterly & Levine (2016) updated their research which was originally released in 2013. Therefore Acemoglu et al., could critique it as early as 2014. The points of critique raised by Acemoglu et al., (2014) have not been accounted for in the updated version of Easterly & Levine's (2016) article.

empirical strategy of this research is modelled after Angeles & Neanidis (2015), who look at the relation between colonial origin and corruption.

In chapter 2 a further discussion will be held about the most relevant literature surrounding the relation between colonial origin, institutions and human capital while also elaborating on how this relationship is theorized by both the institutions view and the human capital view. In chapter 3 the econometric specification, data, variables, and empirical strategy will be discussed. Chapter 4 will discuss the results of the OLS regressions, the 2SLS regressions and will include robustness checks to ensure that the measured results hold across different years and different human capital measures. Chapter 5 consists of concluding remarks along with suggestions for further research. In chapter 6 the various limitations that are inherent to this type of research are discussed to ensure transparency regarding the findings of this research.

2 Theoretical Background

In this chapter the most relevant prior research on the relationship between colonial origin, institutions, human capital, and economic development are discussed along with its theoretical implications. Paragraph 2.1 discusses prior research on colonial origin and economic development in general. Paragraph 2.2 elaborates on the institutions view while paragraph 2.3 discusses the human capital view. Paragraph 2.4 discusses how this research contributes to the existing literature and in paragraph 2.5 two hypotheses are formulated which will be tested in the empirical analysis of this research.

2.1 Colonial origin and economic development

Around two decades ago the field of economics started to produce an increasing amount of literature suggesting that historic events have had significantly persistent effects on economic development (Nunn, 2009, 2020). The origins of this type of literature can be traced back to three lines of research, all of which are focused on the relation between colonial rule and long-term economic development.

Firstly, Engerman et al. (1994) argued that the difference in economic growth paths between countries in the Americas can be significantly attributed to initial factor endowments that made

some states more suitable as slave colonies. They conclude that these slave colonies were characterized by attributing large concentrations of wealth to a small elite leading to slower economic growth.

Secondly La Porta et al. (1997) used legal origin of former colonizers as an instrumental variable to determine how protection of investor rights affect financial development. In their study they found evidence that differences in legal systems set up by colonial powers in colonized regions have had long-term effects on economic development.

Thirdly AJR (2001), look at European mortality rates in colonial regions to establish that extractive institutions set up by colonial powers have had a persistent negative effect on economic development. They argue that high European mortality rates prevented European settlement in colonies which increased the chance of them setting up extractive institutions. They found that European settlement is positively correlated with economic development and therefore conclude that these extractive institutions, and thus institutions in general play a significant role in determining a countries path of development.

After the AJR (2001) article many cross-country, cross-regional and within-country studies focussed on the relationship between colonialism and economic development. The vast majority focussed on the role of extractive institutions (Banerjee & Iyer, 2005; Acemoglu et al., 2012; Bruhn & Gallego, 2012; Naritomi et al., 2012; Michalopoulos & Papaioannou, 2013; Dell & Olken, 2020) while some focussed on the role of human capital that was potentially brought by European settlers (Glaeser et al., 2004; Wantchekon et al, 2013; Easterly & Levine, 2016). Some cross-regional studies concluded that effects on economic development depend per region based on the type of colonial activities conducted (Bruhn & Gallego, 2012), the level of colonial investment (Huillery, 2009) or the type of indirect rule (Iyer, 2010).

When it comes to the role of institutions there seems to be close to a consensus that extractive institutions set up by colonizers tend to have persistent negative effects on current economic development with the only clear exception being Dell & Olken (2020) findings on colonial regions in Java. When it comes to the few studies that have been conducted on the impact that colonialism has had on human capital, findings show a positive relationship. However, the number of studies focussing on human capital specifically is smaller and there have been cases like Dell

(2010) measuring negative effects of colonialism on human capital. In the next two paragraphs the two different theoretical frameworks behind the potential roles of institutions and human capital in explaining the relation between colonialism and economic development are explained.

2.2 The role of Institutions

When the AJR (2001) article on the relation between colonial origin and economic development was released the two variables of interest were essentially institutions and economic development. The relation between colonial origin and economic development is based upon the assumption that the quality of a country's institutions is positively correlated to a country's economic development. To fully understand their argument a closer look at the theoretical framework behind the AJR (2001) research and the relation between institutions and economic development is in order.

In their most recent discussion on the relation between institutions, human capital, and economic development Acemoglu et al. (2014) present two diagrams based on an institutional framework by North and Thomas (1973). Originally, North and Thomas (1973) presented the diagram shown in figure 1. Acemoglu et al. (2014) expand this to a diagram shown in figure 2. Considering this research is focused on European colonial settlement, figure 2 is adjusted to show the full relation between European colonial settlement and economic development as theorized by AJR (2001). This full relation is shown in figure 3. The overall relationship between European Settlement and Economic development is positive, but this is explained as being a negative relation between colonialism in countries with a small European Settlement, current institutions, and therefore current Economic Development (Acemoglu et al., 2001; Acemoglu et al., 2014).

Figure 1: Diagram by North & Thomas (1973)

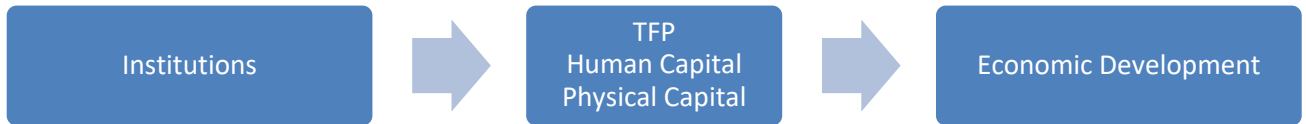
Figure 2: Diagram by Acemoglu et al. (2014)³

Figure 3: Mechanism behind the institutional view



This theory along with the evidence presented in AJR (2001) has played a big role in pushing the idea that ‘Institutions Matter’ for economic development (Acemoglu et al., 2005). In response to this, proponents of the idea that human capital is the main driver behind economic growth have presented a different theory in explaining the relation between European colonial settlement and economic development.

2.3 The role of Human Capital

A few years after the AJR (2001) article an empirical study by Glaeser et al., (2004) challenged the role of institutions in explaining the relationship between European colonial settlement and economic development by instead proposing a theoretical framework in which human capital is placed as the main driver behind economic development. In this theoretical framework institutions play a far smaller role in explaining this relation and its increase is a result of improved economic development and human capital. Glaeser et al. (2004) suggest their results to be more in line with research by Lipset (1960) and Djankov et al. (2003) stating that human capital and

³ TFP stands for Total Factor Productivity

economic growth primarily lead to better institutions instead of the other way around. The human capital view explicitly diverts from endogenous growth theory brought forth by economists such as Lucas (1988) in that it implies that human capital influences economic development through institutions, rather than through technology. A graphical representation of the proposed mechanism behind the relation between European colonial settlement and economic development according to the human capital view is shown in Figure 4.

Figure 4: Mechanism behind the Human Capital view



More recently Easterly & Levine (2016) have elaborated on this view by running regressions on the relation between European colonial settlement and economic development using newly computed colonial settlement data. Easterly & Levine's (2016) findings are in line with the Human Capital view and suggest that human capital has the main role in explaining the relation between European colonial settlement and economic development. Another important new finding potentially in line with the Human Capital view is that Dell & Olken (2020) found significantly positive effects on various human capital outcomes in regions in which Dutch colonizers set up extractive colonial institutions in Java. Dell & Olken (2020) conduct their research based on the institutional view considering they focus specifically on extractive institutions. However, their findings that extractive colonial institutions have had positive effect on development outcomes could potentially be explained by the fact that the Dutch colonizers brought human capital to Java to properly run their extractive institutions.

2.4 Gap in the literature

Prior research looking at the relationship between colonial origin and economic development seems to be standing in the middle of a larger economic debate. On the one hand, studies emphasizing the role of institutions in explaining the relation between the share of European colonial settlement and economic development have been focused at providing evidence for the

new institutionalist idea that institutions matter (North & Thomas, 1973; AJR, 2001; Acemoglu et al., 2014). On the other hand, studies emphasizing the role of human capital in explaining this relation have been focused at discrediting the idea that institutions matter and argue that the main driver behind economic development is human capital. The debate surrounding these studies has primarily been focused on the relationship between human capital, institutions, and economic development and not about the relationship between colonialism and development indicators like human capital and institutions themselves. Therefore, these studies typically use economic development in the form of GDP per capita as a dependent variable.

Economists like Amartya Sen (2009) and development indicators like the Human Development Index (UNDP, 1990) argue that human capital accumulation is not a channel towards economic development but economic development on its own. Furthermore, Acemoglu et al. (2014) have voiced critique on Easterly & Levine's (2016) research emphasizing the potential cross-country relation between European colonial settlement and human capital by addressing that they don't treat for endogeneity and that their self-computed Euro Share variable might not accurately represent European colonial settlement.

This research contributes to the existing literature by for the first time regressing European colonial settlement directly on human capital and by following Easterly & Levine's (2016) strategy of also looking at the results when limiting the sample to countries with a small European colonial settlement. By also treating for endogeneity this research addresses Acemoglu et al.'s (2014) main critique on Easterly & Levine (2016). Acceptance of the hypotheses in paragraph 2.5 will therefore strengthen the evidence suggesting that there is a positive relation between European colonial settlement and human capital. This research is explicitly not aimed at disproving the institutions view and acknowledges the possibility that both views contain truth about the relation between colonial origin and economic development simultaneously. If European colonial settlement has indeed had long-term positive effects on human capital further research on this relationship could help development economists gain further insight in how to improve policies aimed at human capital accumulation.

2.5 Hypotheses

If European colonialism has had significant effects on current human capital than the results of this research should show a positive relation between the share of Europeans that settled during colonization and current human capital. If this positive relation is also present amongst a sample including only countries that have a small share of European settlement this would rule out the possibility that the perceived positive relation between European settlement and human capital is entirely due to inclusive institutions set up in settler colonies (AJR, 2001). Therefore, the following two hypotheses have been formed.

H1: The share of European settlers in former colonies during the early stage of their colonial period is positively and significantly correlated with human capital today.

H2: The share of European settlers in former colonies during the early stage of their colonial period is positively and significantly correlated with human capital today amongst colonies with a low share of European settlers.

The two hypotheses formulated in this paragraph are accepted at a 5% significance level. An acceptance of both hypotheses will be treated as evidence in favor of answering the research question positively.

3 Data & Methodology

In this chapter the data, variables and empirical strategy aimed at answering the research question are discussed. In paragraph 3.1 the composition of the dataset is discussed along with some remarks about sample size. In paragraph 3.2 the Econometric Specification that represents the baseline of the empirical analysis is formulated. Paragraph 3.3 discusses the empirical approach used in this research and why this form has been chosen. Paragraph 3.4 up until 3.7 discuss the main variables used in this research along with the summary statistics and variable definitions shown in table 1 and table 2 respectively. Finally, in paragraph 3.8 and 3.9 the

endogeneity issues that come with this type of research and the instrumental variables used to treat for the endogenous variables are discussed.

3.1 Data

To conduct the empirical part of this research a cross-sectional dataset has been constructed by combining various data sources. The dataset is constructed starting with Easterly & Levine's (2016) data on the share of Europeans in colonized countries during colonial times. They also incorporated the dataset of Acemoglu et. al (2001) which contains some potential instrumental variables that will be discussed in the next paragraph of this chapter. The data regarding human capital is retrieved from the Barro-Lee (2013) dataset which contains data regarding various educational variables measured from 1950-2010 in intervals of 5 years. All other data has been retrieved from the World Bank's World Development Indicators database (2022). All data regarding variables that represent current values are from the year 2010 because this is the most recent year for which data was available for all variables included. In total the dataset consists of 129 countries and the sample used in the regressions vary between 69 and 100. European countries are excluded because Easterly & Levine (2016) focused on the imperial colonial period in which Europeans were the colonizers. Countries that were never colonized are excluded from the sample. Due to the nature of this research the number of observations is relatively low. Sample size therefore plays an important role in variable selection.

3.2 Econometric Specification

The baseline econometric specification used in this research can be denoted as:

$$(1) \quad HC_i = \alpha + \underset{(+)}{\beta} ES_i + \underset{(+)}{\gamma} (\ln) CI_i + \underset{(+)}{\delta} P_i + \varepsilon_i$$

In equation (1) HC_i represents a measure for human capital for country i , ES_i represents the share of Europeans that settled in country i during the colonial period. $(\ln) CI_i$ represents a

measure for current income per capita and is measured in its logarithmic form to ensure a normal distribution. Finally, P_i represents a variable controlling for population density or distribution.

3.3 Empirical Approach

This research essentially builds on Easterly & Levine's (2016) study on the relation between colonial origin and economic development, who in turn build forth on Acemoglu et al. (2001). The empirical approach of this research is heavily inspired by Angeles & Neanidis (2015), who utilize their own Euro Share variable to look at the relation between colonial origin and corruption. Their empirical approach sets a good example for this research because they use a similar explanatory variable and face similar endogeneity problems as this research. In a cross-country analysis looking at the relationship between colonial history and economic development researchers are faced with a trade-off between ignoring endogeneity issues or controlling for endogeneity using imperfect instrumental variables. Considering this dataset deals with multiple endogenous variables, a small sample size and a limited availability of instrumental variables the strength of evidence that can be provided by using only one method can be easily questioned. Therefore, the empirical analysis includes both OLS and 2SLS regressions. By also including 2SLS regressions this research accounts for the critique raised by Acemoglu et al. (2014) that Easterly & Levine (2016) ignored endogeneity and that their Euro Share variable might be imperfectly computed. By combining OLS regressions and the 2SLS regressions replacing this Euro Share variable with an instrumental variable both these points of critique will be accounted for. The hypotheses shown in chapter 2 will only be accepted if the results using both methods are shown to be significant.

The empirical approach of this research can be divided into the following three parts:

- (1) OLS regressions treating both Euro Share and GDP per capita as exogenous
- (2) 2SLS IV regressions instrumenting for both Euro Share and GDP per capita
- (3) Robustness checks

This research uses cross-sectional data as opposed to a panel data to be more comparable with prior research looking at the long-term effects of colonial origin. The use of cross-sectional data

allows to take a 2SLS approach to dealing with endogeneity problems in a similar way it has been done by Angeles & Neanidis (2015). Furthermore, it is not essential to look at variations over time for answering the hypotheses. However, to ensure that the measured results do not exclusively hold in the year 2010 all regressions are rerun in for years 1995, 2000 and 2005. The regressions will also be rerun using an alternative proxy for Human Capital to check for robustness across Human Capital indicators (Stijns, 2006).

3.4 Summary Statistics & Variable Definitions

Table 1 shows the summary statistics of the main variables used in this research. Columns 1-6 respectively display the number of observations, the mean value, the standard deviation, the minimum value, the maximum value, and the median value of each variable. Most notable is the extremely low mean value of 0.07 for the Euro Share variable considering its maximum is 0.91. The median of 0.00 indicates that this is partially because the dataset contains a great deal of countries that have a 0 Euro Share. On the other hand, a few settler colonies like Canada and the United States have an extremely high euro share. By also running regressions with a sample including only countries with a Euro Share below 12.5% insight can be gained into whether these high Euro Share countries heavily influence the results. Furthermore, since countries that were never colonized are excluded from the regressions, many countries with a 0 Euro Share are not included in the regressions.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Obs.	Mean	Std. Dev.	Min	Max	Median
Euro Share	129	0.07	0.166	0	0.91	0.00
Current Income	122	8.92	1.21	6.50	11.79	8.99
Years of Schooling	83	7.23	2.51	1.88	12.32	7.57
Life Expectancy	125	67.97	9.04	45.1	82.98	70.06
Urbanization	127	53.86	24.43	10.64	100	50.96
Malaria Ecology (IV)	115	5.09	7.27	0	31.55	1.42
Indigenous Mortality (IV)	130	0.34	0.48	0	1	0

Table 2 shows an overview of the main variables used in the empirical analysis. The first column presents which variable is being defined. The second column defines the specific indicator that was used to measure each variable. Finally, the third column shows from where the data on each variable has been retrieved. The following paragraphs will specify why each of the variables and indicators have been added.

Table 2: Variable Definitions

Definition		Source
Explanatory Variable		
Human Capital	Average years of schooling in 2010	Barro & Lee (2013)
Euro Share	Proportions of Europeans in colonial population.	Easterly & Levine (2016)
Current Income	Ln of GDP per capita in 2010 (PPP, current international \$)	World Bank World Development Indicators (2022a)
Population Density	Urban population in % of total population	World Bank World Development Indicators (2022b)
Malaria Ecology (IV)	An index of the stability of malaria transmission based biological characteristics of mosquitoes such the proportion of blood meals taken from human host, daily survival of the mosquito, and duration of the transmission season and of extrinsic incubation.	Kiszewski et al. (2004)
Indigenous Mortality (IV)	Dummy variable reflecting high rates of indigenous mortality from European diseases.	Easterly & Levine, (2016).
Human Capital (Alternative)	Life expectancy at birth in 2010	World Bank World Development Indicators (2022c)

Notes

Overlapping variables are defined as shown in Easterly & Levine's Variable Definition Table (2016, p.30).

3.5 Dependent Variable: Human Capital

Human capital measurement by previous researchers can be largely categorized into the cost-based approach, the income-based approach, and the outcome-based (or education-based) approach (Le, 2005; Kwon, 2009).

The income-based approach is based on the returns of education investment and tries to measure human capital by looking at the total discounted values of all future income streams on the individual level. However, since there are a great number of other contributors to individual income other approaches are likely to be more accurate (Le et al., 2005; Kwon, 2009).

In the cost-based approach researchers like Kendrick (1976) and Eisner (1985) measured human capital by summing the investments made by individuals to acquire human capital.

Various ways have been proposed to measure human capital via the cost-based approach and the income-based approach. However, due to the nature of this research I prefer taking an outcome-based approach. This is mainly because the studies examining the relation between colonial history and economic development primarily refer to human capital outcomes in their analyses (Banerjee & Iyer, 2005; Dell, 2010; Easterly & Levine, 2016; Dell & Olken, 2020).

The outcome-based approach measures human capital by looking primarily at education outcomes, but also sometimes at health outcomes (Le et al., 2005; Mousavi & Clark, 2021). Within the outcome-based approach various indicators have been used as a proxy for human capital. Popular examples are school enrollment rates, educational attainments, adult literacy rates and average years of schooling (Kwon, 2009). Which outcome is used to proxy for human capital is usually very dependent on data availability which has historically not been great in developing countries (Le, et al., 2005).

Where economists used to depend on enrolment ratios and adult literacy rates various economists have suggested the use of average years of schooling of the labor force to be a more accurate outcome to use as a proxy for human capital (Psacharopoulos & Arriagada, 1986; Barro & Lee, 1993, 1996, 2013; Wachtel, 1997). For the computation of average years of schooling three approaches have been used: The census/survey-based estimation method, the projection method, and the perpetual inventory method (Le, 2005). Currently the most recent and most widely used measure of average years of schooling is from Barro and Lee (1993, 1996, 2013) who

build forth upon earlier work from Psachoropoulos & Arragada (1986, 1992) and use a combination of all three approaches. Some researchers have critiqued Barro & Lee's (1993, 1996, 2013) computation (De la Fuente & Domenech, 2000; Cohen & Soto, 2001). However, the most recent data computed by Barro and Lee (2013) seems to be the best data currently available and most widely used outcome-based proxy for human capital.

This research considers five Human Capital indicators based on a study conducted by Stijns (2006) on the relation between resource dependence and human capital. The following five indicators were considered: Literacy rate, net secondary education enrolment, public education expenditure, life Expectancy and average years of schooling. Literacy rate and net secondary education enrolment significantly reduced the sample size and public education expenditure creates severe statistical issues due to the inclusion of an income per capita component in the model. Average years of schooling and life expectancy have 83 and 125 observations respectively (See table 1). Although life expectancy has more observations than average years of schooling, the main dependent variable used in this research is average years of schooling. This is because the available instrumental variables which are required for the 2SLS regressions fit average years of schooling better. This will be elaborated on in paragraph 3.9. In paragraph 4.3, all regressions will also be run using life expectancy as a dependent variable to check whether this significantly changes the results.

3.6 Explanatory Variable: Euro Share

This research utilizes a relatively new Euro Share indicator constructed by Easterly & Levine (2016) that measures the share of European settlers in a country during the beginning of their colonial period. Euro Share is an index between 0 and 1 where 0 represents a complete absence of Europeans in the colonial population and where 1 represents a colonial population based entirely out of Europeans. Easterly & Levine (2016) computed the Euro Share variable based on a collection of 46 secondary sources combined with primary data. Based on these sources they determined individual measuring years for each country based on their colonial history. Euro Share is likely to be endogenous to human capital due to omitted variable bias. The endogeneity issues surrounding Euro Share will be further elaborated on in paragraph 3.9.

3.7 Control Variables

As previously mentioned in the first paragraph of this chapter sample size plays a central role in deciding which control variables to include in the model. General knowledge of the determinants of Human Capital dictates that current income and population density are important control variables to include in any regression model on Human Capital (Mousavi & Clark, 2021). Ideally one would also want to control for institutional quality, however, as noted by Acemoglu et al. (2014) this would not make sense because if European settlement would directly affect human capital, it is likely to still also have affected institutional quality directly and potentially through human capital itself. The addition of institutional quality as a control variable would therefore make it very difficult to interpret the results. It would also come with extra endogeneity issues that would be difficult to account for, especially when also treating for endogeneity in current income and European colonial settlement.

For current income the natural logarithm of GDP per capita is used measured in current international dollars based on purchasing power parity. This variable is logged to gain a normal distribution.

The % of the current population that is populated in urban areas is used as a proxy for population density. The main motivation behind favouring urbanization above population density is that the use of urbanization allows for a slightly bigger sample size.

In the OLS regressions *Life Expectancy* is also added to check whether its addition changes the results.

3.8 Endogeneity Issues

The primarily statistical issue with doing empirical research on this topic is the issue of endogeneity. Both Euro Share and GDP per capita are endogenous to years of schooling but for different reasons. In the case of Euro Share the endogeneity issues are due to omitted variable bias as there are expected to be many factors that account for differences across countries in years of schooling, among which cultural differences. There is also the issue of Euro Share influencing institutions, which in turn also influence human capital. GDP per capita is likely to be endogenous due to simultaneity, meaning that besides GDP per capita influencing human capital,

human capital is also likely to influence GDP per capita. Countries that have higher levels of income can afford to invest more in education while countries that are more educated will be more likely to have increased opportunities to generate higher income by being more productive. Some studies like Easterly & Levine (2016) or Glaeser et al. (2004) have ran similar regressions to the one in this research while leaving endogeneity untreated. Others, like Acemoglu et al. (2001) have tried to use instrumental variables to treat for endogeneity.

Both approaches have advantages and disadvantages. The advantage of treating Euro Share and GDP per capita as exogenous is that it is simple and that there is no decreased sample size when the observations of the instrumental variables do not perfectly match the sample. The downside is that it is impossible to tell how the endogeneity is influencing the results as endogeneity bias could be both positive and negative in the case of Euro Share. The advantage for using a 2SLS instrumental variable approach to treat for Euro Share and GDP per capita is that the risk for bias will be highly decreased if the instruments are valid and strong. However, the available instruments are far from perfect, and it is difficult to say with certainty that endogeneity will be properly treated for. Therefore, the empirical part of this research is conducted in two parts, one while treating all variables as exogenous, and one in which endogeneity of both Euro Share and GDP per capita is treated for. In the next paragraph the Instrumental Variables used are discussed.

3.9 Instrumental Variables

In the second part of the empirical analysis Euro Share and GDP per capita are both replaced by instrumental variables. The number of potential instrumental variables is extremely limited due to data availability and a lack of prior research that regress colonial origin on human capital. Angeles & Neanidis (2015) used latitude in absolute value, Population Density in the year 1500, the degree of malaria prevalence, the percentage of a country's territory within 100 kilometres to the sea, and Easterly & Levine's (2016) Indigenous Mortality dummy variable as instrumental variables in their research. For this research only the variables that were included in Easterly & Levine's (2016) dataset were considered because variables taken from other datasets would further reduce the sample size. The instrumental variables for which data was available for most

of the sample size were Malaria Ecology, Population Density in the year 1500 and Indigenous Mortality because Easterly & Levine (2016) used them as control variables. Population Density in the year 1500 causes issues due to its close relation to current population density, which is already in the model as a control variable. Fortunately, Malaria Ecology and Indigenous Mortality meet the two requirements for good instrumental variables, relevance and exogeneity (Angeles & Neanidis, 2015).

Malaria Ecology is used as an instrumental variable for *GDP per capita*. Malaria transmission stability is expected to be negatively correlated to Current Income because malaria directly influences the ability to work and comes with great costs for society in terms of treatment and prevention. It is less likely to be directly correlated to *Average Years of Schooling* as it is a temporary disease that should not directly affect people's long-term educational abilities.

Indigenous Mortality is used as an instrumental variable for *Euro Share*. It is expected to be positively correlated to *Euro Share* because it weakened indigenous resistance and created space for settlers to live as it reduced the indigenous population. It is not expected to be correlated to *Average Years of Schooling* because most European diseases that contributed to indigenous mortality during colonial times are now no longer prevalent or are now easily treatable. Indigenous Mortality is preferred over Settler Mortality, which was used as an instrumental variable by AJR (2001) because it allows for a greater sample size. Indigenous Mortality is a variable computed by Easterly & Levine (2016) and is therefore computed for approximately the same countries as *Euro Share*.

4 Results

In this chapter the results of the empirical analysis are shown, interpreted, and discussed. In paragraph 4.1 the OLS regressions, in which all variables are treated as exogenous, are discussed. In paragraph 4.2 the results of the 2SLS regressions, in which Instrumental Variables are used to treat for the two endogenous variables are shown, interpreted, and discussed. Finally, in paragraph 4.3 and 4.4 the results of the robustness checks are shown, interpreted, and discussed.

4.1 Ordinary Least Squares regressions (OLS)

In the first part of the empirical analysis regressions are run under the assumption that Euro Share and Current Income are exogenous determinants of Human Capital. The results of the OLS regressions shown in table 3 should show unbiased results under this assumption.

In column (1) of table 3 a basic bivariate relation is shown between European colonial settlement and Years of Schooling without any control variables. All countries for which data was available for both variables are included except for the few countries that were never colonized. The coefficients are highly significant at well below the 1% level. Therefore, at first glance the relationship between Euro Share and Human Capital seems very strong, just as suggested by Glaeser et al. (2004) and Easterly & Levine (2016).

In column (2) the log of GDP per capita is added as a control variable. The coefficient of Euro Share drops to about half its original size but both variables are shown to be highly significant at well below the 1% level.

In column (3) Urbanization is added as a second control variable, representing population density. The addition of Urbanization does not change the coefficients of Euro Share and Current Income which both remain significant at the 1% level. Urbanization however turns out to be highly insignificant and shows a very small negative effect, which is opposite of what would be expected. However, regardless of the lack of significance a population density variable should be added as it is a standard control variable for human capital. The negative sign for urbanization is not relevant due to the coefficients being highly insignificant. Other indicators of population density like population density itself have been explored but showed similar issues with significance. Urbanization is included in all regressions in favor of other indicators of population density due to data availability.

Column (4) alternatively measures the relationship between Euro Share and Human Capital with the addition of a Health control variable in the form of Life Expectancy (Mousavi & Clark, 2021). The results however turn out not to be significant and since Life Expectancy is also used as a dependent variable in the Robustness Checks it is being left out of the equation.

Column (5) shows the same OLS regression as shown in column (3) but only includes countries with a Euro Share of below 12.5%. The coefficients of Current Income and Urbanization remain

around a similar level while the coefficients of Euro Share increase massively while remaining significant at a 1% level. These results support the acceptance of hypothesis 2 shown in paragraph 2.5. The over 4 times increase in the coefficient of Euro Share when excluding countries with a Euro Share above 12.5% is in line with the findings by Acemoglu et al. (2014) that European Settlers brought more initial human capital with them to settler colonies. It also suggests that the long-term positive effects from European Settlement were far higher per settler in extractive colonies than in settler colonies. The addition of column (5) also allows to look whether the results shown in column (3) hold when the highly developed new world countries like the US and Canada are excluded. The results shown in table 3 support both hypotheses shown in paragraph 2.5 and contain some evidence toward the human capital view, in line with Glaeser et al. (2004) and Easterly & Levine (2016).

Table 3: OLS regression between Euro Share and Human Capital

	(1)	(2)	(3)	(4)	(5)
	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling
Euro Share	8.022*** (1.077)	3.946*** (0.905)	3.945*** (0.910)	3.709*** (0.867)	16.064** (6.996)
Current Income (ln)		1.396*** (0.202)	1.431*** (0.303)	1.249*** (0.296)	1.370*** (0.328)
Urbanization (%)			-0.002 (0.014)		-0.002 (0.016)
Life Expectancy				0.026 (0.034)	
Countries	73	71	71	71	64
R-squared	0.231	0.611	0.611	0.614	0.553
Prob>F	0.00	0.00	0.00	0.00	0.00
F test:	55.44	136.17	90.22	74.56	26.86

Notes

Robust standard errors in parentheses

, ** and * denote significance at the 10%, 5% and 1% level respectively*

4.2 Two-Stage Least Squares regressions (2SLS)

In this paragraph the same empirical model as in the OLS regressions is used, but this time treating for endogeneity of both Euro Share and Current Income. Table 4 show results treating for endogeneity using a 2SLS using instrumental variables to replace the endogenous variables. Panel A of table 4 shows the second stage results where Euro Share and Current Income are already replaced by its corresponding instrumental variables. Panel B and Panel C show the first stage results for Current Income and Euro Share respectively. Column (2) and (4) of table 4 show results of similar regressions as shown in column (1) and (3) but using Life Expectancy as dependent variables. These columns will be discussed in paragraph 4.3.

The first stage for Current Income shown in panel B of table 4 regresses Malaria Ecology, Indigenous Mortality and Urbanization against Current Income. Malaria Ecology is the intended Instrumental Variable to be used for Current Income. The Malaria Ecology coefficients are negative which is in line with expectations as malaria prevalence is expected to limit economic development due to the costs that come with illness and treatment. The results are also significant at the 1% level which means that Malaria Ecology adheres to both requirements to be used as an instrumental variable for Current Income.

The first stage for Euro Share shown in panel C of table 4 regress Indigenous Mortality, Malaria Ecology and Urbanization against Euro Share. Indigenous Mortality is the intended Instrumental Variable to be used for Euro Share. The coefficient of Indigenous Mortality is positive which is in line with the expectation that Europeans could more easily settle in colonies with high indigenous mortality due to increased land availability and decreased indigenous resistance to settler policies. Indigenous Mortality is significant at the 1% level and therefore it adheres to both requirements to be used as an instrumental variable for Euro Share.

In the second stage shown in Panel A of table 4, column (1) shows the relation between Euro Share and Years of Schooling using the main sample including all countries except countries that were never colonized. The coefficient of Euro Share is positive and significant at the 5% level which is slightly less significant than in the OLS regressions but still significant enough. Current income has a positive coefficient and is significant at the 1% level. Urbanization still shows no significance and has a slightly negative coefficient which is unexpected but insignificant.

Column (3) of the second stage results shown in Panel A of table 4 run the same regressions as in column (1) but using a sample that only includes countries with a Euro Share of below 12.5%. The coefficient of Euro Share is also significant at the 5% level. The coefficients of Current Income and Urbanization are similar as in column (1) except Current Income is only significant at the 5% level in column (3). Although the coefficients cannot be directly interpreted in this type of regression the tripling in Euro Share's coefficient shows a similar pattern as can be seen when comparing the Euro Share coefficients of column (3) and (5) in table 3. The results in table 3 therefore accept both hypotheses formed in paragraph 2.5 and show that the results from the OLS regressions still hold when treating for endogeneity.

Table 4: 2sls IV regressions between Euro Share and Human Capital

	(1)	(2)	(3)	(4)
	Years of Schooling	Life Expectancy	Years of Schooling	Life Expectancy
Panel A: Two-Stage Least Squares				
Euro Share	11.693** (5.395)	33.072** (13.936)	34.223** (13.398)	128.886*** (42.443)
Current Income (ln)	2.374*** (0.860)	11.156*** (2.231)	1.910** (0.871)	9.421*** (2.388)
Urbanization (%)	-0.055 (0.034)	-0.301*** (0.099)	-0.028 (0.327)	-0.198** (0.100)
Countries	69	95	63	86
R-squared	0.325	0.150	0.465	0.334
Prob>F	0.000	0.000	0.000	0.000
F-Test	12.30	24.57	20.19	41.39
Panel B: First Stage for Current Income				
Malaria Ecology	-0.043*** (0.102)	-0.391*** (0.008)	-0.430*** (0.010)	-0.039*** (0.008)
Indigenous Mortality	-0.173 (0.196)	-0.141 (0.175)	-0.215 (0.211)	-0.216 (0.194)
Urbanization (%)	0.038*** (0.004)	0.039*** (0.003)	0.038*** (0.004)	0.382*** (0.003)
Panel C: First Stage for Euro Share				
Malaria Ecology	-0.000 (0.000)	-0.000 (0.000)	-0.001** (0.000)	-0.001** (0.000)
Indigenous Mortality	0.105*** (0.377)	0.145*** (0.041)	0.030*** (0.009)	0.037*** (0.007)
Urbanization (%)	0.001** (0.001)	0.001** (0.001)	0.000* (0.000)	0.000 (0.000)

Notes

Robust standard errors in parentheses

, ** and * denote significance at the 10%, 5% and 1% level respectively*

4.3 Robustness Checks: Repeating regressions with an alternative Human Capital Measure

If the results in paragraphs 4.1 and 4.2 truly show a positive effect of European colonial settlement on human capital one would expect the results to hold amongst various measures of human capital. Unfortunately, most measures of human capital would reduce the sample size to a point where there would be a very high risk of the regression sample not accurately representing the true population. The only other proper human capital measure with a large enough sample size is Life Expectancy which has a higher number of observations than years of schooling. This suggests the results shown in table 5 might have a higher chance of representing the true population than the main results shown in table 3. The 2SLS regressions using Life Expectancy as a dependent variable are shown in table 4.

The OLS regression results shown in table (5) suggest that the relationship between Euro Share and Human Capital hold even when changing the proxy used to measure Human Capital is changed from Years of Schooling to Life Expectancy. For columns (1) up until (4) the results are similar to the results shown in table 3 with the exception that Euro Share is now only significant at the 5% level in column (2) up until column (4). Years of Schooling has been added as a potential control variable in column (4) but is shown to be highly insignificant and significantly lowers the sample size. This decrease in sample size most likely explains the increase in the Euro Share coefficient in column (4). This exemplifies the importance of sample size and that any decrease in sample size increases the chance of biased results.

Most notably, column (5) of table 5 shows no significant coefficients for Euro Share when only countries with a Euro Share below 12.5% are included. This could suggest that in extractive colonies the positive effects of European Settlement are primarily on education outcomes, and not so much on health outcomes.

Interestingly the 2SLS regression results shown in column (3) and (4) of table 4 show robustness to changing the dependent variable from Years of Schooling to Life Expectancy. In the regressions including all observations Euro Share is significant at a 5% level, while Euro Share in the regression that only include countries with a small Euro Share is even significant at the 1% level. The issue with the 2SLS results using Life Expectancy as a dependent variable is however that Malaria Ecology is unlikely to meet the exogenous requirement for being a good instrument as Malaria is

a deadly disease which directly affects one's Life Expectancy. This might bias the 2SLS results and could potentially explain the difference between Euro Share significance in column (5) of table 5 and column (4) of table 4.

Table 5: OLS regression between Euro Share and Human Capital (Alternative HC measure)

	(1)	(2)	(3)	(4)	(5)
	Life	Life	Life	Life	Life
	Expectancy	Expectancy	Expectancy	Expectancy	Expectancy
Euro Share	22.650*** (3.254)	5.723** (2.502)	5.719** (2.518)	7.713** (3.154)	29.699 (25.767)
Current Income (ln)		5.608*** (0.433)	5.558*** (0.785)	5.135*** (0.765)	5.351*** (0.827)
Urbanization (%)			0.003 (0.041)		0.004 (0.047)
Years of Schooling				0.322 (0.384)	
Countries	100	97	97	71	87
R-squared	0.163	0.627	0.627	0.640	0.582
Prob>F	0.000	0.000	0.000	0.000	0.000
F test:	48.45	118.96	78.39	63.98	59.54

Notes

Robust standard errors in parentheses

, ** and * denote significance at the 10%, 5% and 1% level respectively*

4.4 Robustness Checks: Repeating regressions for 1995, 2000 and 2005

Because this research does not use panel data the regressions shown in paragraphs 4.1 up until 4.3 are all based on the same year, namely 2010. To rule out the possibility that the results shown in paragraphs 4.1 up until 4.3 are just coincidentally showing up in the year 2010 due to some

exogenous factor all OLS regressions and 2SLS regressions from paragraphs 4.1 and 4.2 are rerun for the years 1995, 2000 and 2005.

The results in table 6 suggests that the relation between Euro Share and Years of Schooling is robust for all years shown across columns (1) up until column (6). All Euro Share coefficients in the 2SLS and OLS regressions are significant at the 1% level except in Panel D column (4) where Euro Share is significant at the 5% level. However, in column (1), (3) and (5), which include all observations, show a slight sign of significance at the 10% (All between 8 and 10%) while also showing the wrong sign. A possible explanation for the odd sign for urbanization could be the fact that in the small sample available the few countries with extremely high urbanization values in countries such as Qatar in which life is only possible in cities due to most of the landscape being a desert. The addition of these countries could deflate the urbanization coefficients because these geographical characteristics are unlikely to be related to Years of Schooling. Another unexpected finding is the negative r-squared shown in column (3) and column (5). However, R-squared does not give valid information on the goodness of fit in a 2SLS regression because the instruments used are not actually correlated to Years of Schooling by themselves. The significance of the coefficients for Current Income are only significant at the 1% level in column (2), (4) and (6). These are the regressions that only show include countries with a Euro Share below 12.5%. A possible explanation for this decrease in statistical significance could be the slight drop in sample size. All OLS regressions run in Panel D of table 6 are in line with the 2010 results and therefore show robustness.

Table 6: OLS and 2SLS IV regressions between Euro Share and Human Capital 1995-2005

	1995		2000		2005	
	(1)	(2)	(3)	(4)	(5)	(6)
	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling	Years of Schooling
Panel A: Two-Stage Least Squares						
Euro Share	19.991*** (7.751)	39.545*** (15.541)	16.685*** (6.438)	43.574*** (14.199)	16.749*** (6.431)	44.277*** (14.163)
Current Income (ln)	2.444** (1.030)	1.866* (1.115)	2.439** (0.993)	1.732* (1.020)	2.530** (1.042)	1.888* (1.044)
Urbanization (%)	-0.072* (0.042)	-0.036 (0.045)	-0.077* (0.046)	-0.318 (0.047)	-0.081* (0.464)	-0.040 (0.044)
Countries	66	61	68	62	69	63
R-squared	0.184	0.328	-0.027	0.268	-0.008	0.280
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
F test:	8.93	14.04	8.03	14.73	7.71	14.14
Panel B: First Stage for Current Income						
Malaria Ecology	-0.037*** (0.010)	-0.037*** (0.010)	-0.040*** (0.010)	-0.040*** (0.010)	-0.040*** (0.010)	-0.040*** (0.011)
Indigenous Mortality	-0.0191 (0.204)	-0.199 (0.210)	-0.258 (0.194)	-0.285 (0.205)	-0.239 (0.197)	-0.264 (0.211)
Urbanization (%)	0.039*** (0.004)	0.040*** (0.004)	0.041*** (0.004)	0.041*** (0.004)	0.042*** (0.004)	0.041*** (0.004)
Panel C: First Stage for Euro Share						
Malaria Ecology	-0.008 (0.001)	-0.001* (0.000)	-0.001 (0.001)	-0.001** (0.000)	-0.001 (0.001)	0.001** (0.000)
Indigenous Mortality	0.074*** (0.026)	0.028*** (0.009)	0.103*** (0.037)	0.030*** (0.009)	0.104*** (0.037)	0.030*** (0.009)
Urbanization (%)	0.001 (0.001)	0.000 (0.000)	0.002* (0.001)	0.000 (0.00)	0.001** (0.001)	0.000 (0.000)

Panel D: OLS Regression

Euro Share	7.417*** (1.215)	16.511*** (6.002)	4.900*** (1.245)	14.872** (6.599)	4.900*** (1.245)	16.592** (7.138)
Current Income (ln)	1.248*** (0.332)	1.204*** (0.363)	1.187*** (0.353)	1.179*** (0.395)	1.187*** (0.353)	1.233*** (0.377)
Urbanization (%)	-0.000 (0.016)	-0.001 (0.019)	0.003 (0.018)	-0.000 (0.022)	0.003 (0.018)	-0.004 (0.020)
Countries	68	62	70	63	70	64
R-squared	0.591	0.494	0.565	0.471	0.565	0.473
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
F-Test	69.99	26.59	39.25	21.98	39.25	20.65

Notes

Robust standard errors in parentheses

, ** and * denote significance at the 15%, 10%, 5% and 1% level respectively*

5 Conclusion

The findings of this research are in line with the view that colonial European settlement has had positive effects on current human capital, even if the share of European settlers was relatively small. Both the OLS regressions and the 2SLS IV regressions show highly significant results that are robust to changing the human capital indicator and the results hold across 4 time-measurements over 15 years at a 5-year interval. This means that the result support both hypotheses formulated in paragraph 2.5.

Interestingly, the results give further insight into Acemoglu et al.'s (2014) argument that the initial effect of European settlers on human capital was far greater in countries with low settlement as the results show that the effect on human capital today is also consistently 2-3 times larger when only including countries with a low share of European Settlement. This goes against the theory that the measured increase in human capital is a result of settler institutions because according to this theory settlers would have set up weak extractive institutions in most of the countries within this sample. These findings are in line with Easterly & Levine (2016) while accounting for the most recent critique Acemoglu et al. (2014) have risen regarding endogeneity. The 2SLS results of this research also tackle the critique raised by Acemoglu et al. (2014) on the

imperfection of Easterly & Levine's (2016) Euro Share measure by using Indigenous Mortality as an instrumental variable for Euro Share.

The intention of this research is not to argue that institutions do not matter for economic development or that all effects of European colonial settlement on current economic development have come through human capital. Instead, this research aims to highlight that positive long-term effects can be measured of European colonial settlement on human capital, even amongst so-called extractive colonies. This finding on its own should be enough cause to further investigate what has driven these positive human capital effects, preferably in the form of within-country research as economists like Dell & Olken (2020) have already started doing.

It is safe to assume that most people agree that European colonialism was unethical regardless of whether it had positive economic effects. Also, it seems unlikely that any number of positive effects on current development could offset the atrocities that came with European colonization like the enslavement of people in west-Africa or the genocide of native Americans. However, if colonialism has also had positive effects on current human capital, further investigation into the processes that achieved this might be able to help policy makers improve the effectiveness of policies aimed at improving human capital. This could be helpful even if human capital would not be the main driver behind economic development as human capital itself can be a development goal, particularly when policy makers choose to focus on economic well-being in a broad sense.

But within-country studies of each specific colonies and the channels through which institutions and human capital have been affected by European colonial settlement is needed to be able to learn from colonial history for improving policy making. The reason for this is that the observed countries in the sample consist of multiple colonizers using various methods to extract different types of resources in various cultural settings. For example, based on the within-country research conducted by Dell & Olken (2020) that conclude that positive human capital effects came from Dutch settlers setting up sugar factories, policymakers could try to recreate similar economic activity adapted to be in line with the world we live in today. There is however unlikely that policies aimed at e.g. Java are going to work in vastly different former colonies such as Congo or Venezuela.

Future research should primarily focus on working with governments of colonizers and former colonies to create useable data for within-country research such as already been done in countries like Peru (Dell, 2010), India (Banerjee & Iyer (2005) and Indonesia (Dell & Olken, 2020). If useable data became available on a greater number of former colonies researchers can examine the relation between colonial origin and economic development much more precisely. An alternative direction would be for future research to focus on improving data quality and mitigating some of the many limitations that come with looking at colonial history in a cross-country setting like Easterly & Levine (2016) tried to do by computing their Euro Share variable. These limitations will be further discussed in chapter 6. Another potential direction for future research would be to follow the direction of this research and Angeles & Neanidis (2015) by studying the relationship between European colonial settlement and other development outcomes. These studies could also potentially be focused on more contemporary view on well-being like health, safety, and life satisfaction (OECD, 2020). Finally, future research should be done on the possibility that the effects on human capital are not persistent but decreasing over time as Robbert Maseland (2018) suggests might be the case for institutions and economic development. When the goal would be to increase policymaking, it would be best for future economists to set up new within-country analyses.

6 Discussion/Research Limitations

In this chapter the various limitations that come with this research are discussed. They can be summed up as follows:

- The sample size is small and contains outliers
- European colonial settlement is very difficult to accurately quantify
- Number of available instrumental variables are limited and imperfect
- OLS results contain endogeneity bias
- It is impossible to distinguish institutions and human capital effects
- Model contains little controls in order not to further reduce sample size
- Sample is not homogenous. Various colonizers used various methods to extract different types of resources.

Researching the cross-country relation between colonial origin and economic development is quite tricky in terms of statistical issues. Any research on this relationship deals with a small

sample size with a few outliers in terms of economic development and European colonial settlement as most colonies have had a small European colonial settlement and are relatively underdeveloped while a few countries like the United States, Canada and New Zealand have a very high share of European settlement and are highly developed. The small sample size also limits the options of adding extra control variables or instrumental variables as decreasing the sample size further could weaken the reliability of the results. Sample size particularly hinders the possibility in terms of checking for robustness as most human capital variables and potential other control variables decreased the sample size by too much for some robustness checks to be added. There is also the issue of endogeneity when using OLS regressions and the issue of potential weak instrumental variables in 2SLS regressions. Although this research tries to tackle the issues by combining both OLS and 2SLS regressions the potential issues regarding these methods remain. The OLS regressions are likely to be biased due to endogeneity and the 2SLS regressions might not properly treat for endogeneity if the instruments are not strong enough. Particularly the use of Malaria Ecology as an instrument for current income poses as a risk if it turns out to be directly correlated to Years of Schooling because of increased school illness. The issue of not being able to properly separate human capital effects from institutional effects also makes cross-country regressions not ideal to provide proper evidence towards settling the debate amongst proponents of the human capital view and the institutions view.

7 Bibliography

- Acemoglu, D., Gallego, F., & Robinson, J. A. (2014). *Institutions, human capital and development* (No. w19933). National Bureau of Economic Research.
- Acemoglu, D., García-Jimeno, C., & Robinson, J. A. (2012). Finding Eldorado: Slavery and long-run development in Colombia. *Journal of Comparative Economics*, 40(4), 534-564.
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5), 1369-1401.

- Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1, 385-472.
- Angeles, L., & Neanidis, K. C. (2015). The persistent effect of colonialism on corruption. *Economica*, 82(326), 319-349.
- Banerjee, A., & Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American economic review*, 95(4), 1190-1213
- Barro, R. J. and Lee, J.-W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics*, 32(3):363–394.
- Barro, R. J. and Lee, J.-W. (1996). International measures of schooling years and schooling quality. *American Economic Review*, 86(2):218–223.
- Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104, 184-198.
- Bruhn, M., & Gallego, F. A. (2012). Good, bad, and ugly colonial activities: do they matter for economic development? *Review of economics and statistics*, 94(2), 433-461.
- Chang, H. J. (2011). Institutions and economic development: theory, policy and history. *Journal of institutional economics*, 7(4), 473-498.
- Cohen, D. and Soto, M. (2001). Growth and human capital: Good data, good results. Development Centre Technical Papers 179, OECD.
- Djankov, S., Glaeser, E., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2003). The new comparative economics. *Journal of comparative economics*, 31(4), 595-619.
- De la Fuente, A. and Domenech, R. (2000). Human capital in growth regressions: How much difference does data quality make? Working Paper 262, OECD.

- Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6), 1863-1903.
- Dell, M., & Olken, B. A. (2020). The development effects of the extractive colonial economy: The dutch cultivation system in java. *The Review of Economic Studies*, 87(1), 164-203.
- Easterly, W., & Levine, R. (2016). The European origins of economic development. *Journal of Economic Growth*, 21(3), 225-257.
- Eisner, R. (1985). The total incomes system of accounts. *Survey of Current Business*, 65(1):24– 48.
- Engerman, S. L., & Sokoloff, K. L. (1994). *Factor Endowments: Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States* (Working Paper No. 66; Historical Working Paper Series). National Bureau of Economic Research.
- Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do institutions cause growth? *Journal of economic Growth*, 9(3), 271-303.
- Hass, J. L., Brunvoll, F., & Hoie, H. (2002). Overview of sustainable development indicators used by national and international agencies.
- Huillery, E. (2009). History matters: The long-term impact of colonial public investments in French West Africa. *American economic journal: applied economics*, 1(2), 176-215.
- Iyer, L. (2010). Direct versus indirect colonial rule in India: Long-term consequences. *The Review of Economics and Statistics*, 92(4), 693-713.
- Kendrick, J. (1976). *The Formation and Stocks of Total Capital*. Columbia University Press for NBER, New York, N.Y.

- Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. (2004). A global index representing the stability of malaria transmission. *The American journal of tropical medicine and hygiene*, 70(5), 486-498.
- Kwon, D. B. (2009, October). Human capital and its measurement. In *The 3rd OECD world forum on "statistics, knowledge and policy" charting progress, building visions, improving life* (pp. 27-30).
- La Porta, R., Lopez-De-Silanes, F., Shleifer, A., & Vishny, R. W. (1997). Legal Determinants of External Finance. *The Journal of Finance*, 52(3), 1131–1150.
- Le, T., Gibson, J., & Oxley, L. (2005). Measures of human capital: A review of the literature.
- Lipset, Seymour M. 1960. *Political Man: The Social Basis of Modern Politics*. New York: Doubleday.
- Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1), 3-42.
- Maseland, R. (2018). Is colonialism history? The declining impact of colonial legacies on African institutional and economic development. *Journal of Institutional Economics*, 14(2), 259–287.
<https://doi.org/10.1017/S1744137417000315>
- Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, 81(1), 113-152.
- Mousavi, A., & Clark, J. (2021). The effects of natural resources on human capital accumulation: A literature survey. *Journal of Economic Surveys*, 35(4), 1073-1117.
- Naritomi, J., Soares, R. R., & Assunção, J. J. (2012). Institutional development and colonial heritage within Brazil. *The journal of economic history*, 72(2), 393-422.
- North, D., Thomas, R. (1973) *The Rise of the Western World: A New Economic History*, Cambridge University Press, Cambridge UK.

- Nunn, N. (2009). The Importance of History for Economic Development. *Annual Review of Economics*, 1(1), 65–92.
- Nunn, N. (2020, March 27). *The historical roots of economic development*.
- OECD (2020), *How's Life? 2020: Measuring Well-being*, OECD Publishing, Paris.
- Peck, J. (2016). Economic rationality meets celebrity urbanology: exploring Edward Glaeser's city. *International Journal of Urban and Regional Research*, 40(1), 1-30.
- Psacharopoulos, G. and Arriagada, A. M. (1986). The educational composition of the labour force: An international comparison. *International Labour Review*, 125(5):561–574.
- Rodney, W. (1972). *How Europe underdeveloped Africa*. Verso Books.
- Sen, A. K. (2009). *The idea of justice*. Harvard University Press.
- Stijns, J. P. (2006). Natural resource abundance and human capital accumulation. *World development*, 34(6), 1060-1083.
- Sun, H. P., Sun, W. F., Geng, Y., & Kong, Y. S. (2018). Natural resource dependence, public education investment, and human capital accumulation. *Petroleum Science*, 15(3), 657–665.
- The World Bank (2022a). “Gross Domestic Product per capita in 2010 (PPP, current international \$)”. Retrieved from <https://databank.worldbank.org/source/world-development-indicators>, accessed on 10-08-2022.
- The World Bank (2022b). “Urban population (% of total population)”. Retrieved from <https://databank.worldbank.org/source/world-development-indicators>, accessed on 10-08-2022.

The World Bank (2022c). "Life expectancy at birth, total (years)". Retrieved from <https://databank.worldbank.org/source/world-development-indicators> ,accessed on 10-08-2022.

UNDP. (1990). *Human Development Report 1990*. Oxford University Press, New York.

Wantchekon, L., Klašnja, M., & Novta, N. (2015). Education and human capital externalities: evidence from colonial Benin. *The Quarterly Journal of Economics*, 130(2), 703-757.