Complex Relationships:
Income Inequality, Trust and Corruption

Knut Eirik Almelid

Master Thesis

Department of Comparative Politics

University of Bergen

June 2019

Abstract

The goal of this thesis is to investigate the link between income inequality and corruption and poses the following research question: Is there a positive relationship from income inequality to corruption and is this effect dependent on trust? The thesis attempts to contribute to the small but growing literature on this link by using recent data with a global coverage. In addition to the independent variables, income inequality and trust, relevant control variables were added to the model. For the analysis I use the method of Ordinary Least Squares (OLS). The results are surprising in that they reveal that income inequality has both a positive and negative correlation to corruption, depending on the level of trust. These are interesting results and contradict much of the literature. But as there are several issues concerning the quality and availability of the data, I refer from making strong conclusions. Most importantly, my findings reveal that there is a need for additional investigations into this very interesting topic.

Acknowledgments

First, I would like to thank my supervisor, Vegard L. Kolnes, for his patient and positive guidance throughout the whole process. This would not have been a thesis without his insightful supervision.

Second, I want to thank Joakim H. Ulstein for our discussions on method and his indispensable feedback and comments.

Third, I want to thank my friends and family for all their support.

Finally, I would like to thank Sigrid for all her invaluable encouragement and support.

Table of Contents

1 INTRODUCTION

Equaling an estimated 5% of global GDP (Ferguson, 2017), corruption has received an increasing amount of attention from both policy-makers and researchers (e.g., Batabyal & Chowdhury, 2015; Chetwynd, Chetwynd, & Spector, 2003; Gupta, Davoodi, & Alonso-Terme, 1998; J.-S. You & Khagram, 2005). Former UN General Secretary Ban Ki-Moon (2012) claimed that approximately 30 percent of development assistance does not reach its final destination due to corruption. According to Transparency International (2017), nearly a quarter of the world's population has paid a bribe. World Bank President Jim Yong Kim declared that corruption is "public enemy no. 1" and that "every dollar that a corrupt official or corrupt business person puts in his or her pockets is a dollar stolen from a pregnant woman who needs healthcare, or from a girl or boy who deserves an education, or from communities that need water, roads and schools" (Ferguson, 2017, p. 4).

New measurements of corruption, developed over the last two decades, provide the possibility for quantitative research and has yielded an increasing number of studies (see e.g., Rothstein & Uslaner, 2005; Uslaner, 2005; J.-S. You, 2006). Indexes, chief among them the Corruption Perception Index (CPI) by Transparency International, are invaluable tools in global comparative research into the causes and consequences of corruption. Many of these reports include media-friendly scores and rankings, popularizing their use.

According to a report commissioned by NORAD (Mungiu-Pippidi, 2011), the lack of progress in combatting corruption in the last 15 years is due to imprecise measurement instruments, conceptual flaws and inadequate strategies (Ferguson, 2017). Corruption has many forms and many causes, which do not readily provide grounds for a comprehensive universal theory or the subsequent tools and methods that would follow. The current regime of anti-corruption methods focusses in large part on institutional and legal development. Most initiatives fall into the category of anticrime initiatives, based on constraining or otherwise affecting an individual's decision making and propensity to participate in criminal acts (Ferguson, 2017; Søreide, 2014). These are, naturally, important areas in which corruption can be fought, but I do not believe that this focus is sufficient. Institutions and companies do not exist in a void, they are an inherent part

of, and result of, cultures and societies. This is acknowledged in the increase in interest on the underlying socio-political determinants of corruption.

In a relevant, but unrelated, development there has been a growing interest in inequality, especially following the translated versions of Piketty's book Capital (Piketty, 2014). Already the year before, in 2013, the World Bank "twinned its goal for ending poverty with the need for shared prosperity" (Hardoon, 2017, p. 2). Similarly, IMF's Managing Director Christine Lagarde recently spoke of the importance of reducing income inequality while newspapers like the Wall Street Journal and Forbes regularly run articles on economic inequality (Lagarde, 2018). In 2015, the leaders of 193 countries committed to reduce inequality under goal 10 of the UN's Sustainable Development Goals. In the NGO sector, Oxfam created the Commitment to Reducing Inequality Index, which measures 157 countries according to government action on social spending, tax and labor rights – three areas that they find to be important in reducing inequality (Lawson & Martin, 2018). Research on inequality has provided a list of social-ills where income inequality is the primary explanatory variable (e.g., Era Dabla-Norris, Kochhar, Suphaphiphat, Ricka, & Tsounta, 2015; Gallego, 2016; Jordahl, 2007; World Bank, 2016).

With the increasing attention that both corruption and inequality have received I would like to take a closer look at their causal relations. There is a growing literature on this relationship, but an additional variable is usually included to act as an intermediary between the two. One of the most cited researchers on the inequality to corruption link, Eric M. Uslaner, finds trust to be an integral part of what he calls the *inequality trap* (Uslaner, 2005). The inequality trap is supported by several other studies (e.g., Rothstein & Uslaner, 2005; J.-S. You & Khagram, 2005) and argues that income inequality leads to lower levels of generalized trust, leading to more corruption, which in turn leads to higher income inequality and less trust.

If the effect of inequality on corruption is mainly through an intermediary such as trust, it will be interesting to investigate this linkage as well as the literature that can shed light on the constituent causal relations. Additionally, both Rothstein and Uslaner's (2005) and You and Khagram's (2005) published their studies in 2005. Since then there have been many significant advances in the production of global indicators on several of the concepts employed in these and

other studies. It is therefore well worthwhile to re-investigate the inequality-trust-corruption link by testing it against new data.

Following this, my research question is:

Is there a positive relationship from income inequality to corruption and is this effect dependent on trust?

The research question is discussed further in section 3.4.

2 DEFINING THE DEPENDENT AND INDEPENDENT VARIABLES

In this chapter I will present the definitions of the dependent and independent variables. This is a necessary step before delving deeper into the literature in chapter 3. By defining the variables, I will more easily find relevant research and the most appropriate measurements.

2.1 DEFINING CORRUPTION

Corruption is a global phenomenon without a singular universal definition. As a concept it varies depending on cultural, political and economic factors. The UN Convention Against Corruption does not provide a full definition of corruption, instead, it describes a number of behaviors that signatories to the convention must criminalize as corruption and some behaviors that they should at least consider criminalizing (Ferguson, 2017). In a purely legal sense, "corruption is the type of behavior that a state has defined as corrupt" (Ferguson, 2017).

To overcome the lack of a universal concept and definition, it is common to use a broad, generic definition of corruption, such as the World Bank's "[corruption is] the exercise of public power for private gain" (Teorell et al., 2018) or Transparency International's "corruption is the abuse of entrusted authority for private gain" (Transparency International, 2018c). By their very nature, perceptual measures of corruption bypass the narrow legal definitions in a country and instead reflect the cultural definitions (Zhang, Cao, & Vaughn, 2009).

White-collar crime is a term closely associated with corruption. It was first used by the sociologist and criminologist Edwin Sutherland, who wrote in 1940: "crimes in business and the profession consist principally of violation of delegated or implied trust" (Zhang et al., 2009). It is interesting to note that trust plays an important part in this early definition. Although it is not always mentioned specifically, it is a constant, underlying element in any discussion on corruption and is included in Transparency International's definition of corruption as abuse of *entrusted* authority.

Corruption is loyalty to the in-group, not society as a whole, and flouts rules of fairness. Through patron-client relationships and misuse of public funds, either for private gain or rewarding those

loyal to those in power, money is transferred from the public to the elites, from the poor to the rich (Rothstein & Uslaner, 2005; Uslaner, 2005).

Corruption, such as paying a bribe, is for many about safety and not being the fool in a rotten system. It can be assurance that your doctor will have time for you, that your goods will enter the country efficiently, that you receive good grades at school and that you don't pay taxes, taxes that you rarely see beneficial results of anyways. It can be viewed as a rational response to an unresponsive and broken system. These systems incentivize individuals with power to sell it as a service. This decreases the income of the state and its ability to provide public employees with a fair salary, leading to an increasing number of public employees extorting "funds from the public purse" (Uslaner, 2005). Which leads to lower economic growth and inefficient government. In this sense, in lieu of an objective measure of the quality of government, corruption is often considered an accepted measure of the quality of government (Uslaner, 2005).

A definition of corruption should be globally understood and broad enough to include the private and public spheres. OECD's definition of corruption as "abuse of entrusted authority for illicit gain" (OECD, 2012) does this, but by including *illicit gain* it opens up to local notions of what constitutes criminal acts. This is an interesting subject in itself, but one that is beyond the scope of this study[1]. The definition employed by Transparency International (TI) is "corruption is the abuse of entrusted authority for private gain" (Transparency International, 2018c). By including *abuse of entrusted authority,* they look broadly at the behavior of those with power in all sectors of society, both public and private. Unlike with OECD's *illicit gain,* TI's *private gain* seeks to avoid purely legal interpretations. Transparency International's definition meets the above-mentioned requirements of a good definition and is the definition employed in this study.

---

[1] Criticism of modern definitions of corruption are often founded in cultural norms. Many cultures include the idea of a middle-man or contact that provides help to friends and family, in China this is called *guanzi*, in Arabic, *wasta* and in Russia, *blat.* Whether done for personal gain or seen as performing a distributive function, these personal networks exist to circumvent formal procedures and obtain goods and services in short supply (Ferguson, 2017).

2.2 DEFINING INEQUALITY

In popular culture the term *the 1%* is often used to refer to the increasing wealth of the richest 1% of the population, who since 2015 own more wealth than the rest of the world combined. Additional examples of this inequality in recent years are that the eight richest people own as much wealth as the poorest half of the world and that in the US the last 30 years have seen no real growth in incomes for the bottom 50%, while the incomes for the top 1% have grown by 300% (Hardoon, 2017).

Inequality is at its most basic a lack of equality, in rights, status, opportunities or outcomes. Inequality of outcomes concerns itself with inequalities in living conditions and economic conditions, of which inequalities in health, education, nutrition and income have received much attention in research (Afonso, LaFleur, & Alarcón, 2015).

In addition to inequality of outcomes, there is inequality of opportunity. It concerns itself with many of the same issues but focuses on what the individual cannot control. Individuals cannot choose what ethnicity or gender they are born as or what socioeconomic strata they are born into. Rothstein and Uslaner employ a narrow definition of equality of opportunity based on whether governments ensue public policies that are "intended to create equal conditions for citizens regardless of their income, ethnic/religious background, sex, and race in areas such as health care, education, and social security and legal protection" (Rothstein & Uslaner, 2005, p. 42). This definition focuses on individuals not only having access to healthcare and education but whether there are policies actively seeking to create equality of opportunity, recognizing that this requires a fair starting point for all.

In this thesis, inequality denotes income inequality. Income is the "revenue streams from wages, salaries, interest on a savings account, dividends from shares of stock, rent, and profits from selling something for more than you paid for it" (Inequality.org, 2018). Income inequality refers to the uneven distribution of income among individuals in a group, a population, a country or globally. The merits and limitations of the different measurements of income inequality will be discussed in section 5.2.2.

2.3 DEFINING TRUST

The relevance of trust is integral to the definition of corruption employed in this thesis: *the abuse of entrusted power for private gain*. Although trust a popular topic of research, there is currently no general theory or universal definition of trust (Delhey & Newton, 2005). Despite this, we can assume certain general characteristics. Rachel Botsman defines trust as a "confident relationship with the unknown" (Botsman, 2017). This definition is supported by Uslaner (Uslaner, 2013b) who views the unknown as risk. According to him, trust presupposes risk - and gives risk the benefit of the doubt. This can be done consciously through management of risk (strategic trust) or by downplaying dangers through what he calls generalized trust, what some call social trust or moralistic trust. Dangers are not dismissed, but evidence is interpreted in a more positive light - what the German sociologist Niklas Luhmann called "confidence in one's expectations" (Botsman, 2017). Being able to trust is an important tool to navigate the world. This is especially true in a world where urbanization and globalization combine to contribute to an increasing complexity and vulnerability. Through trust, people can downplay the fact that they have little to no control over what happens in the world around them. The complexity of the world and the uncertainty that follows are both reduced through generalized trust. This is a term that concerns the out-group, those beyond the radius of social interactions in an individual's life. When we expect someone to do something there is always the possibility for them to not do that thing, when we trust someone, we have confidence in our expectations that they will. This risk is inherent in any definition of trust. Trust enables us to believe that we are in control of our environment and that we can make it better. As Kenneth Newton writes:

> *"We define trust as the belief that others will not deliberately or knowingly do us harm, if they can avoid it, and will look after our interests, if this is possible. This is consistent with a common sense idea of what trust is—that we trust others when we feel we can walk the streets without being mugged, will not be treated unjustly by officials, exploited at work, deceived by politicians, treated badly by friends, acquaintances, or strangers, or cheated in everyday life"* (Newton, 2007, p. 3).

Generalized trust is closely associated with terms such as mutuality, solidarity, respect, civility and reciprocity. Although not synonyms or definitions, they help delineate a concept. Further delineating a working definition of generalized trust is its opposite: particularized trust (Rothstein & Uslaner, 2005, p. 45). Particularized trust denotes a trust reserved for one's own

specific group in society, defined by its boundaries to other groups. This group can be small, only including family and friends, or large, including those belonging to one's own religion, ethnicity, class or other social sub-groups. In societies with little generalized trust, particularized trust plays an important role[2].

We cannot know the intention of others, least of all those we do not know, yet generalized trust is the belief that we should trust others because they are part of our wider moral community (Uslaner, 2013b). The discussion on measurements of generalized trust is located in section 5.2.3.

---

[2] A statement summarizing this sentiment well might be akin to something I heard several times, from different people, on a visit to Russia: "*90 percent of my fellow citizens are stupid and untrustworthy; I only trust my friends and family*".

3 LITERATURE REVIEW AND RESEARCH QUESTION

Thematically this thesis falls under several broad literatures: the literature on inequality, the literature on corruption, and the literature investigating the relationship between these two variables and trust. The latter relationship deals with the inequality trap and is of particular interest to this thesis. However, before getting there I will situate this study in the broader contexts of the literature on corruption and inequality.

3.1 CORRUPTION

In very broad strokes, the literature on corruption can be subdivided into causes and consequences of corruption. In other words, research that treats corruption as a dependent variable and research that treats it as an independent variable. This dissertation falls into the former category, but a short review of the consequences of corruption is warranted.

3.1.1 Consequences of Corruption

As mentioned in the Introduction, the negative consequences of corruption are many, both for a country's economic growth and measures on human development. By looking at the World Value Survey from 1981, 1990 and 1995-97, Eric M. Uslaner (2005) found that respondents in countries where corruption was believed to be widespread were more likely to believe that luck and connections were more important than hard work, less likely to believe in a bright future and more likely to buy stolen goods, cheat on taxes or take bribes. The main detrimental effects of corruption, according to Gerry Ferguson in his book *Global Corruption: Law, Theory and Practice* (2017) are that it; increases the cost of doing business; leads to waste or the inefficient use of public resources; excludes poor people from public services and perpetuates poverty; corrodes public trust; undermines the rule of law, ultimately delegitimizing the state. More broadly, he adds that corruption has negative effects on human rights, gender equality, global security and climate change and environmental degradation (Ferguson, 2017, p. 7).

Many studies have examined corruption in relation to its many adverse effects, with both economic development and governance indicators receiving attention. Gupta, Davoodi and Alonso-Terme (1998) argue that corruption increases poverty and income inequality and therefore "policies that reduce corruption will also lower income inequality and poverty".

In his 2008 book *The Poverty of Corrupt Nations*, Roy Cullen writes that "Nations where corruption is rampant also tend to have a large proportion of the population living in poverty […] while the countries' leaders may be diverting millions from national wealth to Swiss bank accounts for their personal benefit" (Ferguson, 2017) The book addresses the link between corruption and poverty, which is strong, but as many others have he only mentions inequality as something that is, along with poverty, exacerbated by corruption. In order to understand corruption, the role of inequality must be understood, and not confused with poverty as elements of the same phenomena (e.g., Ünver & Koyuncu, 2016).

3.1.2   Causes of Corruption

Research on the causes of corruption has provided many possible causes with varying degrees of explanatory power and significance. The causal directions, indirect effects and causal relations of these variables are mapped in a steadily increasing number of studies. Yet isolating variables that cause corruption is challenging because many are themselves affected by corruption.

In the broader corruption research, there is a strong focus on institutional causes. When corruption is considered a sign of bad government, it makes sense to think that improving the institutions will improve the situation. A more responsive and transparent government, the thinking goes, works against corruption by design. Therefore, structural reform should cause levels of corruption to decrease (see e.g., Graf Lambsdorff, 2005; Rose-Ackerman, 2004; Søreide, 2014). But institutional quality or type alone cannot explain the level of corruption. Uslaner (2005) lists four challenges that institutional explanations face. 1) Internal consistency, in that both centralized and decentralized power can facilitate corruption. 2) Institutions are products of their cultures. Nearly 60 percent of countries with Presidential systems with closed party lists are in South America or in former communist countries, possibly reflecting the low levels of trust. 3) A long-term fight against corruption requires an institutional culture that can continue even when the media loses interest and politicians actively work against it. 4) Corruption has a *persistent stickiness* over time to it that institutional explanations don't explain.

As mentioned above, research on corruption has revealed many correlates of corruption. In this paper, it is the potential link between inequality and corruption that is of interest. Yet any paper dealing with the causes or consequences of corruption should include a short review of the literature on corruption and economic development[3]. This is especially true for a paper focusing on inequality, a term that is closely related to, and sometimes confused with, economic development.

There is no agreement on the main direction of causality, but there is general agreement in the literature reviewed that the strongest and most consistent finding of research in the last two decades is that higher economic development correlates closely with lower perceived corruption (Chetwynd et al., 2003; Treisman, 2007, p. 223). In Ünver and Koyuncu's analysis of 154 countries with data from the period 2000 to 2013, they found that "corruption is strongly and positively influenced by poverty" (Ünver & Koyuncu, 2016). The strength of the correlation holds true even for levels of GDP per capita in 1820 and 1900 (Treisman, 2007). Treisman also finds that the correlation survives the inclusion of control variables for democracy, trade, inequality, ethnolinguistic fractionalization, latitude, religion, culture, inflation, various policy variables and region. Only the variables for democracy and press freedom attenuated the effect of GDP per capita on corruption to a significant degree.

Yet the answer to why there exists a relation does not have a definitive answer or consensus. Despite the many studies showing a strong correlation between GDP per capita and corruption, the effect should be understood as being mainly indirect[4]. For corruption to have a direct effect

---

[3] This relationship has been attempted modeled before, most famously in the Kuznets curve. But the Kuznets curve is not included because of Piketty's (Piketty, 2006), Acemoglu's (Acemoglu & Robinson, 2002) and Kuznets' own criticism of its use. As Kusnets himself wrote: "This is perhaps 5 per cent empirical information and 95 per cent speculation, some of it possibly tainted by wishful thinking" (Kuznets, 1955).

[4] Ünver and Koyuncu (2016) find that countries with a higher poverty level experience higher levels of corruption, this corresponds to findings in other studies. However, their use of measurements for the poverty variable is not common in the literature reviewed for this thesis. To represent poverty, they include five variables: the UN's Human Development Index (HDI), which is a composite index of several variables; the WB's GINI coefficient, which is a measurement of income inequality; the WB's mean log deviation (MLD), which is another measure of income inequality; the World Bank's Headcount, which measures the percentage of a population living below the poverty line; and the WB's Watts' poverty index. By including a composite index as well as two inequality variables in their operationalization of poverty, they risk confusing causal relations and ignoring conceptual and measurement overlap.

on GDP per capita at a country-level, money and other values must be siphoned off from the many various economic activities that make up the gross value of domestic output. This is conceivable in countries with a very low GDP output and in countries where a few individuals control large parts of the economy and/or state apparatuses. How GDP per capita can have a direct effect on the levels of corruption is more difficult to think of and the literature did not provide many examples. The one factor that can explain a direct effect is the bias towards poorer countries among the experts participating in a survey-based index of corruption, such as Transparency International's Corruption Perceptions Index. But this does not explain the correlation to indexes of corruption that do not include foreign "experts" or those that do not include surveys at all.

3.2 INEQUALITY

3.2.1 Consequences of Inequality

The detrimental consequences of inequality are many. Some of the detrimental effects of inequality include poor educational performance, violence and low levels of wellbeing, health and trust (Stephany, 2017). At its worst, economic and social stratification of a society can lead to anomie, in which people experience society as unpredictable and without order and normative constraints lose strength. In such a society, a sense of futility and unfairness will be more prevalent (Ariely & Uslaner, 2017). This can lead to higher levels of crime, including corruption (Zhang et al., 2009). And although the definitions may vary, corruption is considered a crime in most countries. It is therefore interesting to investigate the criminal consequences of inequality.

Ecological theories of crime look at the differing incentives, pressures and deterrents that individuals face in different environments (communities and societies) that explain variations in levels of crime. Three of the most influential ecological theories of crime in sociology - social disorganization theory, the strain theory and the economic theory – predict that inequality increases crime (Kelly, 2000).

When talking about crime it is natural to talk about violence. Homicide-rates, in particular, are oft-cited examples of violence (Wood, 2006). Additionally, there are many studies that tie inequality to several specific sub-categories of violence such as ethnonationalist civil war,

insurgencies and domestic violence (Cederman, Weidmann, & Gleditsch, 2011; Fearon & Laitin, 2003; Sanz-Barbero et al., 2015).

Morgan Kelly (2000), perhaps a bit counter-intuitively, finds that while there are strong correlations between violent crime and inequality, no such correlation exists for property crime and inequality. Phrased differently, crimes with economic incentives are impacted mainly by poverty while crimes of passion are inexorably linked to inequality and the associated social disorganization. This has some implication for the thesis. If corruption is more of a property crime than a violent crime, then in Kelly's model of crime and economy it should not correlate with inequality, but rather with poverty. There are two points that need to be considered before accepting the model's relevancy to the thesis. Firstly, it is a criminological model originating in studies on urban areas in the US. However, this does not negate the potential for generalization to a global level. Secondly, Kelly did not include corruption in his examples of property crime. Yet, when considering that corruption is linked to rational and economic incentives, corruption can be considered as within the sphere of property crime.

3.2.2   Inequality as a Cause to Corruption

Research on the correlation between inequality and corruption has often examined how corruption causes inequality (Batabyal & Chowdhury, 2015; Chetwynd et al., 2003; Gupta et al., 1998; Treisman, 2007). As to the possibility of the reverse direction of causality – that inequality has a causal effect on corruption – there is an increasing amount of research. A few of these even build on regression analysis on global datasets (e.g., J.-S. You & Khagram, 2005; J. S. You, 2012; Zhang et al., 2009). Their positive findings indicate a correlation between inequality and corruption, but this does not determine the direction of causation.

Uslaner finds that the fairness of the legal system has both indirect and direct effects on corruption: "The only institutional variable that matters is the fairness of the legal system, the perception (of elites from the Economist Intelligence Unit) that the rich and the poor receive equal treatment under the law. This, of course, is a form of inequality more than it is of legal structure" (Uslaner, 2013a, p. 3606).

Jong-Sung You and Sanjeev Khagram (2005) argue that inequality should be considered as equally important in increasing corruption as economic development is. Their main argument is that through material and normative mechanisms, inequality is socially conducive to high levels of corruption. In a very unequal society, the rich will own a disproportionally large part of the wealth and political influence. To avoid losing this they will work against perceived threats and chief among these are fair political, administrative and judicial processes. Undermining legal processes carries great potential reward for those who can afford it. In such a society the barriers to participation will be higher and the poor will see their influence, beyond the vote, decrease. This phenomenon is called state capture and concerns itself with economic inequality and, as a result, with inequality of influence (Dutta & Mishra, 2013). The poor cannot afford the bribes of the rich and lack the resources to organize effective political campaigns and push for methods of monitoring corruption. More so than the rich, the poor and middle classes are dependent on the functioning state apparatus as they cannot afford to circumvent it. To increase the chances of receiving certain services and goods, as well as avoiding certain taxes, many will resort to petty corruption and bribes (J.-S. You & Khagram, 2005).

Indranil Dutta and Ajit Mishra (2013) look at inequality and corruption from an economics perspective. They link corruption in one market, or sector, to inequality and resulting imperfections in related markets. Informal markets have a higher number of less productive firms, and they contend that less productive firms are under certain conditions more likely to be corrupt. When outside the regulated market, firms can afford to be less productive as the legal requirements and regulations for production and quality are easily ignored. Although they generally do not have access to the loans, investments and legal protections of their competitors in the legal market, the boundaries between the informal and formal markets blur when government inspectors and monitors can be bribed (E. Dabla-Norris & Inchauste, 2008). By bribing their way out of taxes and regulations corrupt firms can compete in the formal market. An increase in the number of these firms in a market increases the level of corruption by forcing the departure of productive, law-abiding firms (Dutta & Mishra, 2013).

Zhang et al. (2009) argue in their study that income inequality fosters corruption indirectly through reduced social support and reduced human development. They look at causes of

corruption from a criminological viewpoint and include social support theory in their analysis. Social support is something positive that can prevent or lessen occurrences of crime. In their study they define social support as "the willingness of governments to commit scarce resources to the aid and comfort of their residents" (Zhang et al., 2009, p. 208) and have measured it as public expenditure on healthcare[5]. As found in several other studies, the Gini index has no direct effect on corruption, but they do find that it has a robust, moderate indirect effect through social support and human development. In his review of corruption studies, Treisman (2007) also did not find support for a strong correlation between inequality and corruption.

The literature supports a correlation between inequality and corruption, but it is contingent on an intermediary.

3.3 COMPLEX RELATIONSHIPS: INEQUALITY, TRUST AND CORRUPTION

3.3.1 Inequality to Trust

In "All for All", Rothstein and Uslaner argue that inequality is causally prior to generalized trust, saying that generalized trust "both depends upon a foundation of economic and social equality and contributes to the development of a more egalitarian society" (Rothstein & Uslaner, 2005, p. 45). Equality of opportunity implies the application of fair rules and strong supporting structures, enabling among other things social mobility. Universal programs, such as universal healthcare and free public education, are associated with the rights of all citizens and not only a few groups. In contrast to means-tested social programs - where the beneficiary has to prove the need for aid, sometimes on a regular basis - universal social programs increase social trust through a more general and transparent redistribution of wealth. Equality of opportunity and equality of income correlate with a sense of optimism for a shared future, which in turn makes trusting strangers easier. The opposite should be true for strong hierarchical societies stratified by class or caste where particularized trust is strong:

> *"Where there is a dearth of social solidarity due to class envy, the social bonds of*
> *generalized trust will be weak, and so will the propensity (especially from the middle*

---

[5] This is similar to Rothstein and Uslaner's definition of equality of opportunity as public policies intending to create equal conditions for citizens (see section 2.2).

class) to pay high taxes. People will identify more with their class or ethnic group (or both) than with members of the larger society." - (Rothstein & Uslaner, 2005, p. 55)

By looking at levels of generalized trust and age- and education-specific measures of income inequality in 22 countries, Fabian Stephany (2017) finds inequality to be an important indicator of the level of trust[6]. He also finds that all countries, independent of GDP per capita, have age-specific income imbalances that decrease trust, especially in countries with generally high inequality. Stratification and the perception of inequality are two concepts that he argues are important to explaining how inequality leads to low trust. When societies become more unequal, or stratified, separate spheres of daily life emerge. In the end, large differences between in-groups and out-groups will act as barriers to trust. Perception of inequality is primarily made by socio-economic benchmarking and does not measure the actual level of inequality (Stephany, 2017). Importantly, this also works the other way around, economic measures of inequality do not measure perception of inequality - which should be an important factor in the level of trust. When people do not perceive society to have a shared fate, due to the economic differences in a stratified society, they will lose trust in out-groups, i.e. most people. The group that we compare ourselves with depends on many factors, but it also changes as we age. The link between inequality and trust should be different for a 20 year-old, who socializes with and compares themselves to peers who have recently moved from home for the first time and just started work or studies, and most 80 year-olds, who traditionally socialize with family and a smaller group of friends.

You (2012) argues that generalized trust is negatively affected by inequality through a sense of unfairness, rather than a sense of difference. He uses Rawls' *justice as fairness* concept to argue that fair distributive, procedural and formal justice leads to higher levels of trust. The measurements of these theoretical variables are "corruption (breach of formal justice), democracy (procedural justice), and income distribution (distributive justice)" (J. S. You, 2012, p. 703). Inequality might also violate a culture's *norm of distribution*, which regulate "what is seen as a fair allocation of income or other goods" (Elster, 1989, p. 101). Smerdon and Blauw

---

[6] His methodology shows the interesting results that this kind of research offers. By looking closer at sub-national patterns of inequality he shows the importance of including such complimentary measurements to the standard cross-national aggregated measures, that alone do not capture or explain the deeper patterns within a society.

find that the distribution of income and the process behind it is important: "that inequality's effect on trust is strongest when income is randomly allocated, is consistent with Uslaner's (2002) well-established theory of the psychological foundations of trust" (Smerdon & Blauw, 2017, p. 23).

According to You (J. S. You, 2012), rich people generally think they are *justly rewarded* while poor people are more prone to think they are *unjustly under-rewarded*. Gould finds that in many advanced countries the increase in income inequality in the last two decades has led to increased social gaps in society and that overall "inequality seems to reduce an individual's sense of trust, fairness, and helpfulness in others, and this finding is rather consistent across gender and education groups" (Gould & Hijzen, 2016). In the United States, inequality at the bottom reduces trust mainly for those who are impacted by this inequality, less educated people and those in the lower third of the income distribution. In Europe however, inequality at the bottom leads to a general reduction of trust among a broader section of society, including those who are not directly impacted by inequality. Despite the differences, the results suggest inequality at the bottom of the distribution lowers the level of trust in others, both in the United States and in Europe (Gould & Hijzen, 2016).

3.3.2 Trust to Corruption

In nearly all human interactions trust is involved. When humans successfully cooperate, this is a result of trust. When cooperation is not certain to be successful, the social sciences provide us with many metaphors: the prisoner's dilemma, the provision of public goods, the problem of collective action, and the tragedy of the commons (Uslaner, 2013a). Rawls explains the causal link:

> *"For although men know that they share a common sense of justice and that each wants to adhere to existing arrangements, they may nevertheless lack full confidence in one another. They may suspect that some are not doing their part, and so they may be tempted not to do theirs. The general awareness of these temptations may eventually cause the scheme to break down. The suspicion that others are not honoring their duties and obligations is increased by the fact that, in absence of the authoritative interpretation and enforcement of the rules, it is particularly easy to find excuses for breaking them."* – Rawls, 1971, in Rothstein (2013)

In other words, a culture of mistrust increases the level of perceived trust-violations, such as corruption, which in turn leads more people to not honor their duties and obligations. Findings from several cross-national studies using aggregate data lend empirical support to this view: "a lack of trust […] prevents the adoption of a universalistic ethos and cooperative behavior and favors instead instrumental and individualistic approaches to problems" (Morris & Klesner, 2010). If cooperation is costly to all parties in the short run but rewarding in the long run, trust is integral to keeping agents on-board. Without trust, the likelihood of agents reneging on their commitment increases.

Individuals "form their system of beliefs from the imperfect information available to them" (Rothstein, 2011, p. 173). The hearsay, rumors, collective narratives and personal experiences that are a part of this imperfect information affects both the perceived and real prevalence of corruption. Survey respondents will answer surveys based on the perceived prevalence of corruption, just as many of those participating in corruption will justify their behavior by the perceived high level of corruption: "*everyone else is doing it*" (Rose-Ackerman, 2001). A perception of corruption, both high-level and petty, is linked to lower levels of generalized trust (Rothstein & Uslaner, 2005).

3.3.3   Inequality to corruption via trust

As Eric M. Uslaner writes in *The Moral Foundations of Trust*: "Countries with more trusters have better functioning government, more redistributive policies, more open markets, and less corruption. What distinguishes countries that are trusting from those that are not is the level of economic equality" (Uslaner, 2002). The opposite of this, high inequality and low levels of trust, can ultimately lead to a view that "society is a zero-sum game between conflicting groups" (Rothstein & Uslaner, 2005, p. 46). You and Khagram argue that people in unequal societies are more likely to consider political institutions and rules as favoring the rich and as lacking legitimacy. People are more likely to circumvent laws and regulations when they are considered illegitimate (J.-S. You & Khagram, 2005, p. 139).

Several studies support what can be termed an *inequality trap*, where inequality, trust and corruption are viewed as mutually reinforcing (e.g., Rothstein & Uslaner, 2005; Uslaner, 2005; J.

S. You, 2012). Inequality leads to lower trust, which leads to higher levels of corruption, which in turn leads to more inequality (Uslaner, 2013a):

High inequality −> low trust −> high corruption −> high inequality

Each persists over long periods of time and countries face serious difficulties in escaping this trap - either to advance upward or to fall downward (Uslaner, 2005). If the persistent stickiness of corruption is not explained by institutional changes very well, the argument of corruption being a bottom-up phenomenon is strengthened.

Inequality, low levels of trust and corruption all undermine the ideals of what John Rawls called the *well-ordered society*, a society in which "everyone accepts and know that the others accept the same principles of justice, and the basic social institutions satisfy and are known to satisfy these principles" (Uslaner, 2005).

3.3.3.1   *A note on regional dynamics*

Latin American countries, countries with socialist legal origins and Eastern European countries are in several studies (e.g., Rothstein & Uslaner, 2005; Uslaner, 2005; J.-S. You & Khagram, 2005) held as examples of regions where the dynamic of the inequality trap is weakened[7]. You and Khagram (2005) and Uslaner (2002) found countries with a socialist legal origin to have a significantly more equal income distribution than others while at the same time being significantly more corrupt. The different dynamics are considered a result of the socialist legacy; despite scoring low on inequality, 80 percent in Bulgaria, Hungary and Russia believe that high incomes are a result of dishonesty (Uslaner, 2005). Similarly low levels of trust are normally only seen in countries with high inequality. You and Khagram warn that "failure to consider different conditions between socialist and nonsocialist legal origins obscures the effect of inequality on corruption" (J.-S. You & Khagram, 2005, p. 146).

---

[7] It should be noted that Latin American and Eastern European countries comprise 38 percent of the sample.

3.4 SUMMARY AND RESEARCH QUESTION

As shown in the discussion above the links between inequality, corruption and trust have been the subject of academic inquiries before. Central authors like Rothstein, Uslaner and You all see inequality through trust as an important causal model for corruption. Challenging this understanding is some of the literature on inequality and crime. That inequality leads to social ills is not challenged. Kelly finds that social disorganization and alienation are inexorably linked to inequality, and certain types of crime are as well. Property crime, or economic crime, however, is not found to be explained by inequality. The idea that property crime is linked to rational, economic incentives and that violent crime is linked to anomie and social disorganization contradicts the trust centered models of Rothstein, Uslaner and You. Yet all seem to receive support from the data.

Rothstein, Uslaner and You favor complex theories focusing on structural and normative explanations, but the trust-centered model could be an overly complicated model for what might turn out to be a simpler causal dynamic, or none at all. If Kelly is right, then there should be no correlation from inequality to corruption. Yet if there is a correlation, it does not disprove Kelly's theory, but rather provides further evidence of why corruption is a special case.

In order to investigate these issues, my research question is:

Is there a positive relationship from income inequality to corruption and is this effect dependent on trust?

4 THEORY AND HYPOTHESES

Answering the research question will be a small addition to the growing literature on the relations between inequality, corruption and trust. In order to answer the research question, I will generate testable hypotheses based on the theories and assumptions presented in the literature review.

Below is the model of causality in my hypotheses. The hypothesized causal link from income inequality to corruption is much stronger when it is indirect through trust. This is represented with thicker arrows than that of the direct effect of income inequality to corruption.

Independent Variables   Dependent Variable

Income Inequality → Corruption

Income Inequality → Low Trust → Corruption
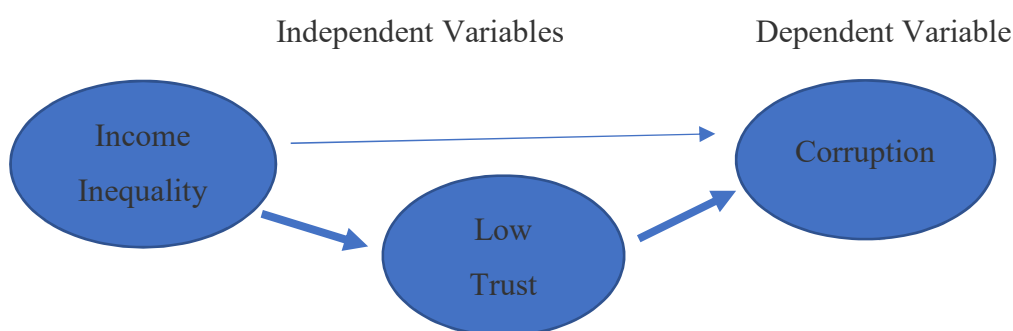
Hypothesis 1:

The first part of the research question asks *Does inequality have an effect on corruption*. Much of the literature, as presented in sections 3.3.3, supports a link between inequality and corruption. Even if much of the effect is indirect through trust, it will show as a direct relationship from income inequality to corruption, also when trust is in the model. Accordingly, my first hypothesis is:

Higher levels of income inequality lead to higher levels of corruption

Hypothesis 2:

The second part of the research question asks whether the relationship between inequality and corruption is *dependent on trust.* To answer this, I have to investigate the effect of trust on

corruption. The literature tends to support the argument that low trust societies experience higher levels of corruption, as presented in sections 3.3.2 and 3.3.3. Cooperation and adherence to a common sense of justice suffers where there is distrust. A culture of mistrust increases the level of perceived corruption, which in turn might increase the actual level of corruption. In other words, the conditions required for John Rawls' well-ordered society to function are absent in a low trust society: "a lack of trust […] prevents the adoption of a universalistic ethos and cooperative behavior and favors instead instrumental and individualistic approaches to problems" (Morris & Klesner, 2010). Following this, my second hypothesis is:

Trust is a significant variable in explaining corruption

Hypothesis 3:

The other hypotheses above cover individual relations in the research question. H3 covers the causal relations from inequality to corruption, via trust. In section 3.3.3, theories pertaining to this link were presented. The inequality trap describes the causal relations between the three variables. Additionally, several studies by other researchers, such as Rothstein, Ariely and You, have found support for this causal chain. Considering the strong theoretical support for the casual relations between the three variables, resulting in the other hypotheses, I expect H3 to be true:

There is a positive relationship from income inequality to corruption which is dependent on trust

5 DATA

In this chapter I will present the data used in this thesis. Additionally, I will discuss the variables'
measurements, sources and their validity to the theoretical phenomenon I wish to investigate.

The dataset in this thesis was created with data from three datasets. The Quality of Government
Standard Dataset, version jan2018 (Teorell et al., 2018) was the main dataset, but several
variables are from the 2019 and 2016 datasets (Teorell et al., 2019; Teorell et al., 2016). The
2018 Quality of Government dataset (2018 QoG dataset) includes data from and around 2014. If
a country does not have data for 2014, data for 2015 is included. If there is no data for 2015, data
for 2013 is included, this continues up to a maximum of any +/- 3 years. In other words, data for
2014 is prioritized but where this is missing the dataset includes data from the closest year,
ranging from 2011 to 2017. The same system applies to the 2019 and 2016 datasets, but with the
years 2015 and 2012 as the starting point.

5.1 CHOICE OF UNITS

The units in this thesis are countries. This is limited by the number of countries included for the
selected variables in the QoG datasets. By including as many units as possible, selection bias is
avoided, external validity is strengthened, and the validity of a generalization increases.

5.1.1 Data: single year or averaged

Rothstein and Uslaner (2005) and You and Khagram {, 2005 #11) write about the stickiness
problem of corruption, trust and inequality: they rarely change much over time. Because of this,
and to reduce measurement errors, You and Khagram {, 2005 #11) use averaged inequality data
for the years 1971-1996. Although averaging data might increase the accuracy of estimates,
benefits are offset by ignoring intertemporal variations. Further discussions of sample quality
and cross-sectional data are in section 6.3.

5.2 VARIABLES

5.2.1 Corruption

As there is no way of accurately measuring the level of corruption, indexes with data from multiple surveys represent the best method - such as TI's Corruption Perception Index (CPI) (Transparency International, 2018a). However, one issue with survey-based ratings is the risk for systemic bias and measurement error in the data. Survey respondents will always have subjective and possibly biased judgements. Survey answers may be biased by social desirability effects, where respondents answer what they believe is socially desirable and not what they have experienced (Justesen & Bjørnskov, 2014). As the CPI is an index of the perception of corruption, it avoids some of the problems that experience-based surveys might have, such as in countries where reporting on corruption might lead to negative consequences. However, there is good reason to ask whether surveys measure the countries' corruption levels or the countries' reputations, influenced by media coverage, stereotypes and rumors. In a review on the causes of corruption, Treisman (2014) finds that numerous studies show a discrepancy between expert and popular evaluations of the level of corruption, with experts more likely to overestimate the extent of corruption, especially in low income countries. Exports from richer countries might be negatively biased towards poorer countries, resulting in a stronger correlation between corruption and poverty than is true. Despite this, the CPI index is the most widely used indicator of corruption. And although there are some issues relating to the reliability of its data, the CPI correlates well with more objective indicators of corruption, such as citizen experiences with bribery (Transparency International, 2018b).

As mentioned in section 2.3, TI's definition of corruption is "the abuse of entrusted power for private gain". The CPI measures the perception among experts and business executives of corruption in the public sector by combining data from 13 surveys from 12 independent institutions. The scale goes from 0 (highest level of perceived corruption) to 100 (lowest level of perceived corruption). By using the average of at least three sources, potential errors in the individual surveys are compensated for. Another benefit of this composite index is that different manifestations and dimensions of public sector corruption are reconciled into one indicator (Transparency International, 2018b). Depending on the source data, the surveys capture many

different aspects of corruption, see appendix B, section 11.1 for more information on this. They write that the CPI does not capture "citizen perceptions or experience of corruption", but the experts providing answers to the survey are also citizens. Although the experts are experts in their field and far more knowledgeable than the average person, they are also subject to cultural norms, political and religious ideologies and prejudice. They, just as other citizens, "form their system of beliefs from the imperfect information available to them" (Rothstein, 2011, p. 173).

Additional measures of corruption are the Global Corruption Barometer (GCB) by Transparency International and the Public Sector Corruption Index and the Political Corruption Index by the Varieties of Democracy Project[8]. Each measurement provides benefits the others do not have, but the CPI has some strengths the others do not have. Unlike the World Bank's Control of Corruption, which uses different surveys measuring different aspects of corruption, the CPI measures aspects of corruption in line with the definition *the abuse of entrusted power for private gain*. Use of the same methodology for all countries provides quantifiable indicators. This is a reason for its popularity and by using the CPI it will be easier to compare the results of this study to those of other studies.

The CPI data in the regression is from the 2019 Quality of Government dataset (Teorell et al., 2019) and covers 178 countries with observations ranging from 2014 to 2017.

A histogram of the variable revealed that it was not normally distributed. According to econometric literature a logarithmic transformation is recommended. Following this a natural logarithmic transformation was performed on the variable, with good effect.

5.2.2 Inequality

To measure income inequality, I will use the World Bank's Gini Index from the 2019 QoG dataset. This index is widely used, and the data is easily available. Data is not available for every country in every year and is one of the weaknesses of using a single year. If I were to use data

---

[8] When both are run as independent variables against the natural logarithm of CPI's corruption variable, the adjusted R2 is 0.728 and the Variance inflation Factor (VIF) for both is 10.382 (see section 10.5.1). All three variables are from the 2019 QoG dataset (Teorell et al., 2019).

from a 10-year period, 2008-2017, the number of countries would increase. However, using Gini measures from 10 years prior to data for other variables might be problematic for investigating correlations.

The main weakness of the Gini coefficient is that it ignores different kinds of inequality. The Gini coefficient is less sensitive to inequalities in the bottom and top parts of the income distribution and most sensitive to the middle part (De Maio, 2007). In a global comparison of inequality, the different kinds of inequality should be of statistical and theoretical interest and importance. This is especially relevant for an investigation of its effect on trust: "Different measures of inequality will capture different mechanisms through which inequality may influence trust" (Jordahl, 2007, p. 8). De Maio argues that the GINI coefficient is best seen as one of many options for operationalizing income inequality.

Despite the above-mentioned weaknesses of the Gini index, I have chosen to use it for several reasons. All indexes of income inequality are attempts at a currently impossible feat: quantifying a pattern of income among a certain number of individuals, all while lacking both a mathematical definition of inequality and access to all the data. The lack of complete data is due to the, and quite understandable, absence of a universal ledger that records and stores every individual's official and unofficial income. The mathematical definition of inequality should be much easier to produce, but the large number of inequality indexes proves that we have not yet arrived at an agreement on a universal definition, and much less its estimate (Kokko et al., 1999). The Gini index overcomes these challenges by measuring easily accessible data according to a simple method, the Lorenz curve. In a graph showing complete income equality, with the cumulative proportion of population on the horizontal axis and the cumulative proportion of income on the vertical axis, there would be a 45-degree line going from the lower left corner to the upper right corner, where the poorest 1% earn 1% of the income, the poorest 20% earn 20% of the income, and so on. The Lorenz curve plots the actual distribution of income. Visually, the Gini coefficient is defined by the area between the diagonal 45-degree line and the Lorenz curve (McKay, 2002). Thus, the Gini index is expressed as the percentage that this area covers of the total area under the hypothetical 45-degree line.

The measure for inequality is the World Bank's Gini index as found in the 2019 Quality of Government dataset (Teorell et al., 2019, p. 693) and covers 114 countries with observations ranging from 2012 to 2017.

5.2.2.1 *Equality of Opportunity*

In addition to inequality of outcomes, of which income inequality is one measure, there is inequality of opportunity. According to Rothstein and Uslaner, inequality of opportunity can be measured by policies that intend "to create equal conditions for citizens regardless of their income, ethnic/religious background, sex, and race in areas such as health care, education, and social security and legal protection ("equality before the law") (Rothstein & Uslaner, 2005)". They argue that unequal access to opportunity for different groups in society is another important measure of inequality – which is also correlated to mistrust and corruption. This is a very interesting aspect of inequality and a variable measuring this would be a good addition to my model. The Bertelsmann Tranformation Index (Teorell et al., 2018) includes a variable called Equality of Opportunity, but as the index only covers developing and transition countries it excludes a majority of developed countries. This will further decrease the sample of countries and will be of little use in a global comparative study.

5.2.3 **Trust**

Generalized trust is the belief that most people can be trusted. Surveys like the World Values Survey (WVS), European Values Study (EVS) and European Social Survey (ESS) measure levels of generalized trust with the statements *Most people can be trusted* correlating to high generalized trust and *Needs to be very careful* corresponding to low generalized trust. Trust can also be measured with the question *Generally speaking, do you believe that most people can be trusted or can't you be too careful in dealing with people?* The question does not mention context and it asks respondents for their views on *most people*, which most respondents will understand includes a variety of people in their society (Rothstein & Uslaner, 2005).

Trust can be viewed as a measurement of survey respondents' "assessment of the moral standard of society", their sense of social solidarity and belief that the citizens of their country share a shared fate (Rothstein & Uslaner, 2005, p. 42). When viewed as a measurement of people's

evaluation of the moral standard of a society, it is easy to see the importance it plays in what an individual considers acceptable behavior[9].

There are limits however to what a single question can measure, especially concerning such an intricate and non-tangible term as trust. The term *most people* is vague and subjective. People distinguish others into two groups of social interactions and trust, in-group and out-group. *Most people* refers to the out-group, but the radius of social interactions varies according to cultures. For example, in Confucian countries the radius is narrow while in wealthy countries and European countries it is wider (Stephany, 2017). This poses a challenge, however, for cross-national comparisons. If the question is interpreted differently, how can we compare generalized trust in two different countries? Several studies (e.g., Delhey, Newton, & Welzel, 2011; Glaeser et al., 2000; Stephany, 2017) have found that the ambiguity of the question is its strength. If the term is culturally defined the variation should be low.

The measure for Trust is the Human Understanding Measured Across National (HUMAN) Surveys' *Social Trust* variable as found in the 2019 Quality of Government dataset (Teorell et al., 2019, p. 367) and covers 109 countries with observations ranging from 2012 to 2016. The HUMAN survey combines public opinion data from 19 surveys and averages the scores to create unique country-year observations. The score ranges from 0, the lowest possible level of trust, to 100, the highest.

5.2.4   Control variables

In the studies reviewed for this paper, all regressions included several control variables. By including theoretically and empirically relevant control variables we can control for their effect on variables in the model, thereby avoiding both specification bias and underspecifying the model (Kolnes, 2016; Jeffrey M. Wooldridge, 2012). At the same time, I do not wish to include irrelevant variables as they generally increase the variances of the other independent variables due to multicollinearity (Jeffrey M. Wooldridge, 2012). This problem is referred to as overfitting

---

[9]A study by Glaeser et al. (2000) combined two experiments and a survey and found that the question about trust is better at predicting trustworthy behavior (trustworthiness) than trusting behavior (Glaeser et al., 2000).

or overspecifying the model. As the causal direction cannot be known, including variables that are not directly relevant to my model is of questionable benefit. When related variables that are not directly relevant to the model are included, the R square (R2) increases near-automatically and won't decrease even when an irrelevant variable is added. Before including any control variables, I will consider their theoretical and empirical importance and avoid conceptual- and measurement overlap. An additional consideration when including a new variable is the number of observations. Most variables considered for the regression have observations around 100, with trust having the lowest number of 88[10]. The issue of sample quality and data availability is discussed further in section 6.3.

5.2.4.1   *Economic Development*

The attention economic development receives in studies is understandable when considering the many consequences of poverty, such as poor health, high levels of crime and low levels of education (C. Kellogg, 2018). Due to the importance of economic development in the corruption literature (see section 3.1.2) I believe it will be interesting and enlightening to include economic development as a control variable.

In the literature, Gross Domestic Product per capita (GDP per capita) is the standard measure of the level of economic development (e.g., Lučić, Radišić, & Dobromirov, 2016; Treisman, 2014; J.-S. You & Khagram, 2005). Following this, GDP per capita will be the measurement of economic development. Bias from government estimates and definitions of poverty is avoided.

The data is from the World Bank's World Development Indicators, as in the 2018 QoG dataset, and covers the years 2011 to 2014 and N=190. GDP per capita is the gross domestic product divided by the country's midyear population. It is the sum of the gross value added by all resident producers in the economy and all product taxes, minus subsidies not included in the value of the products. Data are in current US dollars (Teorell et al., 2018, p. 656).

---

[10] At one point in the data preparation the number of observations across all variables was eight. Meaning that only eight countries would be included in the analysis. This is a significant limiting factor and a reason for including fewer variables that measure larger dynamics.

A histogram of the variable revealed that it was not normally distributed. Following econometric literature, a logarithmic transformation was performed. This is quite common in corruption studies where GDP per capita is included (e.g., Treisman, 2014; J.-S. You & Khagram, 2005).

5.2.4.2 Democracy

In studies on corruption, democracy is often included as an institutional variable. Democracies, the reasoning goes, protect civil liberties, such as free speech, and nurture an independent press and public institutions, such as the judiciary. This should result in an independent judiciary continuously fighting corruption as well as protecting the right of the press to expose corrupt practices (Zhang et al., 2009). Following this reasoning, several studies on corruption have included democracy and a score for the independence of the judiciary as control variables. Some research has shown that for democracy to have an impact on corruption and trust, it must be long-term and stable. On trust, one study found that democracy has a negative effect for countries below the mean score of democracy and a positive effect for higher scores. This is supported by Uslaner, who finds that only after 46 years of continuous democracy does democracy produce a positive effect on trust, and that it has a negative effect in the beginning (J. S. You, 2012). The same has been found for the duration of democracy on corruption (Treisman, 2000, 2007). A recent study found that democracy reduces corruption only in countries with a GDP per capita of above 2000 USD (Jetter, Agudelo, & Hassan, 2015). The importance of both the duration of democracy as well as income per capita finds support in another study (Rock, 2009), but the number of years before a democracy has a positive effect on corruption is between ten to twelve years, significantly lower than what the findings of Uslaner and Treisman suggest.

To make my thesis more readily comparable with other studies, I have chosen to measure democracies with one single variable. Based on Polity IV's own categorization, I created a dummy variable that contains all countries with a score of seven or higher. According to Polity IV's Dataset User's Manual, all countries with a score of seven or higher are full democracies (Marshall, 2018, p. 35). A high score entails a well-functioning democracy and should weigh up for not including separate variables for how old or how rich the democracies are.

5.2.4.3   Rule of Law

In addition to democracy it will be beneficial to include a measure of institutional and governance quality. Of interest is the independence of the judiciary, respect for property rights, contract enforceability and the extent to which people abide by the rules of society and consider their enforcement predictable. There are many variables that measure each individual factor and variables that include many. As mentioned in the beginning of this chapter (see section 5.2.4), including variables that do not have observations for a large number of countries will result in a lower total number of observations that can be used in the regression. The Rule of Law is an aggregate indicator on the extent to which a society is successful in "developing an environment in which fair and predictable rules form the basis for economic and social interactions and the extent to which property rights are protected" (Teorell et al., 2018, p. 625).

The disadvantage of subjective variables, which comprise the Rule of Law variable, is the conceptual overlap with the perception of corruption. This is something that Treisman also noted in his review of literature on the causes of corruption (Treisman, 2007). The independence of the judiciary or the predictable enforcement of the law are conceptually quite similar to the measurement of corruption, whether there is "abuse of entrusted power for private gain". Despite this overlap, the analysis will benefit from this variable. However, due to a very high VIF and in order to control for the effect of low rule of law, I have created a dummy variable (see section 10.4) where a score above the median is 0 and a score on or below the median is 1.

The data is from the World Bank's Worldwide Governance Indicators, as in the 2018 QoG dataset, with 194 countries covering the years 2013 to 2014.

5.2.4.4   HDI

In addition to GDP per capita, the Human Development Index (HDI) can function as a second and different measure of development. A high HDI score indicates a high level of human development. The HDI includes three key dimensions of human development: the health dimension, measured by life expectancy at birth; the education dimension, measured by expected years of education for children starting school and the average years of education among those

25 years and older; and the standard of living, measured by gross national income per capita (GNI). The HDI is the geometric mean of the normalized indices. Although closely related to GDP, GNI includes net wages, salaries, property income, taxes and subsidies received from abroad (OECD, 2019). As an example, due to the high number of foreign companies in Ireland, the GNI in 2016 was nearly 30% lower than the GDP (Boland, 2017). Most countries do not experience this extreme difference, but many people in poor and middle-income countries rely on remittances that are included in GNI, but not GDP. And importantly, because the importance of income diminishes with an increasing GNI, the HDI uses the natural logarithm of income (Teorell et al., 2018, p. 589).

Previous research has shown that high levels of corruption are correlated to lower adult literacy rates, higher infant mortality rates, lower expenditures on education and higher dropout rates in primary schools (J.-S. You, 2006). And from the sections above, it is easy to understand these variables' theoretical relevance and correlation to inequality, trust and GDP per capita.

To capture the link between low human development and corruption I have created a dichotomous dummy variable. The country groupings are according to UNDP's cutoff points presented in the Technical Notes to the Human Development Report (UNDP, 2018). Countries with a score of 0.69 or lower are considered less developed and coded 1, while countries scoring 0.7 or higher are considered to be developed countries and are coded 0.

5.2.4.5 Natural resources for exports as part of GDP

The "resource curse" is not of direct theoretical interest to this thesis, but an increase in the shares of natural resources and fuels in exports have been shown to have a substantial effect on the levels of perceived corruption (Treisman, 2007). Countries where the extractive sector constitutes 25 percent or more of the GDP are highly dependent on this sector. Simply put, this dependence increases the potential for corruption by putting a small workforce, as compared to other sectors, in proximity to large amounts of wealth. Conceivably, such a large part of GDP will give the sector an outsized influence on political affairs, and should therefore correlate not only to corruption, but also to Rule of Law. By including this variable this interaction will be controlled for.

The data is from the World Bank's World Development Indicators, as in the 2016 QoG dataset, with 179 observations from 2012. Total natural resources rents are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents (Teorell et al., 2016).

5.2.4.6 Interaction Variable

H3 states that *there is a positive relationship from income inequality to corruption which is dependent on trust*, based on the assumption that inequality has a stronger effect on levels of perceived corruption in countries with low levels of trust than in countries with high levels of trust. In order to investigate this specific effect, I have created an interaction variable by multiplying two variables with each other (see Appendix A, section 10.4). The first variable is a dichotomous dummy variable for trust where countries scoring lower than the median have the value of 1 and the high-trust countries have the value of 0. This variable is then multiplied with the inequality variable to create the interaction variable.

5.3 DATA SUMMARY

Dependent Variable	Measure	Source
Corruption	A compilation of expert interviews and surveys. 0 = highly corrupt, 100 = not corrupt Min. year: 2014 Max year: 2017 N: 178	Transparency International's Corruption Perceptions Index as in the 2019 QoG dataset (Teorell et al., 2019).
Independent Variables		
Income Inequality	Gini index of distribution of income. 0 = perfect equality, 100 = perfect inequality Min. year: 2012 Max. year 2017 N: 114	The World Bank's GINI index as in the 2019 QoG dataset.
Trust	Generalized Social Trust question in surveys. 0 = no trust, 100 = highest score Min. year: 2012 Max. year: 2016 N: 109	The HUMAN Survey's variable for Social Trust as in the 2019 QoG dataset.
Control variables		
Economic Development	GDP per capita in current US dollars Min. year: 2011 Max. year: 2014 N: 190	World Development Indicators as in the 2018 QoG dataset (Teorell et al., 2018)
Democracy	Combined Polity Score Min. year: 2012 Max. year: 2016 N: 163	Polity IV Project as in the 2018 QoG dataset
Rule of Law	Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, in particular the quality of contract enforcement, property rights, the police and the courts, as well as the likelihood of crime and violence. Min. Year:2013 Max. Year: 2014 N: 194	World Bank's World Governance Indicators as in the 2018 QoG dataset
Resource Revenue	Total natural resources rents are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents. As a percentage of GDP. Min. Year: 2012 Max. Year: 2012 N: 179	World Bank's World Development Indicators as in the 2016 QoG dataset (Teorell et al., 2016)
Human Development Index	A long and healthy life, being educated and having a decent standard of living. Year: 2014 N: 186	United Nations Development Programme's Human Development Index as in the 2018 QoG dataset.

6 METHOD

For the analysis I have chosen to use the method of Ordinary Least Squares (OLS) within the framework of the classical linear regression model (CLRM). OLS is a method to model a linear correlation, through the best possible fitted line, between the dependent variable (Transparency International's Corruption Index, see section 5.2.1) and the independent variables (see section 5.2.2 through section 5.3). Essentially, this method attempts to explain how variations in the dependent variable might be related to changes in the independent variables - allowing us to test our hypotheses (see chapter 4). OLS is a recognized method for large-N comparative studies and a primary advantage is that it allows for us to control for the effect of different variables (Midtbø, 2007; Jeffrey M. Wooldridge, 2012).

6.1 OLS REGRESSION - ASSUMPTIONS

The basis of a regression equation is usually like the following:

$$\gamma = \beta_0 + \beta_1 X + \varepsilon$$

In the equation, Y represents the dependent variable, X the independent variable, b_0 is the intercept (or constant) which is where the fitted line crosses the Y-axis, b_1 is the angle of the slope coefficient, and e is the error term/residual which contains the variation of Y that is not explained by X (Ringdal, 2013).

I want the regression analysis to produce unbiased estimates with the smallest possible variance. The Gauss-Markov theorem states that if a linear regression satisfies the six classical CLRM assumptions, then it is a Best Linear Unbiased Estimator (BLUE).

> *"The Gauss-Markov Theorem is telling us that in a regression model, where the expected value of our error terms is zero, $E(\epsilon_i) = 0$ and variance of the error terms is constant and finite $\sigma^2(\epsilon_i) = \sigma^2 < \infty$ and ϵ_i and ϵ_j are uncorrelated for all i and j, the least squares estimator b_0 and b_1 are unbiased and have minimum variance among all unbiased linear estimators"* (A.D., 2015).

Below I will present the assumptions and the degree to which the main variables in my model satisfy these assumptions. This is to make sure my regression analysis produces unbiased estimates with the smallest possible variance and that it satisfies the assumptions.

6.1.1   Assumption 1

The linear regression model is linear in parameters

The dependent variable (Y) needs to be the combination of independent variables (X) and the error term. Additionally, the model must be correctly specified to avoid specification bias and specification error (Gujarati & Porter, 2010, p. 54). This means that the model is linear in parameters, but not necessarily linear in variables. The model below satisfies the assumption:

$$Y = b_0 + b_1 X + e$$

Scatterplots for the relationships between inequality and corruption and trust and corruption were created (see Appendix A, section 10.2). Additionally, if residuals are normally distributed (see assumption 3) and homoscedastic (see assumption 5), I do not need to worry about linearity (Goetsch, 2018).

6.1.2   Assumption 2

Independent variables are uncorrelated with the error term

This assumption refers to what is called exogeneity. Endogeneity is the opposite and refers to when there exists a correlation between the independent variable and the error term. A correlation to the error term is also a correlation to the dependent variable.

Although endogeneity occurs in several forms, the distinctions are not always sharp, and several sources of endogeneity may occur in the same equation (Jeffrey M Wooldridge, 2002, p. 51). The most relevant forms to this thesis are omitted variables, measurement error and reversed causality and are presented in section 6.2.1.

Unless we can explain all the variance, we can never be certain that the independent variables are uncorrelated to the error term. But according to Gujarati and Porter (Gujarati & Porter, 2010, p. 97), if the variables in my model are non-stochastic (non-random, i.e. fixed numbers in repeated sampling) assumption 2 is fulfilled. As this is the case, and to the extent that I can test this assumption, there is nothing that indicates that assumption 2 has failed. Assumption 2 has been met.

6.1.3   Assumption 3

The error term follows a normal distribution and has a population mean of zero

The error term is the variation in the dependent variable that the independent variable does not explain. As these values should be determined by random chance, the expected average value should equal zero (Gujarati & Porter, 2010, p. 55). Following this is the expectation that the error term follows a normal distribution. Yet the assumption of normality is optionable because it is not necessary for the validity of the OLS method nor for it to be BLUE. However, when error terms follow a normal distribution, we can perform statistical hypothesis testing and generate reliable prediction intervals and confidence intervals (Gujarati & Porter, 2010, p. 62). In other words, the error terms should be IID (Independent and Identically Distributed).

To test this assumption, a histogram with a superimposed normal curve and a Predicted Probability (P-P) plot were created (see Appendix A, section 10.2). In the histogram, the mean and standard deviation should have values of approximately 0 and 1, respectively. This is the case for the histogram of the dependent variable (non-logarithmic) where the mean is 1.43E-14 and the standard deviation is 0.943. As the histograms' appearance largely depends on the selection of the correct bin/column width, which can be deceptive, I will additionally check the P-P plot (Laerd Statistics, 2015).

In the P-P Plot, the residuals are not perfectly aligned along the diagonal line but are quite close. Because multiple regression analysis is fairly robust against deviations from normality, the residuals are close enough to normal for an analysis to proceed. With the mean in the histogram at 1.43E-14, the assumption of normality is met.

6.1.4 Assumption 4

There is no perfect collinearity

In case of more than one IV, there should be no exact linear relationships between them, meaning no perfect collinearity. Although there is no absolute number at which multicollinearity officially becomes a problem, it is best to avoid high correlations between the independent variables (Jeffrey M. Wooldridge, 2012, p. 95). If the degree of collinearity is high it affects the estimation procedure, producing high standard errors and coefficient estimates that become sensitive to small changes in the model, reducing the precision and explanatory power of the model. A solution may be to drop one of two variables that are measuring the same variance. It should be noted that this assumption allows for collinearity, just not perfect collinearity. As some collinearity will always exist between variables, excluding all collinearity would decrease the benefit of multiple regression analysis (Jeffrey M. Wooldridge, 2012, p. 84).

In Model 2 in section 7, with all the independent variables included, only one variable has a VIF[11] value above 5. GDP per capita has the highest VIF score, 6.097. Collinearity might therefore be at a problematic level (Christophersen, 2009, p. 161; Jeffrey M. Wooldridge, 2012, p. 84). A VIF score between 5 and 10 indicates high multicollinearity and might warrant a corrective measure. Yet there are several reasons why a corrective measure is not necessary in this case. A VIF of 6 is only slightly above 5 and it is already the natural logarithmic of the original variable. Additionally, the variables in the model are broad measurements that should correlate to economic development. As the remaining VIF scores range between 1 and 3.6, there is no perfect collinearity and assumption 4 has been met.
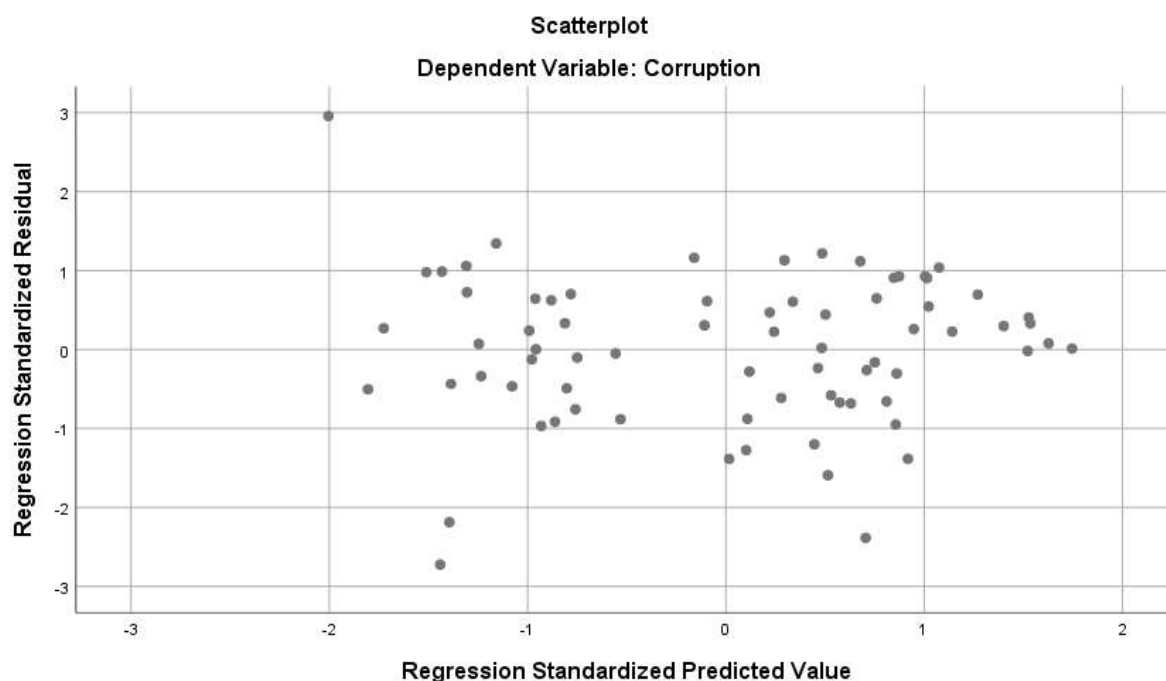
6.1.5 Assumption 5

The variance of the error term is homoscedastic

For the error terms to be homoscedastic the error terms must have a constant and equal variance. If this is not the case, the regression model has heteroscedastic errors and will likely result in incorrect estimates, weakening conclusions concerning the statistical significance. This will not

---

[11] As VIF is the reciprocal of Tolerance, I only need to consult one of these measures (Christophersen, 2009, p. 161).

alter the consistency or unbiasedness properties of OLS, but OLS estimators will not be efficient or of minimum variance (Gujarati, 2011, p. 82).

Scatterplot

Dependent Variable: Corruption

The scatterplot of standardized residuals versus standardized predicted values appear random, suggesting that the assumption of homoscedasticity has been met.

6.1.6   Assumption 6

There is no autocorrelation

No autocorrelation means the error terms of different observations should not be correlated with each other. This is an assumption especially relevant for time-series regressions and is rarely included in cross-sectional regressions (Jeffrey M. Wooldridge, 2012). Despite this, I used the Durbin-Watson statistic to test this assumption. The Durbin-Watson value is 1.961, quite close to the preferred value of 2. Assumption 6 has been met.

6.2 OLS Regression - Other Deliberations

6.2.1 Sources of Endogeneity

In section 6.1.2 I mentioned that endogeneity can occur in several forms and that the distinctions are not always sharp. Below are the most relevant ways in which endogeneity might occur in this thesis.

6.2.1.1 *Measurement error*

Measurement errors arise when what we want to measure can only be observed through an imperfect measurement. When this happens, part of the causal effect on the DV might be excluded from the model and will therefore become part of the error term (Jeffrey M Wooldridge, 2002, p. 51). To avoid this, I have chosen to trust the expertise behind the pre-constructed indexes. The disadvantage is that it is beyond the scope of this study to perform an independent evaluation of measurement error.

6.2.1.2 *Simultaneity and reversed causality*

Simultaneity occurs when an independent variable and the dependent variable are simultaneously determined. If this is so, then the independent variable and the error term are correlated (Jeffrey M Wooldridge, 2002, p. 51). Reverse causality occurs when the causal direction runs the opposite way to what we hypothesized. As this thesis a cross-sectional investigation of correlations, using the latest data for the three main variables, this is expected.

6.2.1.3 *Omitted Variables*

Omitting relevant variables leads to underfitting, or underspecifying, the model. To avoid this, a mis-specified model, it is important to consider all theoretically relevant variables when developing a model. At the same time, I do not wish to include irrelevant variables as they generally increase the variances of the other independent variables due to multicollinearity (Jeffrey M. Wooldridge, 2012). This problem is referred to as overfitting or overspecifying the model.

To resolve this balance, I have generally included variables that are commonly accepted to have a significant relationship with corruption. In other words, I have used the same control variables

as those used in other relevant studies (see e.g., Chetwynd et al., 2003; Gupta et al., 1998; Rothstein & Uslaner, 2005; Treisman, 2007; Ünver & Koyuncu, 2016; J.-S. You & Khagram, 2005). Using the same major control variables as other studies in the field has the added advantage of possibly generating comparable results.

6.3 CROSS-SECTIONAL DATA

There are several reasons why I have chosen to run a regression using cross-sectional data and not time-series data. Several of the variables in this thesis, especially corruption and inequality, have been shown not to vary much over time (Rothstein & Uslaner, 2005, p. 65; J.-S. You & Khagram, 2005). Additionally, and most importantly, sporadic reporting from states on major variables such as GINI and the irregular surveys of trust makes a time-series analysis difficult to impossible. The World Values Survey time-series variable for generalized trust has an average of ten countries per year and three years of data per country (Teorell et al., 2019, p. 787). The source for trust in this thesis, the HUMAN Survey, has a time-series variable with an average of 32 countries per year and eight years per country.

In the surveys used for the CPI, year-to year changes can result from a changing methodology, different samples and different respondents. To counter this the CPI score is a 3-year moving average, meaning that changes are reflected in the score gradually. In a long-term perspective this should not invalidate the relevance of the data, but a year to year analysis or using lagged variables do not provide enough substantive insight to challenge the usefulness of a cross-sectional analysis in the case of this thesis (Teorell et al., 2019, p. 619). The direction of causation is another important matter where the use of time-series data could be useful, but due to the lack of sufficient data this is not possible.

As mentioned in chapter 5, all variables, except Resource Revenue and HDI, include data from several years. The number of years ranges from two to five, and because this is not controlled for the correlations might be obscured.

6.3.1 Dealing with missing values

Missing values is a challenge with cross-sectional data from a single year. Using a global dataset with many variables from several indexes will inevitably lead to missing values. Different indexes focus on different regions while some major measures, such as Gini and GDP per capita, depend on countries' self-reporting - and countries don't always report[12]. There are several ways of dealing with missing values, either by excluding them or substituting with the mean. Substituting with the mean would increase the N, but this must be weighed against the potential for biasing the coefficients towards the mean. It will also decrease the standard deviation and, following this, the standard error. The potential drawback is that the regression analysis provides results that are more accurate than they actually are, falsely providing significant coefficients (Christophersen, 2009, p. 164). Considering that only 73 of 194 countries have data for all the variables, I would have to impute values for the remaining 121 countries. In this study, excluding values listwise is the most valid approach for a rigorous analysis.

6.4 DIRECTION OF CAUSALITY

The causal direction cannot be determined in this study as I am not conducting a longitudinal study. While correlation is not sufficient to claim causality, the lack of a correlation can be taken as an indication of a lack of a causal relation. Although this approach cannot prove a causal link, it does open for the possibility of disproving it. So unless otherwise noted, whenever causality is mentioned it is informed and supported by theory, and not data.

6.5 LEVEL OF SIGNIFICANCE

Due to the differing number of observations per variable, only 73 countries are included in the regression. As a result, the p-values might be higher than what a more complete dataset would have. This is because the p-value calculations assume the null-hypothesis is true and that all variation in the sample is caused by chance. In general, sample variations and errors in cross-

---

[12] The data availability for income inequality was surprisingly low, with no data for many countries, even countries such as Japan, Algeria and Serbia (Teorell et al., 2019).

national comparisons are rarely random, and in many cases the ability of a state to report data is correlated with other underlying variables, such as corruption, GDP per capita or Rule of Law.

A p-value does not directly prove or disprove a hypothesis, but a low p-value signifies statistical significance and that the null hypothesis can be rejected. If there are several alternatives to the null hypothesis, a rejection does not help in discerning which alternative is the correct one. Following this, I have chosen to formally test my hypotheses with a 95 percent threshold, that is a minimum significance of 0.05 or lower. This is done to avoid a type-II error. But due to the low number of countries and the issues surrounding data quality (as mentioned in chapter 5) a dynamic interpretation is preferable to a rigid disregard of variables with high p-values.

6.6 SAMPLE QUALITY

The goal of this paper is to perform a global comparison, but as mentioned above in section 6.3 and 6.3.1, the lack of data precludes this possibility.

Region	Count
Western Europe and North America	19
Latin America	16
Africa, sub-Saharan	15
Eastern Europe and Ex-Soviet states	12
Middle East and North Africa	7
East Asia	2
South East Asia	1
South Asia	1
Pacific	None
The Caribbean	None

Figure 1: List of regions with the number of countries included in the regression analysis.

The table above[13] reveals an imbalance in the sample of countries. National and regional variations increase in importance with an incomplete selection of countries, this is especially true if the selection is unbalanced. The four regions with the highest representation, though none are complete, are Western Europe and North America, Latin America, sub-Saharan Africa, and Eastern Europe and Ex-Soviet states. The Middle East and North Africa, East Asia, South East

---

[13] Hadenius and Teorell classified the world into ten politico-geographic regions, based on two considerations: geographical proximity and demarcation by area specialists (Teorell et al., 2019, p. 364).

Asia and South Asia are underrepresented. While no countries in the Pacific or Caribbean regions are included. For a complete list of the included countries, see Appendix A, section 10.1.

Looking at the two descriptive statistics tables for corruption[14] in section 10.6.1 we see a 0.69 decrease in range, from 2.43 in the original variable with 178 countries to 1.74 for the 73 countries in the study. This is a 28.4 percent decrease in range. The maximum value of 4.51 is the same for both variables, but the minimum value increases from 2.08 to 2.77. From the lower half of the original range (countries who scored less than 3.30) only 13 out of 34 countries are included in the 73-country sample.

By only including 73 countries, the potential for generalizing to a global level is weakened. All results must be considered in light of the regional imbalance and the potential change that every additional country would bring.

6.7 INTERACTION TERMS

Many causal arguments require certain conditions to be met before an effect is present. It is therefore quite common to use interaction terms in quantitative political science. H3, which can be considered the main hypothesis of this thesis, is conditional in nature because the effect of X on Y is strongly, but not completely, conditioned by the value of Z. In other words, the effect of income inequality on corruption is different in a country with high levels of generalized trust than it is in a country with low levels of trust. Because the effect of income inequality on corruption varies according to the value of trust, an interaction term testing the specific context conditionality should be created (Brambor, Clark, & Golder, 2005, p. 2).

According to Brambor et al. (2005), 90% of surveyed articles[15] ignored well-established methodological points when including an interaction term. They subsequently created a checklist to avoid the three most common mistakes made in multiplicative interaction models. Firstly, all

---

[14] Note that both the full and 73-country corruption variables are the Natural Logarithms of the CPI.
[15] The survey included 156 articles from 1998 to 2002 from three political science journals: American Political Science Review, Journal of Politics, and American Journal of Political Science (Brambor et al., 2005).

constitutive terms should be included. In other words, all variables that constitute the interaction term must be included as independent variables in the model (Brambor et al., 2005, p. 4). A simple regression equation is similar to that of a normal regression equation (see section 6.1) and looks something like this:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon$$

Secondly, constitutive terms should not be interpreted as independent or unconditional effects as they would be in a linear-additive regression model. The independent effects of X and Z in the model cannot be interpreted in isolation. In a multiplicative interaction model, the effect of independent variable X on dependent variable Y depends on a modifying independent variable Z. When an interaction term (X*Z) is present in the model, and if Z is a dichotomous 0-1 variable, the independent variable X can only be interpreted as the effect of X on Y when Z is absent, i.e. Z=0 (Brambor et al., 2005; Jeffrey M. Wooldridge, 2012), simplifying the equation to:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Likewise, the effect of the interaction term can only be interpreted as the effect of X on Y when Z is present, i.e. Z=1.

Finally, substantively meaningful marginal effects must be presented in a more informative and detailed manner. This will be done by including a graph (see section 10.8) that illustrates the interaction terms' marginal effects and substantial significance. Because the conditioning variable is a dichotomous dummy variable, I will not be able to interpret the effect at many different levels of trust, but the graph will help us see the interaction term's substantive and statistical significance in a more informative and detailed manner (Berry, Golder, & Milton, 2012; Kolnes, 2016).

Additionally, because the interaction term is standardized along with the other variables the standardized betas cannot be interpreted. Instead the unstandardized betas (B) should be used instead (Christophersen, 2009).

Due to the high collinearity between the constitutive variables and the interaction term, multicollinearity will be higher in such a model. However, this also shows the lack of sufficient data to provide accurate estimations of the parameters, which is correctly reflected in the larger standard errors (Brambor et al., 2005, p. 62). Centering the variables is suggested as a remedy for this issue (e.g., Christophersen, 2009), but this has been criticized as substantively ineffectual by others, amongst them Brambor et al. (2005).

7 RESULTS

The standardized beta coefficient standardizes the variables before the regression is run. By placing them on the same scale it is possible to compare the relative effects of each variable on the dependent variable. However, as written in section 6.7, a model that includes an interaction term complicates the interpretation of coefficients. Following this, model 3 will be interpreted with caution, focusing on the unstandardized betas (B), statistical significance, the direction of the correlation and any other results of interest.

When reading the results, it is important to remember that a higher score on the corruption variable equals lower levels of perceived corruption[16]. This means that while a .2 beta will increase the corruption score it will be written as a .2 decrease in corruption.

---

[16] Think of the score as an evaluation or test, where a higher score is better.

Running the OLS regression described above in SPSS provides the following results:

Adjusted R Squares
Model 1: .756
Model 2: .767
Model 3: .808

Model		Unstandardized Coefficients		Standardized Coefficients	Sig.	VIF
		B	Std. Error	Beta		
1	(Constant)	3.082	.416		.000	
	Income Inequality	-.005	.004	-.093	.203	1.530
	GDP per capita	.114	.035	.415	.002	4.881
	Rule of Law	-.433	.073	-.500	.000	2.099
	Democracy	.069	.069	.075	.320	1.668
	HDI	.115	.098	.129	.244	3.533
	Natural Resources	-.204	.128	-.110	.116	1.413
2	(Constant)	3.060	.406		.000	
	Income Inequality	.000	.004	-.007	.936	2.077
	GDP per capita	.079	.039	.287	.045	6.097
	Rule of Law	-.442	.071	-.511	.000	2.107
	Democracy	.104	.069	.113	.141	1.773
	HDI	.094	.096	.105	.333	3.574
	Natural Resources	-.232	.126	-.125	.070	1.430
	Trust	.005	.002	.186	.045	2.574
3	(Constant)	3.073	.369		.000	
	Income Inequality	-.010	.005	-.187	.037	2.904
	GDP per capita	.064	.035	.233	.074	6.170
	Rule of Law	-.461	.065	-.532	.000	2.119
	Democracy	.097	.063	.105	.130	1.774
	HDI	.043	.089	.048	.628	3.656
	Natural Resources	-.235	.114	-.127	.044	1.430
	Trust	.014	.003	.554	.000	5.986
	Trust (dummy) x Inequality	.010	.003	.547	.000	7.569

Model 1:

The adjusted R square, or R^2, of model 1 is .756, which means that 75.6% of the variation in corruption is explained by the independent variables of the model[17]. At -.093, the standardized beta coefficient (beta) for inequality indicates that if there is a relationship between the two it is not a very strong one. The correlation between income inequality and corruption is surprisingly low and is not statistically significant (p=.203). Looking at the other beta coefficients we see that Rule of Law has the strongest correlation with corruption with a beta of -.500 while GDP per capita has the second strongest correlation with a beta of .415 – both are significant at the 1% level. This means that for every unit increase in GDP per capita there will be a .415 unit decrease in the level of perceived corruption. Democracy, HDI and Natural Resources are not statistically significant.

Model 2: Introducing Trust

The adjusted R^2 of model 2 is .767. This is a modest change from an R2 of .756 in model 1. However, an R2 close to 8 for both models means their explanatory power is very high and indicates some good choices with regards to the selection of control variables.

Income inequality's beta has increased by .086 to -.007, but its p-value has had a significant increase to .936. The new variable, Trust, has a moderate effect with a beta of .186 and is statistically significant at the 5% level (p=.045). Rule of Law still has the highest correlation with a beta of -.511 and a p-value of .000. GDP per capita still has the second strongest beta but has fallen to .287, with a p-value of .045. At 6.097, GDP per capita's VIF reveals that its variance is 509.7% more likely to be inflated than it would be if it was not correlated to the other independent variables[18,19], slightly weakening the reliability

---

[17] Unlike the R Square, the adjusted R2 penalizes a higher number of variables and only increases if a variable adds to the explanatory power of the model more than what would be expected by chance (Jeffrey M. Wooldridge, 2012).
[18] Rule of Law is an apparent source of this inflation as it has a high correlation to corruption, trust and GDP per capita. The scatterplot in section 10.7.6 reveals an R2 of .602 between GDP per capita and Rule of Law.
[19] GDP per capita's beta decreased with .128 and its VIF increased by 1.216 in model 2, signifying a correlation to trust. This high collinearity should not be considered surprising for several reasons. A strong beta and the scatterplot in section 10.2 show that GDP per capita's explanatory power is high. As mentioned in section 2.1, bias among survey respondents might increase the correlation between economic wealth and corruption, by rating wealthier countries as less corrupt and poorer countries as more corrupt. Low levels of perceived corruption and high economic development both correlate to stronger rule of law, strong institutions, efficient bureaucracies and high

of its estimates (Allison, 2012). Yet it is not high enough to warrant a corrective measure[20].

The coefficients of Democracy, HDI and Natural Resources have slight changes in model 2, but they are still not statistically significant.

Model 3: Introducing the Interaction Term

The adjusted R2 of model 3 is .808, the highest of the three models. Including the interaction term has also changed the estimates of several variables, but the standardized betas should not be interpreted as reflecting the actual standardized betas for the reasons mentioned in section 6.7. To avoid drawing incorrect conclusions I will interpret the unstandardized betas, referred to as B.

As the non-modifying constitutive term of the interaction term, Income Inequality's B has strengthened in effect from 0 to -0.01 and is now significant at the 5% level (p=.037). Trust, the source for the dichotomous variable which serves as the modifying constitutive term of the interaction term, has increased from a B of .005 to .014 and is significant at the 1% level (p=.000). The changes in effect size and lower p-values must be seen in light of the interaction term and their roles in it.

The interaction term has a B of .01 and is statistically significant at the 1% level (p=.000). But contrary to much of the literature reviewed for this thesis it has a positive sign, meaning that income inequality in low-trust countries is associated with an improvement on the corruption score, which is a decrease in the level of perceived corruption. In other words, income inequality in low-trust countries correlates to *less* corruption[21]. This is very interesting. In models 1 and 2,

---

levels of trust. Yet the strong correlation does not imply causal direction, and it may be that corruption leads to less economic development. Despite a very strong correlation, the direction and specifics of this relationship is not fully understood, and neither are the roles played by income inequality and trust.

[20] If it were higher, there are several corrective measures I could use, such as performing a different kind of regression that handles multicollinearity better, removing the variable or transforming the variable. As GDP per capita is already the natural logarithm of the original data and running a different kind of regression is beyond the scope of this thesis, removing the variable would be the most relevant option. However, I should consider that the high VIF is not problematic insomuch as it is emblematic of the complex interdependence of the real world and of the difficulty of trying to isolate the effects of specific measurements.

[21] Since the interaction term was first created and used in the model, it has had a positive sign. That it took some time before I noticed this is a great example of why it is important to know the data one is using – independent of theory.

inequality's variance was all in one variable and was statistically insignificant – indicating that the variance was too dispersed. Now that its variance is split in two, high-trust and low-trust countries, both the interaction term and income inequality are statistically significant.

As expected, the VIF-values of Income Inequality and Trust have increased with the introduction of the interaction term. Trust's VIF nearly doubled to 5.986 while Income Inequality's VIF increased only slightly from 2.077 to 2.904. As the interaction term is an *interaction* of effects already present in the model, a high VIF is expected, and at 7.569 it is not problematic.

The B's of Rule of Law, Democracy and Natural Resources have not changed much and are statistically significant at the 5%-level, while GDP per capita (p=.074), Democracy (p=.13) and HDI (p=.628) are not.

8 ANALYSIS

Here I discuss how the hypotheses hold up in view of the regression results.

8.1 HYPOTHESIS 1

Higher levels of income inequality lead to higher levels of corruption

Model 1 provides weak and inconclusive results for the direct correlation between income inequality and corruption. As can be seen in the regression results above, this hypothesis does not find support in the data. The threshold of significance is at 0.05 (see section 6.5) and a p-value of .203 indicates a 20.3 percent likelihood of erroneously discarding the null hypothesis. Since both income inequality and corruption are natural logarithms, the beta can be interpreted as a 9.3% increase in corruption, which in linear algebra terms is not a steep slope. Strengthening the null-hypothesis is the constant, which at 3.082 means that even with perfect equality the model predicts a moderate level of corruption[22]. The beta of Income Inequality in model 1 is very weak, add to this the high p-values and it is easy to conclude that there is no relationship[23].

In model 2, the inclusion of trust has decreased Income Inequality to statistical insignificance and the constant is nearly unchanged. Although a lack of direct correlation does not negate or cancel a hypothesis, the data in model 1 and model 2 does not provide grounds for rejecting the null-hypothesis of H1.

Discussion and theoretical implications:

The lack of a statistically significant direct correlation between inequality and corruption finds support in some studies but is contrary to the findings of others. Several studies (Treisman, 2000, 2007; J.-S. You & Khagram, 2005) found that there is a weak direct correlation, but when certain economic and institutional factors are controlled for, especially GDP per capita, this correlation

---

[22] The range of the Corruption variable goes from 2.08 to 4.51, with the mean and median both at 3.6. See descriptive statistics tables for Corruption in section 10.6.1.

[23] But this might be due to a lack of information. The interaction term in model 3 reveals a correlation that is not possible to see in models 1 or 2.

weakens even further. The results in models 1 and 2 seem to support these findings with both GDP per capita and Rule of Law having the strongest correlations to corruption.

Data availability might explain part of this, as mentioned in section 6.6 the variance of corruption is smaller among the 73 countries in the regression than it is in the original variable. Yet a true global study might not be necessary to produce a more accurate relationship. The scatterplot between Income Inequality and Trust in section 10.6.3 reveals a clear split in the spread at the median of trust, with all but one country clearly belonging to either the higher-trust group or the lower trust group. If this effect is not controlled for, the variance will naturally result in a high p-value. This interaction is further explored in model 3.

In Kelly's model of crime and economy (see section 3.2.1), inequality does not correlate to corruption, but poverty does. The findings both here and in the literature on corruption supports the strong and significant correlation between GDP per capita and corruption. But Kelly's model cannot disprove the potential for an indirect correlation from income inequality through another variable or, as model 3 also reveals, a direct correlation.

8.2   HYPOTHESIS 2

Trust is a significant variable in explaining corruption

The second hypothesis predicts a positive relationship between trust and corruption – higher levels of trust correlate to higher levels of corruption. The regression results in model 2 show that trust has a moderate positive correlation that is statistically significant. Yet, considering the beta of 0.186 in terms of linear algebra, the slope is not very steep. And at 3.06, the constant in model 2 has barely changed to that of model 1 and signifies that even with no trust, there is only a moderate level of corruption. Despite the strength of the correlation being slightly weaker than expected[24], the data in model 2 provides grounds for rejecting the null-hypothesis of H2.

---

[24] The scatterplot between trust and corruption in section 10.2 reveals an R2 of .201, but this scatterplot includes many countries that are excluded in the models above. The scatterplot with countries only in the regression result in an R2 of .298.

Discussion and theoretical implications:

H2 did not fail, but as the hypothesized causal drive and main link to corruption the correlation was weaker than expected. According to Morris (Morris & Klesner, 2010), a culture of mistrust precludes the creation of a universalistic ethos. As humans, we infer on the world our experiences and beliefs. Low generalized trust means we do not expect others to honor their duties and obligations - violating Rawls' well-ordered society - and leading to individualistic approaches to problems. In the "absence of the authoritative interpretation and enforcement of the rules" individuals are more likely to break them (Rothstein, 2013). Low trust should therefore correlate to more corruption. However, the institutional control variables are included in the models due to their correlation to corruption. As seen in chapter 7, they also correlate to trust and modify the effect of trust further. The moderate effect is therefore not necessarily much lower than what should be expected.

8.3 HYPOTHESIS 3

There is a positive relationship from income inequality to corruption which is dependent on trust

In contrast to inconclusive results in models 1 and 2, model 3 reveals a statistically significant correlation from income inequality to corruption, in both high- and low-trust countries. In low-trust countries the interaction term has a weak but statistically significant correlation to corruption. Because the interaction term is included, the variable Income Inequality captures the variance that the interaction term does not (see section 6.7) and reveals a statistically significant moderate correlation to corruption in high-trust countries. The p-values for inequality in the first two models indicated a very low possibility that inequality is correlated to corruption. This changes in model 3, where the interaction term and inequality are statistically significant. Low p-values are expected for the constitutive terms of the interaction term and this is especially true here. As mentioned above, the scatterplot in 10.6.3 reveals that there are two variances for Income Inequality that are pulling in different directions – depending on the level of trust. By separating the variance in two, the p-values for Income Inequality and the interaction term reflect the statistical significance of inequality in these two groups of countries.

To further interpret the interaction term, I produced a graph (see section 10.8) that illustrates the interaction term's marginal effects and substantial significance. I did this in the statistical program Stata with the "grinter" command[25]. It shows the marginal effect of the interaction term on corruption on the vertical axis with values of the conditioning variable[26], trust, on the horizontal axis. The angle of the slope reveals that the effect is not substantively significant, and because the upper and lower bounds of the confidence interval cross the zero value of the vertical axis and are not, except for trust when it is zero, fully above or below, the marginal effect of the interaction term is not statistically significant. However, because the interaction term is significant at zero, we can say that income inequality in high-trust countries is statistically significant (Berry et al., 2012).

The most interesting result, however, is the direction of the sign. The interaction term reveals that income inequality leads to *lower* levels of corruption in low-trust countries, while decreasing corruption in high-trust countries. In other words, H3 is only partially rejected. There is a correlation from inequality to corruption and it depends on trust, but the correlation is not only positive - in this study. Depending on the level of trust, the correlation is either positive or negative. A partial rejection is still a rejection and we can conclude that the data does not give grounds to reject the null-hypothesis of H3.

Discussion and theoretical implications:

While the interaction term highlighted a link between inequality and corruption through trust, the results contradicted much of the literature. The inequality trap argues that income inequality leads to lower levels of generalized trust, leading to more corruption, which in turn leads to higher income inequality and less trust, creating a loop - or trap. The interaction term was hypothesized to support this theory through a positive correlation to corruption, in that inequality in both low- and high-trust societies correlates to more corruption - and therefore a lower score

---

[25] I was not able to find a simple and similar method of graphically illustrating interactions in SPSS. The variables used in Stata were the same as those used in SPSS. A check of regression ouputs revealed the results were the same.
[26] Although I consider the dichotomous dummy variable for trust to be *the* conditioning variable, its effect also varies with the value of inequality. This is due to the inherent symmetry of interactions (Berry et al., 2012; Brambor et al., 2005).

on the corruption scale. However, what the data reveals are two distinctly different correlations: positive in high-trust countries and negative in low-trust countries.

Countries with a high level of trust have a stronger sense of unity and a belief in a bright future (Rothstein & Uslaner, 2005). Thus, one could theorize, higher inequality leads to a stronger sense of economic disenfranchisement than in a low-trust society, which in turn leads to a lower threshold for violating the laws and participating in corrupt practices.

Considering GDP per capita's correlation to corruption, income inequality and trust, as seen in the regression and in the scatterplots in section 10.2 and 10.7, it is an important part of the dynamic. But the correlations are different for the full-N variables and the variables only including countries in this regression. Inequality and GDP per capita's very low R^2 of .099 increases to .223 for the variables used in the regression. For these countries the correlation is more than twice as strong as the global average, significantly inhibiting the potential for generalizing the findings to a larger population. The difference in R^2 for Trust and GDP per capita is moderate, increasing from .291 to .365 in the variables used for the regression. In other words, for the countries in this regression economic development plays a larger role than what can be observed in the full-N sample. But despite the outsized role of economic development and the interaction term's weak unstandardized beta, the correlation is present and significant.

Corruption is expensive. In low-trust societies higher levels of inequality also mean lower economic development. This means that while many poor people might be too poor to participate, the lower number of richer people simply do not have to participate, or at least they do not contribute to a high level of perceived corruption.

The dynamic that we see in low-trust societies might be partially explained by looking at literature on regional variations. Despite the theoretical and empirical support mentioned in section 3.3.3.1, my data provided only some support for these findings. Two separate regressions, one excluding countries with a socialist legal origin (see section 10.5.2) and the other excluding Eastern European and ex-Soviet countries (see section 10.5.3), reveal a stronger coefficient for Income Inequality and a very weak GDP per capita. Although these results are

interesting and should be investigated further, they do not significantly challenge the findings of the regression in chapter 7 and are – beyond this mention – outside the scope of this paper.

The interaction term's negative correlation is weak and should be interpreted in light of the many weaknesses with the study; the most corrupt countries are underrepresented; there are regional imbalances in the country selection; corruption is not a direct measure of corruption, but is the perception of corruption among survey respondents; data for most of the variables is not from a single year, but from several years. Should any of these issues be improved, or modified, the B of 0.01 will change – in either direction. Yet the coefficient stands, and the dynamic of the inequality trap is both supported and challenged by the findings of this thesis.

9 CONCLUSION

In this thesis I have investigated the link between income inequality, generalized trust and corruption. The goal was to investigate underlying socio-political determinants of corruption. To answer the research question *Does inequality have an effect on corruption and is this effect dependent on trust*, I created three hypotheses that investigated different aspects of the research question. H1 failed and H2 rejected its null-hypothesis. The last one, H3, can also be considered the main hypothesis: *There is a positive relationship from income inequality to corruption which is dependent on trust.* What I found was that H3 was only partly true. The interaction term provides counter-intuitive results, contradicting much of the literature on the inequality trap and highlights the need for additional investigations. As to whether this invalidates the underlying theories of the thesis, the answer is no; failing to prove a positive is not sufficient evidence for a negative. But my findings do question the universality of the inequality trap. However, the results show that inequality is more than an aspect of economic development and is an important socio-economic variable in studies on corruption and trust.

The lack of evidence supporting H3, generated from the inequality trap and Rothstein and Uslaner's stickiness model, indicates that their importance in explaining persistent corruption may be overstated or misunderstood. Ideally, this should be further investigated through longitudinal studies drawing on panel data.

The theories supporting my hypotheses have therefore not been disproven, I have only demonstrated that with the data in my regression and the method of analysis there is insufficient grounds to accept them all.

9.1 SUGGESTIONS FOR FURTHER RESEARCH

In the process of writing this thesis there have been many interesting variables and interactions that I would like to have investigated further, but due to resources, capacity and space I did not.

In terms of data, improvements in data availability will enable time-series analysis, and improvements in data quality, especially measurements for trust, corruption and different kinds of inequality, will benefit future studies greatly. Income inequality as measured through a Gini index is only one measurement and any results should be considered in light of this. As more and different types of data become available and more research is being done, other kinds of inequality can reveal new insights – for example age- and education specific measure of income. A good example is the skewness of the distribution of income; You, Gould and Morris all found that inequality at the bottom of the income distribution leads to lower trust than inequality at the top or middle. Additionally, measures of inequality do not measure the perception of inequality, which should correlate to the level of trust better than a simple Gini index. Studies including these kinds of inequality will provide interesting insights to the dynamic inequality-trust-corruption.

More studies on regional and country-specific variations would be of great added value to the field. The multidimensionality and conditional complexity of the variables and their relationships necessitate not only further quantitative studies, but more qualitative studies.

10 APPENDIX A: CHARTS, GRAPHS, SYNTAX AND OTHER OUTPUT

10.1 COUNTRIES IN REGRESSION

N=73

Argentina	Ecuador	Latvia	Romania
Austria	Egypt	Liberia	Russia
Belgium	El Salvador	Lithuania	Rwanda
Bolivia	Estonia	Luxembourg	Slovakia
Brazil	Finland	Madagascar	Slovenia
Bulgaria	France	Malaysia	Spain
Burundi	Georgia	Mauritius	Sweden
Cameroon	Germany	Mexico	Switzerland
Canada	Ghana	Morocco	Togo
Chile	Greece	Mozambique	Turkey
China	Guatemala	Namibia	Uganda
Colombia	Guinea	Netherlands	United Kingdom
Costa Rica	Honduras	Nicaragua	United States
Cote d'Ivoire	Hungary	Niger	Uruguay
Croatia	Iraq	Norway	Yemen
Cyprus	Ireland	Pakistan	Zambia
Czech Republic	Israel	Panama	
Denmark	Italy	Peru	
Dominican Republic	Korea, South	Portugal	

According to region:

Region	Count
Western Europe and North America	19
Latin America	16
Africa, sub-Saharan	15
Eastern Europe and Ex-Soviet states	12
Middle East and North Africa	7
East Asia	2
South East Asia	1
South Asia	1
Pacific	None
The Caribbean	None

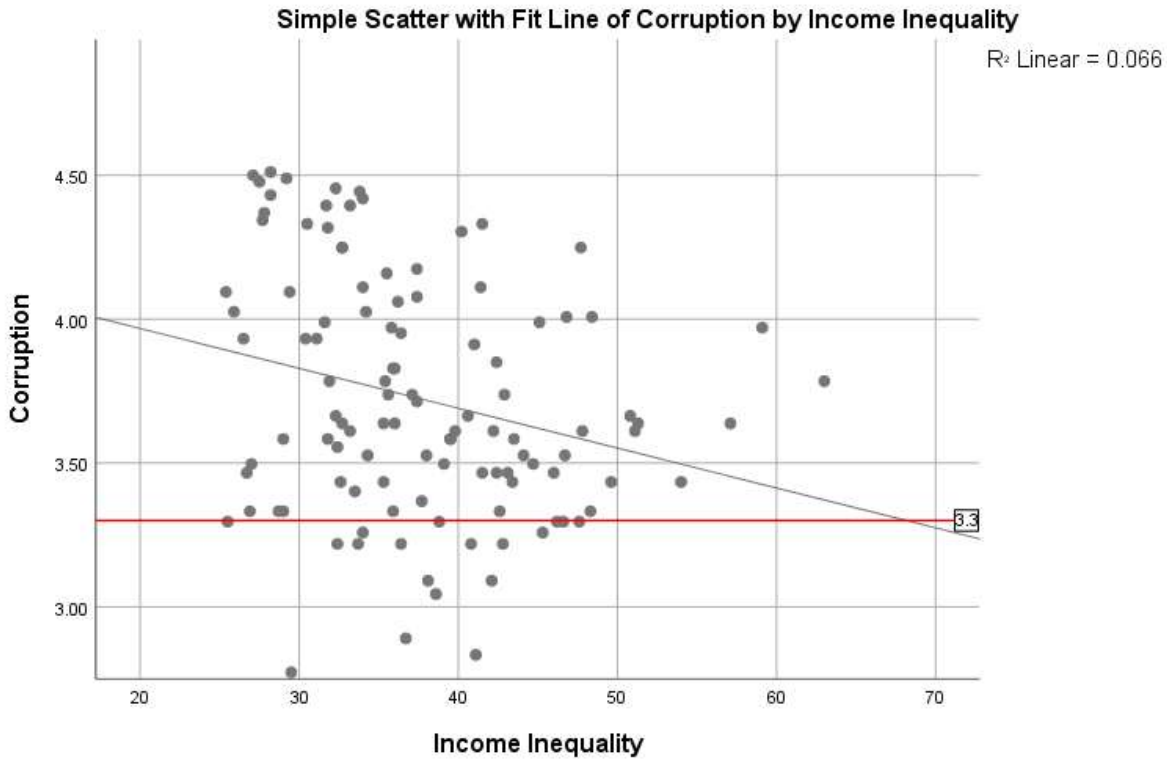## 10.2 ASSUMPTIONS

Assumption 1:
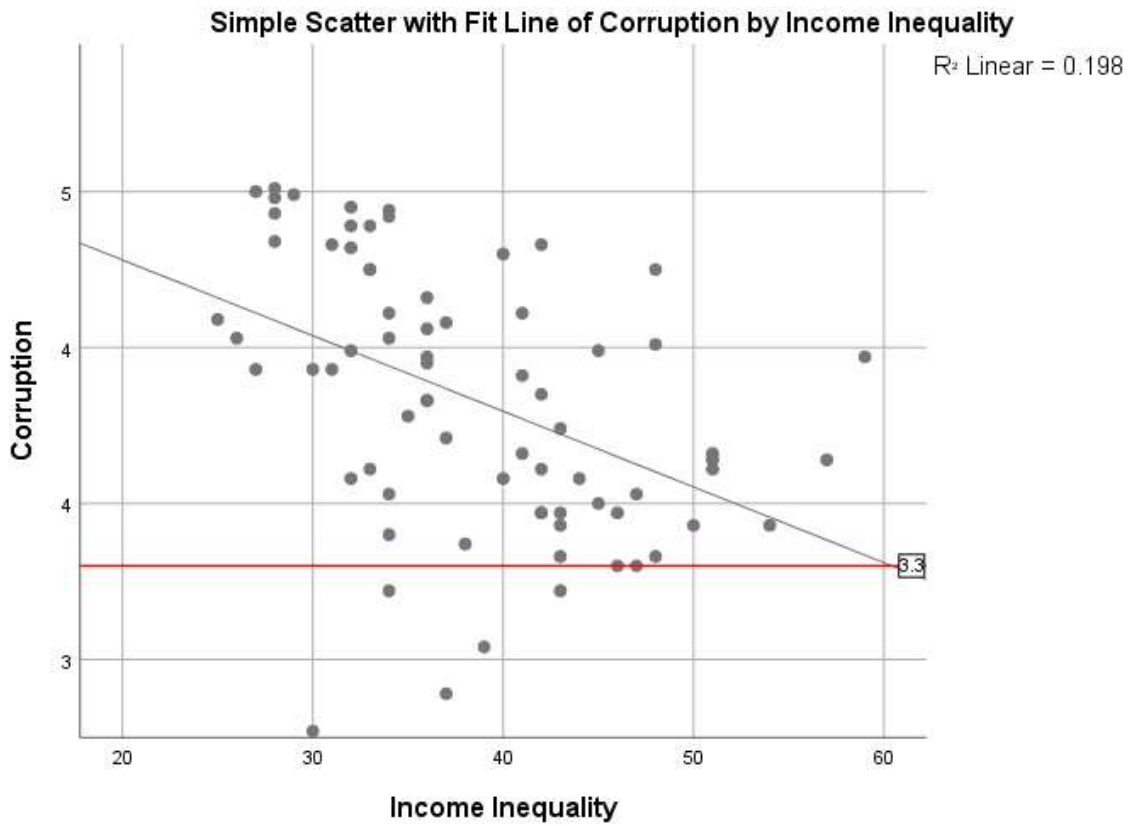

Figure 2: Full-N variables.

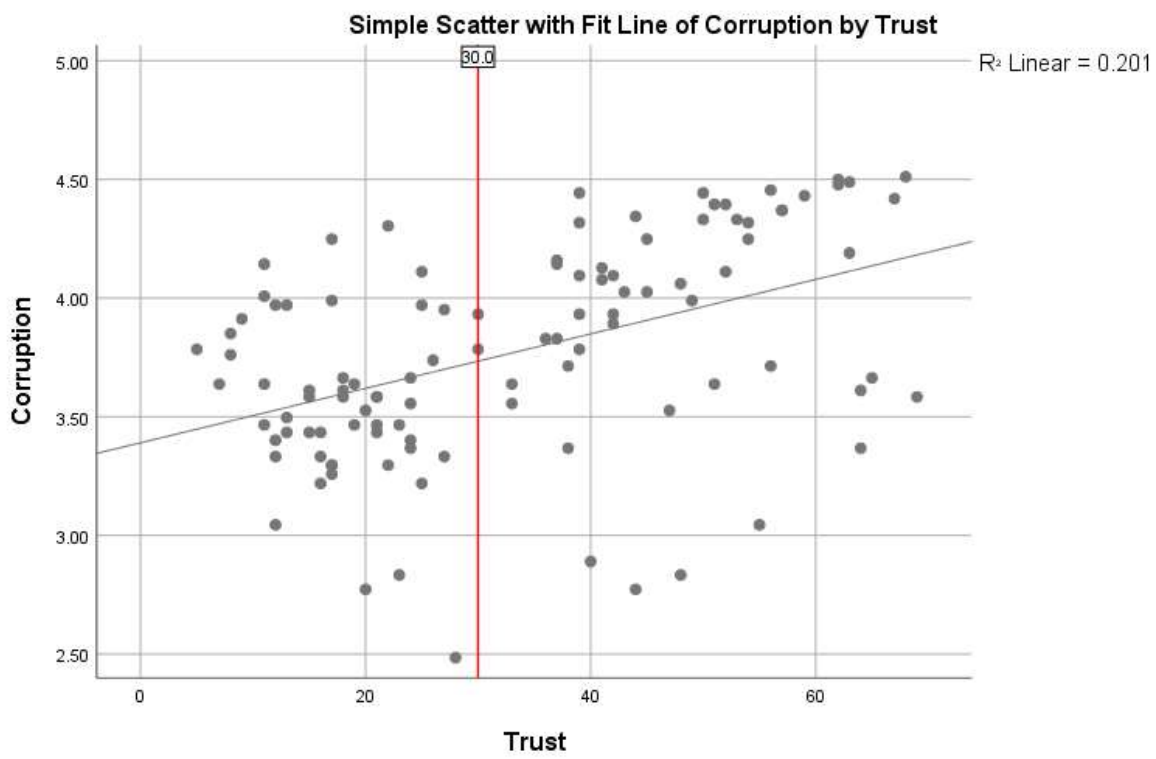*Figure 3: Scatterplot only including the 73 countries in the regression.*
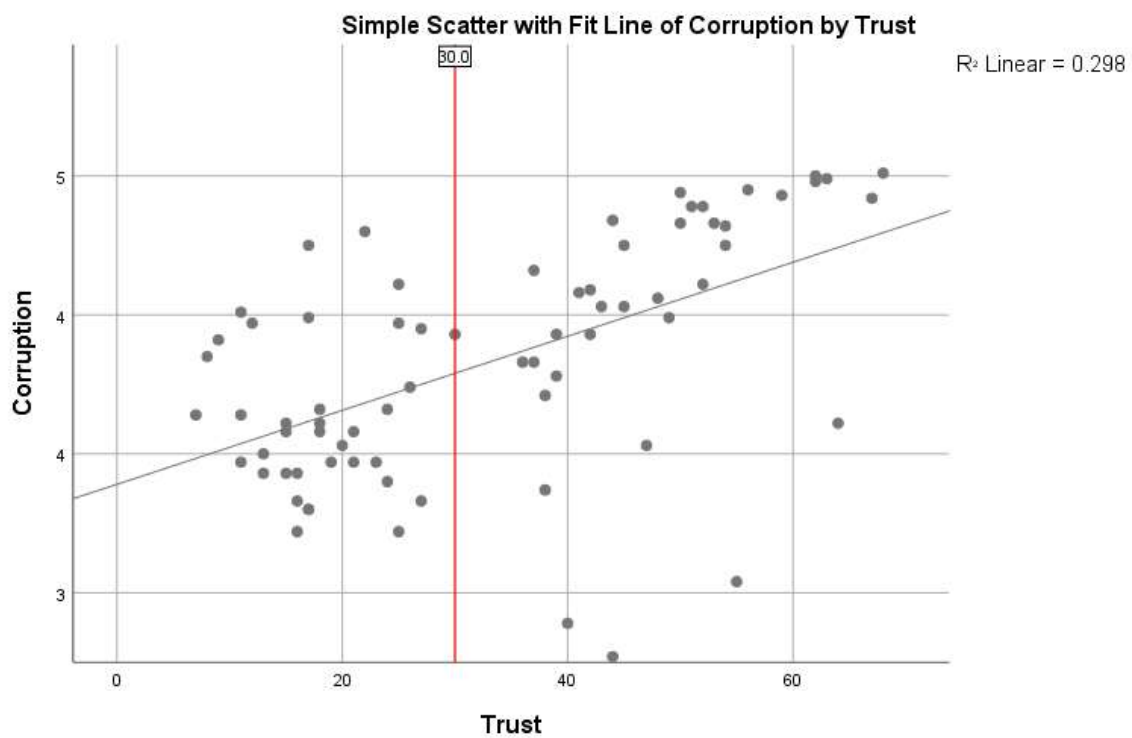

Figure 4: Full-N variables.

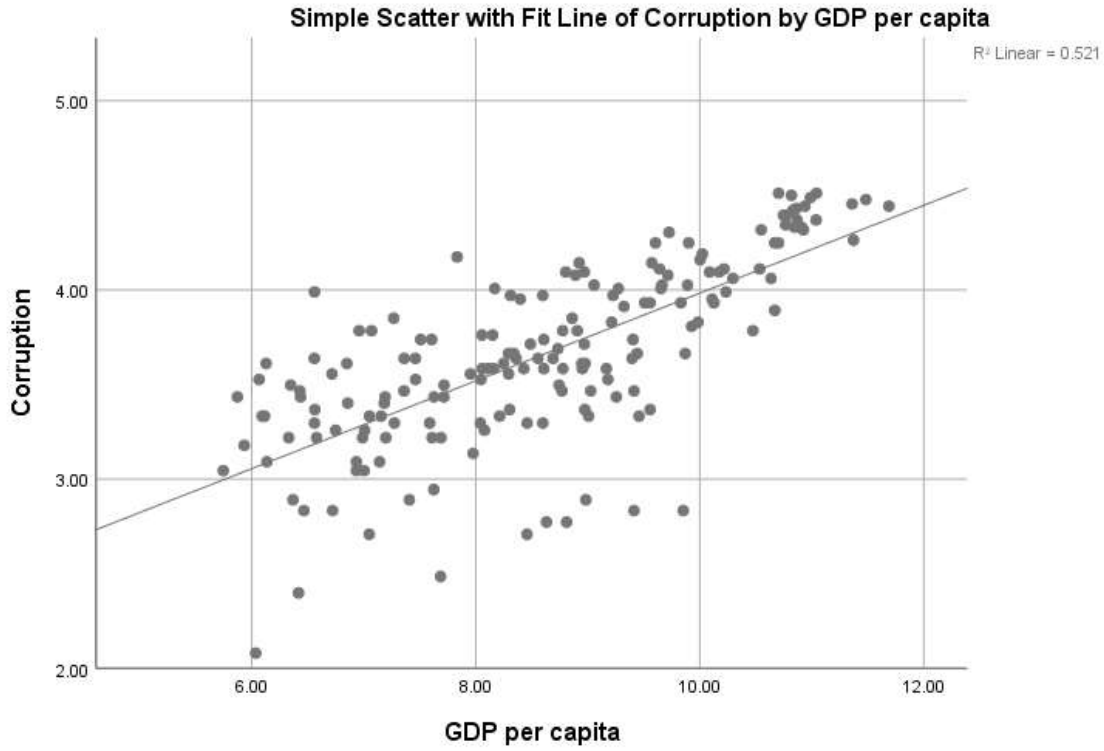*Figure 5: Scatterplot only including the 73 countries in the regression.*

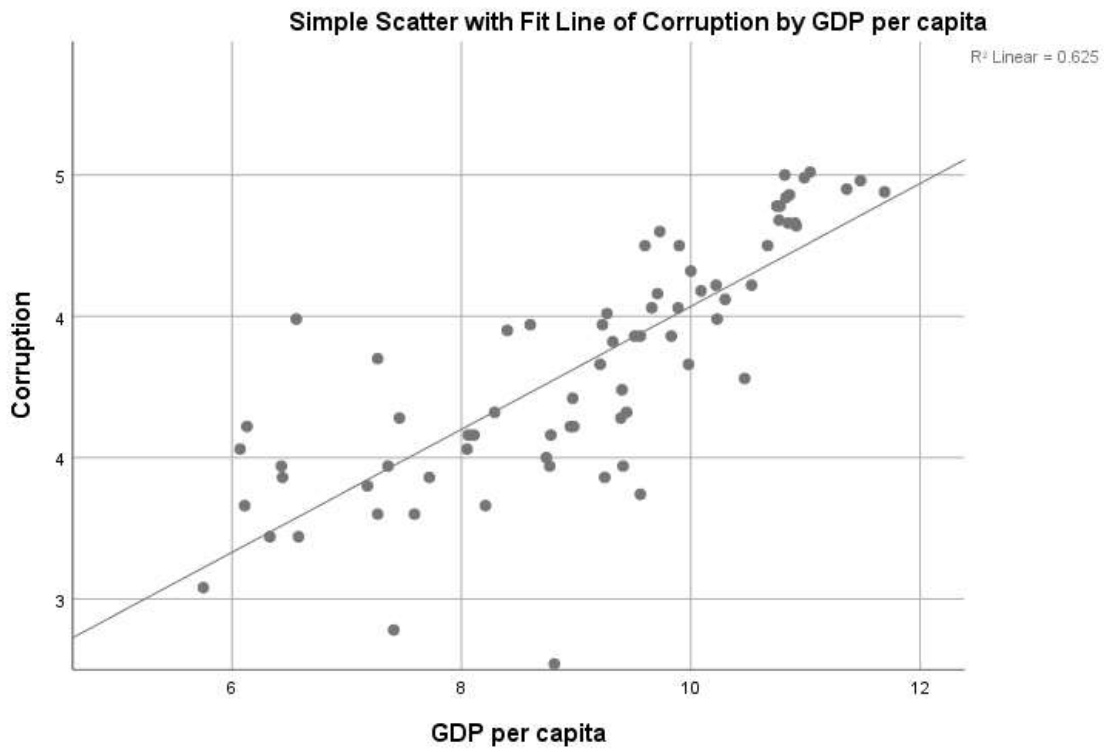*Figure 6: Both are the Natural Logarithm of their respective variables.*
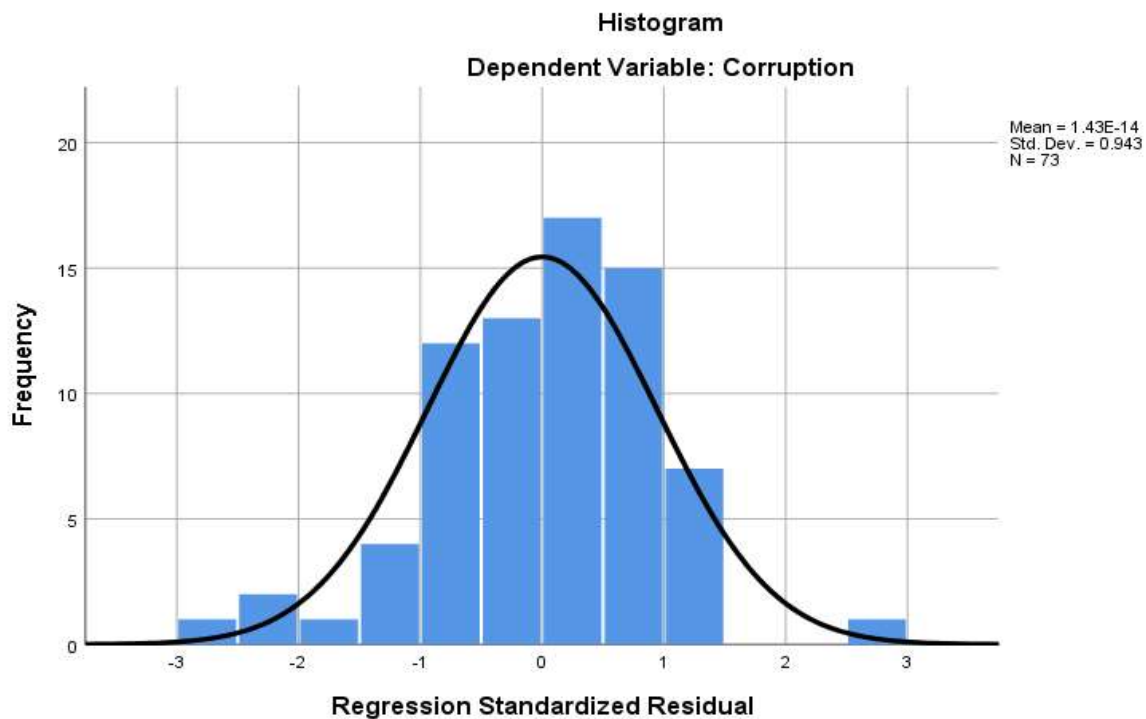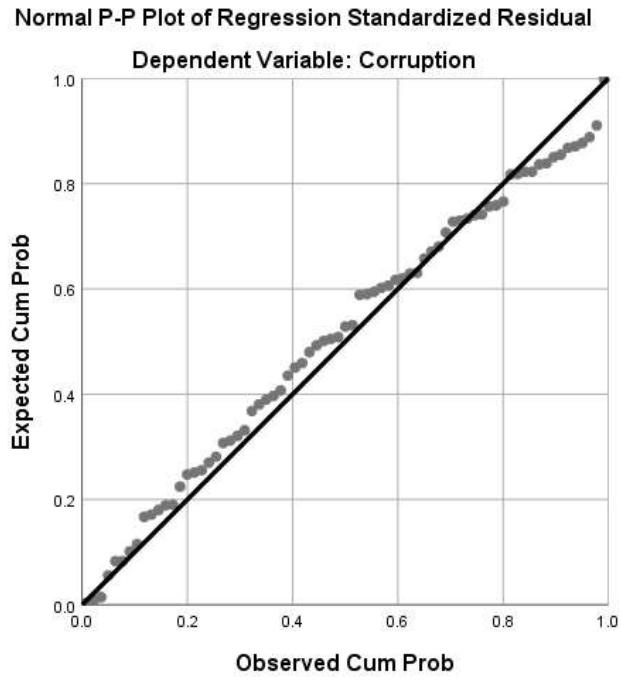

Figure 7: Scatterplot only including the 73 countries in the regression. Both are the Natural Logarithm of their respective variables.

Assumption 3:
REGRESSION
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT lnCorruption_2019
 /METHOD=ENTER Gini_2019 lnGDPpc dRuleofLaw_median dPolityIV dHDI dNatrr
 /METHOD=ENTER Trust_2019
 /METHOD=ENTER dTrustxInequality_2019
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).

Histogram

Dependent Variable: Corruption

Mean = 1.43E-14
Std. Dev. = 0.943
N = 73

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Corruption
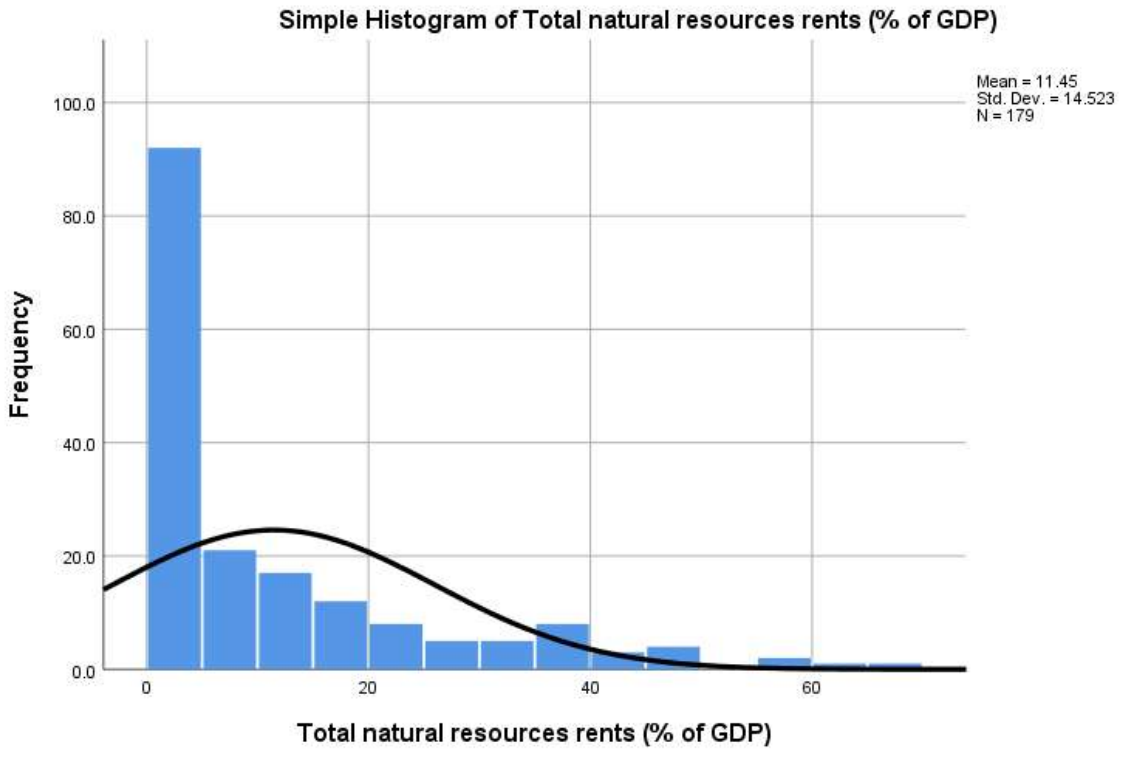
10.3 HISTOGRAM FOR NATURAL RESOURCES

Figure 8: Before LN

10.4 NEW VARIABLES

LN Corruption (2019)
COMPUTE lnCorruption_2019=LN(ti_cpi_2019).
EXECUTE.

LN GDP per capita
COMPUTE lnGDPpc=LN(GDPpc).
EXECUTE.

Dummy for low Trust by median (2019)
RECODE SocialTrust (Lowest thru 30=1) (30.1 thru Highest=0) INTO dTrust_2019.
VARIABLE LABELS dTrust_2019 'Trust (dummy) 2019'.
EXECUTE.

Interaction term for low Trust and Income Inequality (2019)
COMPUTE TrustxInequality_2019=dTrust_2019 * Gini_2019.
EXECUTE.

Dummy for Low Rule of Law by median
RECODE wbgi_rle (Lowest thru -.23=1) (-.22 thru Highest=0) INTO dRuleofLaw_median.
VARIABLE LABELS dRuleofLaw_median 'Rule of Law'.
EXECUTE.

Dummy for full democracy from polity IV
RECODE p_polity (Lowest thru 6=0) (7 thru Highest=1) INTO dPolityIV.
VARIABLE LABELS dPolityIV 'Democracy'.
EXECUTE.

Low HDI dummy
RECODE HDI (Lowest thru 0.69=1) (0.7 thru Highest=0) INTO dHDI.
VARIABLE LABELS dHDI 'HDI'.
EXECUTE.

Dummy for Natural Resources at 25 % of GDP
RECODE wdi_natrr (25 thru Highest=1) (Lowest thru 24=0) INTO dNatrr.
VARIABLE LABELS dNatrr 'Natural Resources'.
EXECUTE.

10.5 OTHER REGRESSIONS

10.5.1 Regression of three corruption variables

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.855[a]	.731	.728	.25603

a. Predictors: (Constant), Political corruption index, Public sector corruption index

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	29.266	2	14.633	223.224	.000[b]
	Residual	10.751	164	.066		
	Total	40.017	166			

a. Dependent Variable: Corruption

b. Predictors: (Constant), Political corruption index, Public sector corruption index

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.347	.041		107.234	.000		
	Public sector corruption index	-.823	.213	-.503	-3.861	.000	.096	10.382
	Political corruption index	-.622	.224	-.362	-2.777	.006	.096	10.382

a. Dependent Variable: Corruption

10.5.2 Regression excluding countries with Socialist Legal Origin

Model 1: .788
Model 2: .803
Model 3: .816

Model		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error	Beta	t	Sig.	VIF
1	(Constant)	4.086	.553		7.382	.000	
	Income Inequality	-.014	.005	-.237	-2.810	.007	1.886
	GDP per capita	.048	.045	.184	1.077	.287	7.710
	Democracy	.099	.079	.105	1.245	.219	1.868
	Rule of Law	-.432	.084	-.488	-5.165	.000	2.360
	HDI	-.011	.117	-.012	-.091	.928	4.470
	Natural Resources	-.337	.131	-.196	-2.584	.013	1.525
2	(Constant)	4.025	.534		7.530	.000	
	Income Inequality	-.008	.006	-.131	-1.377	.175	2.564
	GDP per capita	.010	.047	.038	.211	.833	8.986
	Democracy	.118	.077	.125	1.529	.133	1.892
	Rule of Law	-.436	.081	-.492	-5.405	.000	2.361
	HDI	-.047	.114	-.052	-.411	.683	4.567
	Natural Resources	-.377	.127	-.219	-2.964	.005	1.557
	Trust	.005	.002	.222	2.183	.034	2.950
3	(Constant)	3.849	.523		7.363	.000	
	Income Inequality	-.014	.006	-.232	-2.244	.030	3.253
	GDP per capita	.012	.045	.047	.272	.787	8.992
	Democracy	.117	.074	.124	1.573	.122	1.892
	Rule of Law	-.470	.080	-.531	-5.912	.000	2.463
	HDI	-.041	.110	-.045	-.370	.713	4.570
	Natural Resources	-.342	.124	-.199	-2.754	.008	1.585
	Trust	.012	.004	.483	3.072	.004	7.541
	Trust (dummy) x Inequality	.008	.004	.387	2.125	.039	10.096

10.5.3 Regression excluding Eastern European and ex-Soviet countries.

Model 1: .774
Model 2: .785
Model 3: .816

Model		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error	Beta	t	Sig.	VIF
1	(Constant)	3.760	.529		7.110	.000	
	Income Inequality	-.010	.005	-.164	-2.105	.040	1.616
	GDP per capita	.066	.044	.241	1.495	.141	6.903
	Democracy	.095	.077	.101	1.240	.220	1.750
	Rule of Law	-.467	.082	-.518	-5.673	.000	2.207
	HDI	.008	.111	.009	.072	.943	3.888
	Natural Resources	-.261	.134	-.144	-1.952	.056	1.447
2	(Constant)	3.765	.515		7.310	.000	
	Income Inequality	-.005	.005	-.078	-.891	.377	2.145
	GDP per capita	.026	.047	.094	.540	.591	8.441
	Democracy	.136	.078	.145	1.759	.084	1.885
	Rule of Law	-.485	.081	-.537	-6.004	.000	2.232
	HDI	-.019	.109	-.021	-.177	.860	3.950
	Natural Resources	-.287	.131	-.158	-2.190	.033	1.461
	Trust	.005	.002	.193	1.985	.052	2.643
3	(Constant)	3.687	.477		7.726	.000	
	Income Inequality	-.015	.006	-.253	-2.574	.013	3.156
	GDP per capita	.019	.044	.069	.426	.672	8.462
	Democracy	.135	.072	.143	1.876	.066	1.885
	Rule of Law	-.509	.075	-.563	-6.775	.000	2.256
	HDI	-.037	.101	-.041	-.368	.714	3.963
	Natural Resources	-.274	.121	-.151	-2.262	.028	1.463
	Trust	.014	.004	.558	3.800	.000	7.032
	Trust (dummy) x Inequality	.011	.003	.536	3.145	.003	9.475

10.6 DESCRIPTIVE STATISTICS

10.6.1 Corruption

Statistics

Corruption

N	Valid	178
	Missing	15014
Mean		3.6488
Median		3.6376
Std. Deviation		.49083
Range		2.43
Minimum		2.08
Maximum		4.51

Statistics

Corruption

N	Valid	73
	Missing	0
Mean		3.8347
Median		3.8500
Std. Deviation		.42363
Range		1.74
Minimum		2.77
Maximum		4.51

10.6.2 Trust

Statistics

Trust

N	Valid	109
	Missing	15083
Mean		33.49
Median		30.00
Std. Deviation		17.757
Range		64
Minimum		5
Maximum		69

Statistics

Trust

N	Valid	73
	Missing	0
Mean		33.36
Median		30.00
Std. Deviation		17.369
Range		61
Minimum		7
Maximum		68

10.6.3 Income Inequality

Statistics		
Income Inequality		
N	Valid	114
	Missing	15078
Mean		37.67
Median		36.40
Std. Deviation		7.640
Range		38
Minimum		25
Maximum		63

Statistics		
Income Inequality		
N	Valid	73
	Missing	0
Mean		38.38
Median		37.00
Std. Deviation		7.776
Range		34
Minimum		25
Maximum		59

10.7 ADDITIONAL SCATTERPLOTS

10.7.1 Income Inequality and Trust

Figure 9: With only the 73 countries in the regression. Red line is median of Trust.

10.7.2 Trust and GDP per capita
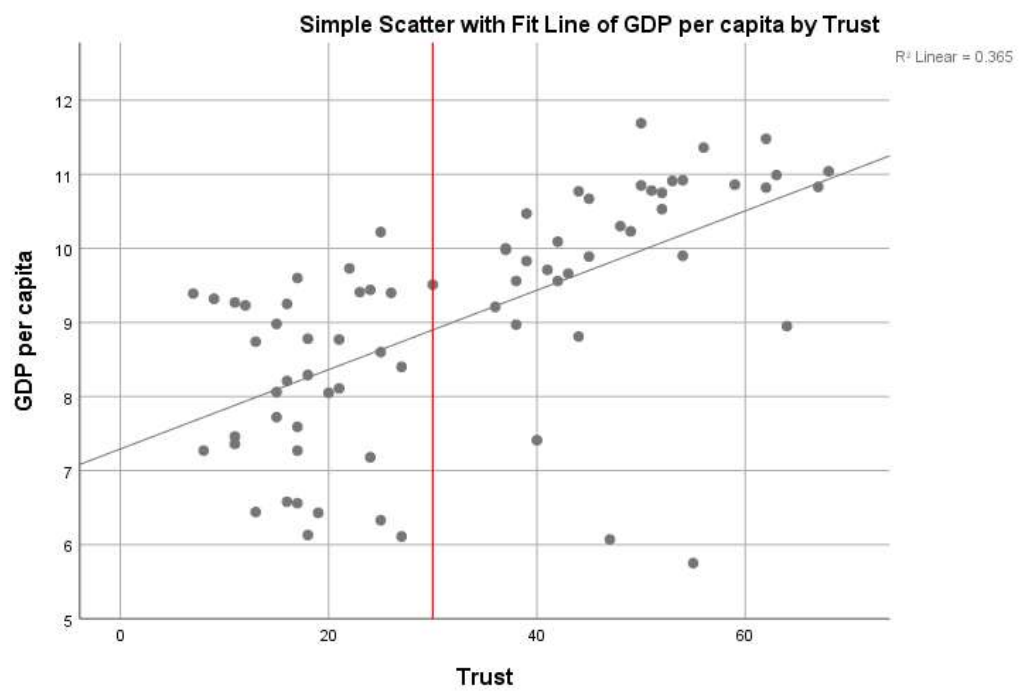
Figure 10: With only the 73 countries in the regression. Red line is median of Trust.
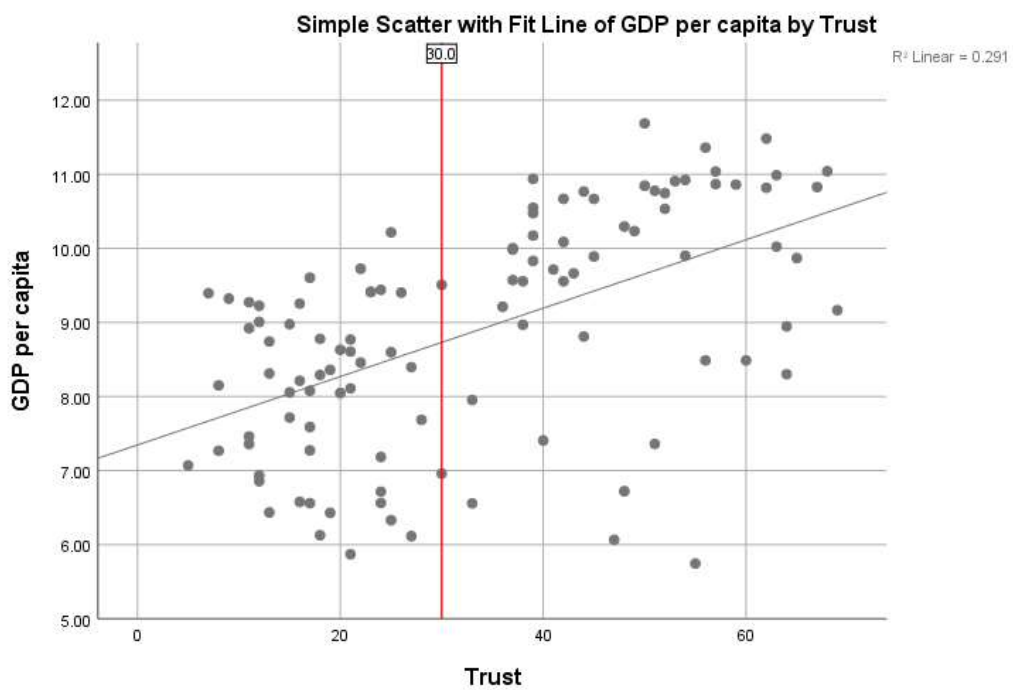
Figure 11: Full-N variables. Red line is median of Trust.
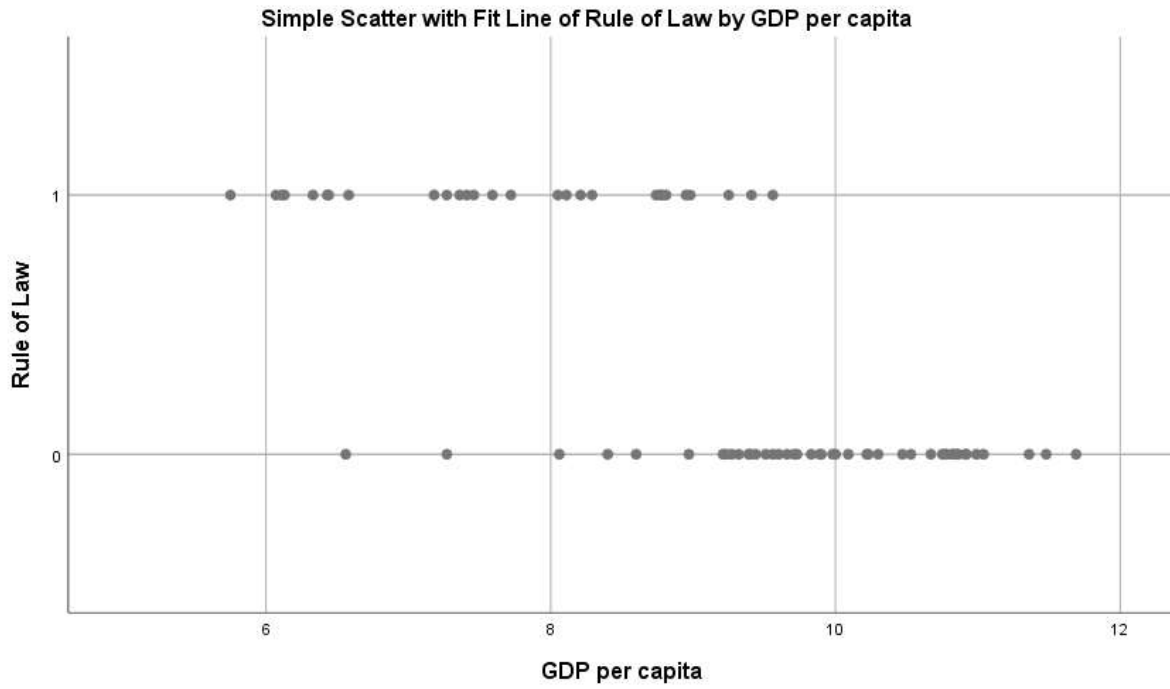
10.7.3 Income Inequality and GDP per capita
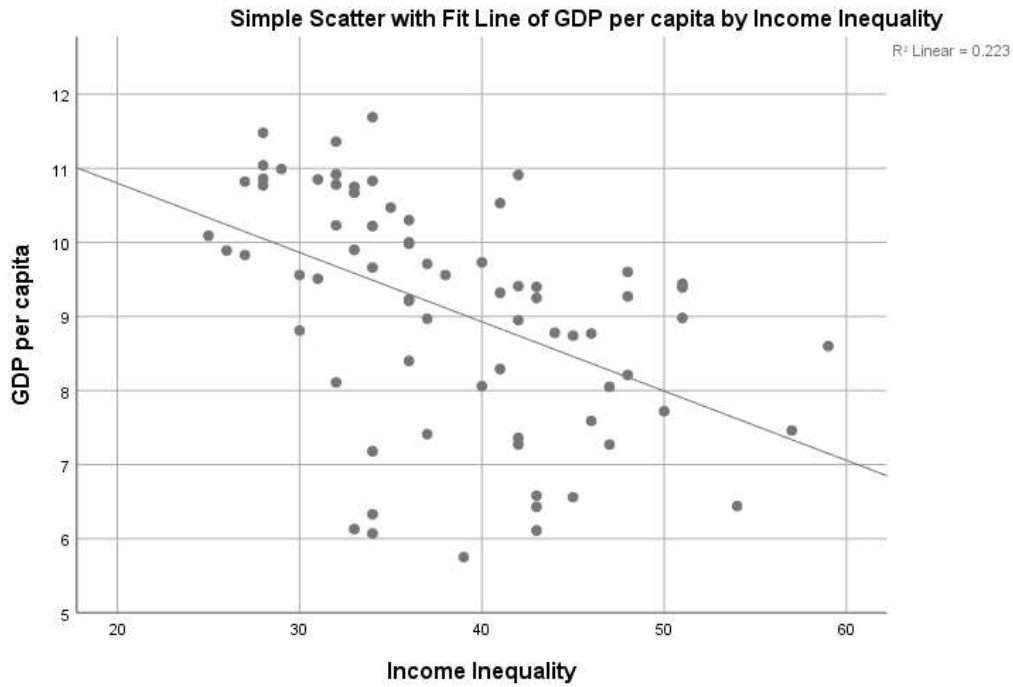
Figure 12: With only the 73 countries in the regression.
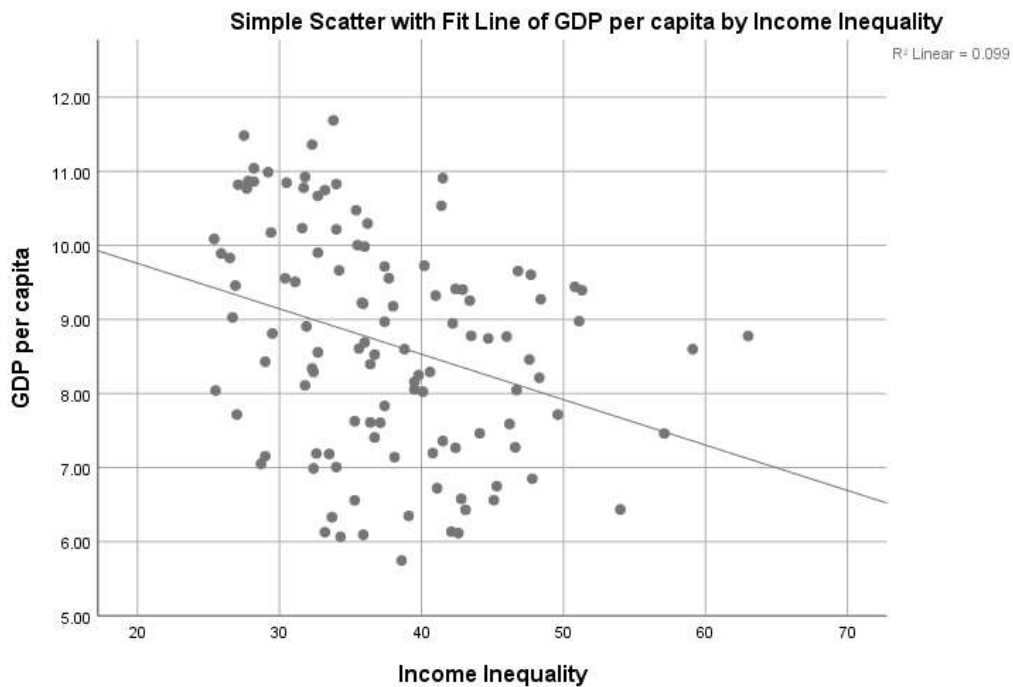
Figure 13: Full-N variables.

10.7.4 Trust and Democracy
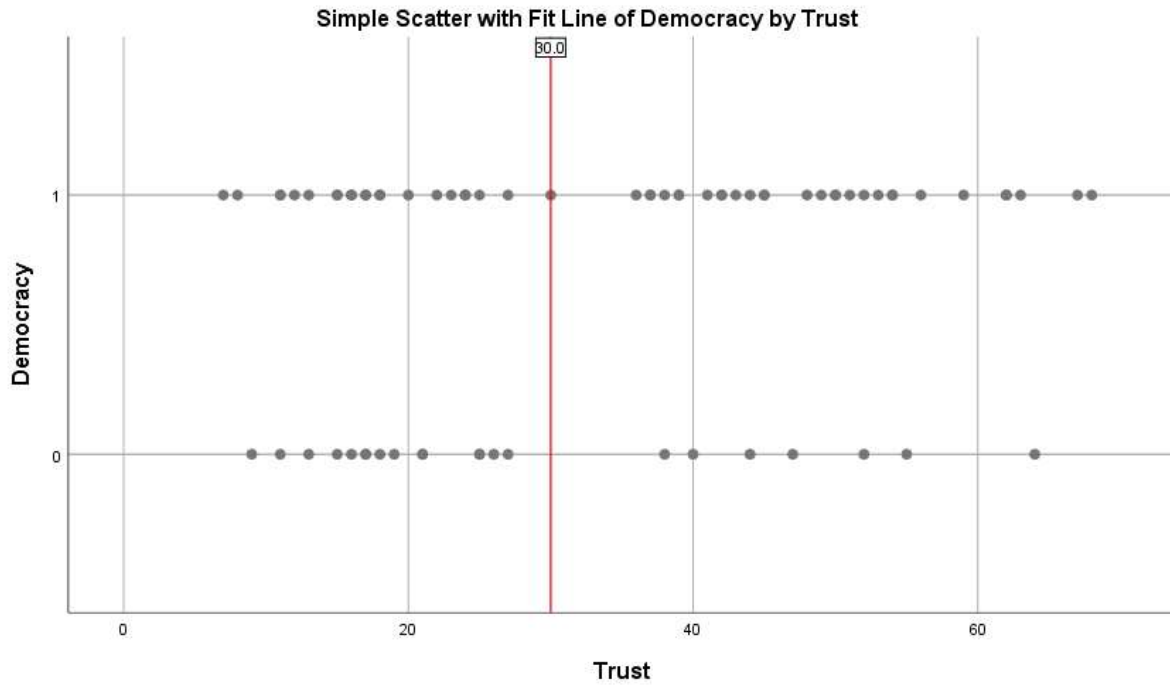
Figure 14: With only the 73 countries in the regression.

10.7.5 Trust and HDI
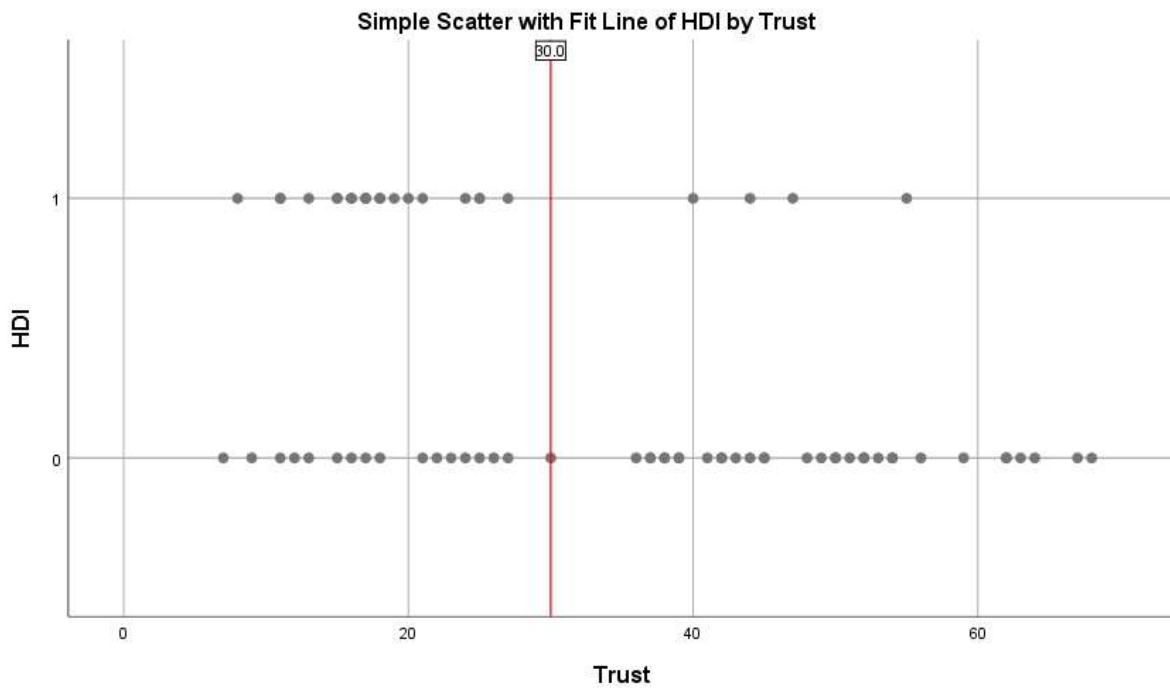
Figure 15: With only the 73 countries in the regression.

10.7.6 GDP per capita and Rule of Law

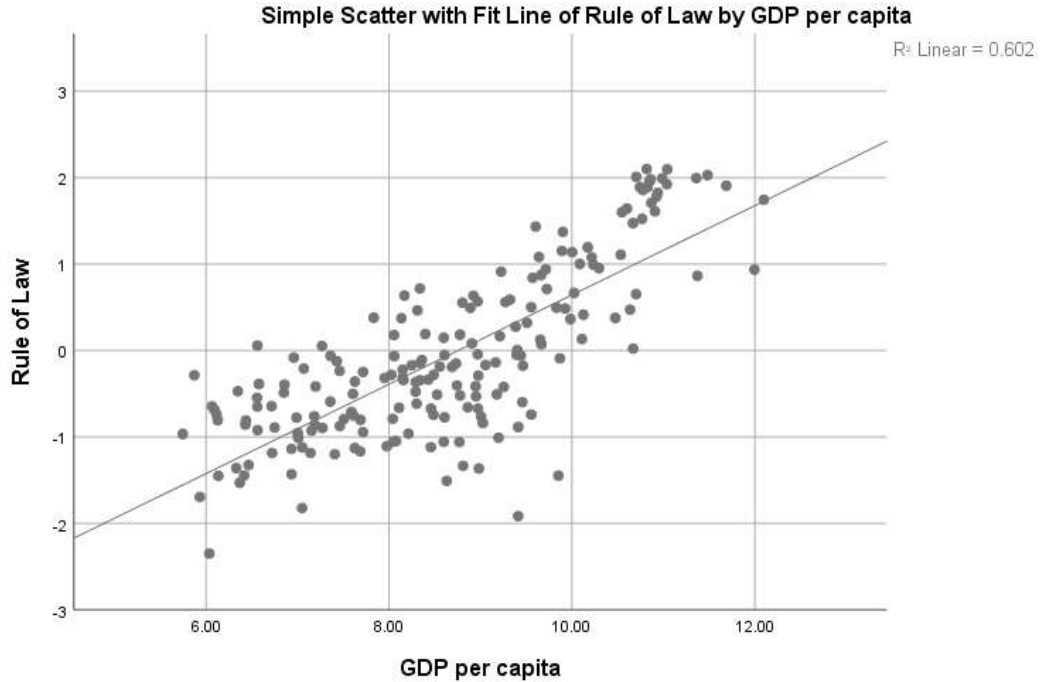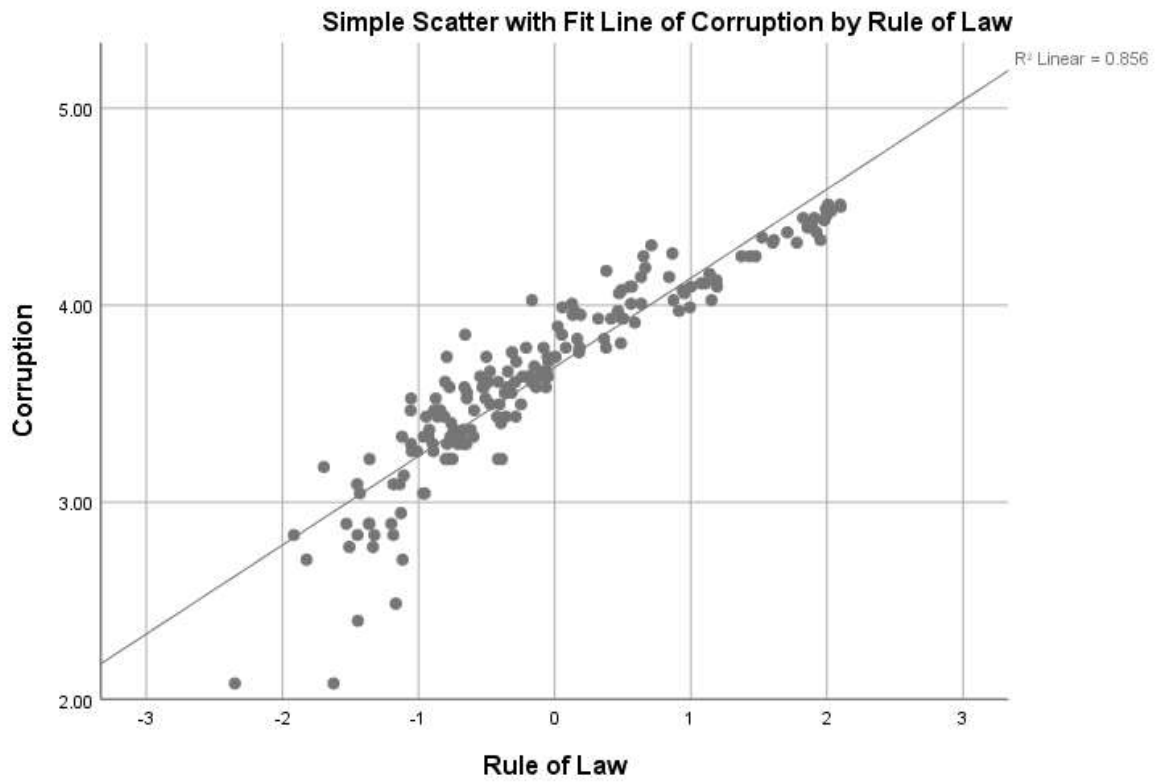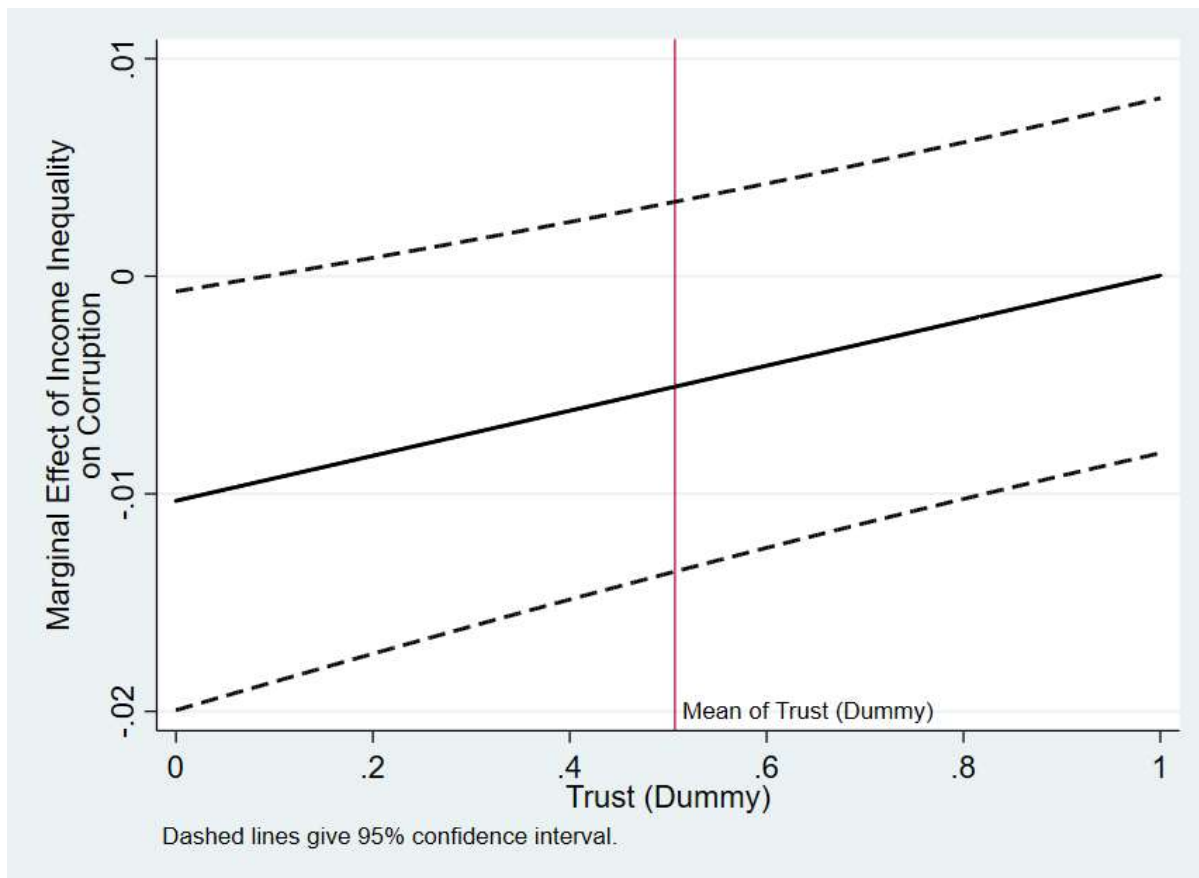*Figure 16: With only the 73 countries in the regression.*

Figure 17: Full-N original variables.

10.7.7 Corruption and Rule of Law

Simple Scatter with Fit Line of Corruption by Rule of Law

10.8 GRAPHICAL ILLUSTRATION OF INTERACTIONS

Command in Stata:

grinter Gini, inter(dTrustxInequality) const02(dTrust) depvar(lnCorruption)

11 APPENDIX B

11.1 CPI MEASUREMENTS

The source data captures the following aspects of corruption:
- Bribery
- Diversion of public funds
- Prevalence of officials using public office for private gain without facing consequences
- Ability of governments to contain corruption and enforce effective integrity mechanisms in the public sector
- Red tape and excessive bureaucratic burden which may increase opportunities for corruption
- Meritocratic versus nepotistic appointments in the civil service
- Effective criminal prosecution for corrupt officials
- Adequate laws on financial disclosure and conflict of interest prevention for public officials
- Legal protection for whistleblowers, journalists, investigators when they are reporting cases of bribery and corruption
- State capture by narrow vested interests
- Access of civil society to information on public affairs

(Transparency International, 2018b, p. 2)

The source data does not capture the following manifestations of corruption:

- Citizen perceptions or experience of corruption
- Tax fraud
- Illicit financial flows
- Enablers of corruption (lawyers, accountants, financial advisors etc)
- Money-laundering
- Private sector corruption
- Informal economies and markets

(Transparency International, 2018b, p. 2)

12 BIBLIOGRAPHY

A.D. (2015). Ordinary Least Squares (OLS). *Economic Theory Blog.* Retrieved from
https://economictheoryblog.com/contents-ordinary-least-squares/ [November 14th, 2018]

Acemoglu, D., & Robinson, J. A. (2002). The political economy of the Kuznets curve. *Review of development economics, 6*(2), 183-203.

Afonso, H., LaFleur, M., & Alarcón, D. (2015). Concepts of Inequality. Retrieved from
https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/dsp_policy_01.pdf

Allison, P. (2012). When Can You Safely Ignore Multicollinearity? Retrieved from
https://statisticalhorizons.com/multicollinearity [March 19th, 2019]

Ariely, G., & Uslaner, E. M. (2017). Corruption, fairness, and inequality. *International Political Science Review, 38*(3), 349-362. doi:10.1177/0192512116641091

Ban Ki-Moon. (2012). Secretary-General's closing remarks at High-Level Panel on Accountability, Transparency and Sustainable Development [Press release]. Retrieved from
https://www.un.org/sg/en/content/sg/statement/2012-07-09/secretary-generals-closing-remarks-high-level-panel-accountability

Batabyal, S., & Chowdhury, A. (2015). *Curbing Corruption, Financial Development and Income Inequality* (Vol. 15).

Berry, W. D., Golder, M., & Milton, D. (2012). Improving Tests of Theories Positing Interaction. *The Journal of Politics, 74*(3), 653-671. doi:10.1017/s0022381612000199

Boland, V. (2017). Ireland's 'de-globalised' data calculate a smaller economy. *Financial Times*. Retrieved from https://www.ft.com/content/dd3a6f1c-6aea-11e7-bfeb-33fe0c5b7eaa

Botsman, R. (2017). *Who can you trust? : how technology brought us together and why it might drive us apart*.

Brambor, T., Clark, W. R., & Golder, M. (2005). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis, 13*(1), 1-20. doi:10.1093/pan/mpi014

C. Kellogg, N. (2018). Top Effects of Poverty. Retrieved from https://borgenproject.org/5-effects-poverty/ [June 3rd, 2018]

Cederman, L.-E., Weidmann, N. B., & Gleditsch, K. S. (2011). Horizontal inequalities and ethnonationalist civil war: A global comparison. *American Political Science Review, 105*(3), 478-495.

Chetwynd, E., Chetwynd, F., & Spector, B. (2003). *Corruption and Poverty: A Review of Recent Literature*. Retrieved from https://www.semanticscholar.org/paper/Corruption-and-Poverty%3A-A-Review-of-Recent-Chetwynd-Chetwynd/1614ebee8a2931dec5c956bf083683239c98f9f4?tab=abstract:

Christophersen, K.-A. (2009). *Databehandling og statistisk analyse med SPSS*. [Oslo: Unipub.

Dabla-Norris, E., & Inchauste, G. (2008). Informality and Regulations: What Drives the Growth of Firms? *IMF STAFF PAPERS, 55*(1), 50-82.

Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F., & Tsounta, E. (2015). *Causes and Consequences of Income Inequality; A Global Perspective*. Retrieved from
https://EconPapers.repec.org/RePEc:imf:imfsdn:15/13

De Maio, F. G. (2007). Income inequality measures. *Journal of epidemiology and community health, 61*(10), 849-852. doi:10.1136/jech.2006.052969

Delhey, J., & Newton, K. (2005). Predicting Cross-National Levels of Social Trust: Global Pattern or Nordic Exceptionalism? *European Sociological Review, 21*(4), 311-327. doi:10.1093/esr/jci022

Delhey, J., Newton, K., & Welzel, C. (2011). How General Is Trust in "Most People"? Solving the Radius of Trust Problem. *American Sociological Review, 76*(5), 786-807. doi:10.1177/0003122411420817

Dutta, I., & Mishra, A. (2013). Does Inequality Foster Corruption? *Journal of Public Economic Theory, 15*(4), 602-619. doi:doi:10.1111/jpet.12027

Elster, J. (1989). Social Norms and Economic Theory. *Journal of economic perspectives, 3*(4), 99-117.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review, 97*(1), 75-90.

Ferguson, G. A. (2017). *Global Corruption : Law, Theory and Practice*. In. Retrieved from https://dspace.library.uvic.ca//handle/1828/9253

Gallego, A. (2016). Inequality and the erosion of trust among the poor: experimental evidence. *Socio-Economic Review, 14*(3), 443-460. doi:10.1093/ser/mww010

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring Trust. *The quarterly journal of economics, 115*(3), 811-846. doi:https://doi.org/10.1162/003355300554926

Goetsch, R. (2018). Testing Assumptions of Linear Regression in SPSS. Retrieved from https://www.statisticssolutions.com/testing-assumptions-of-linear-regression-in-spss/

Gould, E., & Hijzen, A. (2016). *Growing Apart, Losing Trust? The Impact of Inequality on Social Capital*. Retrieved from https://EconPapers.repec.org/RePEc:imf:imfwpa:16/176

Graf Lambsdorff, J. (2005). *Consequences and causes of corruption: What do we know from a cross-section of countries?* Retrieved from https://EconPapers.repec.org/RePEc:zbw:upadvr:v3405

Gujarati, D. N. (2011). *Econometrics by Example*. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan.

Gujarati, D. N., & Porter, D. C. (2010). *Essentials of Econometrics*. Boston (Mass.): McGraw-Hill Irwin.

Gupta, S., Davoodi, H., & Alonso-Terme, R. (1998). *Does corruption affect income inequality and poverty?* Washington (D.C.): International monetary fund.

Hardoon, D. (2017). An Economy for the 99%. *Oxfam Briefing Paper*, 76.130-121.

Inequality.org. (2018). Income Inequality. Retrieved from https://inequality.org/facts/income-inequality/#

Jetter, M., Agudelo, A. M., & Hassan, A. R. (2015). The Effect of Democracy on Corruption: Income is Key. *World Development, 74*, 286-304. doi:https://doi.org/10.1016/j.worlddev.2015.05.016

Jordahl, H. (2007). *Inequality and Trust*. Retrieved from https://ideas.repec.org/p/hhs/iuiwop/0715.html

Justesen, M. K., & Bjørnskov, C. (2014). Exploiting the Poor: Bureaucratic Corruption and Poverty in Africa. *World Development, 58*, 106-115. doi:http://dx.doi.org/10.2139/ssrn.2168119

Kelly, M. (2000). Inequality and crime. *Review of economics and Statistics, 82*(4), 530-539.

Kokko, H., Mackenzie, A., Reynolds, J. D., Lindström, J., Sutherland, W. J., & Associate Editor: Patsy, H. (1999). Measures of Inequality Are Not Equal. *The American Naturalist, 154*(3), 358-382. doi:10.1086/303235

Kolnes, V. L. (2016). *Foreign Direct Investment and Corruption.* (Master in Comparative Politics), University of Bergen, Retrieved from http://hdl.handle.net/1956/12818

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review, 45*(1), 1-28.

Laerd Statistics. (2015). Multiple regression using SPSS Statistics. Retrieved from https://statistics.laerd.com/

Lagarde, C. (2018). 'Steer, Don't Drift': Managing Rising Risks to Keep the Global Economy on Course. Retrieved from https://www.imf.org/en/News/Articles/2018/09/27/sp100118-steer-dont-drift

Lawson, M., & Martin, M. (2018). The Commitment to Reducing Inequality Index 2018: A global ranking of governments based on what they are doing to tackle the gap between rich and poor. doi:10.21201/2018.3415

Lučić, D., Radišić, M., & Dobromirov, D. (2016). Causality between corruption and the level of GDP. *Economic Research-Ekonomska Istraživanja, 29*(1), 360-379. doi:10.1080/1331677X.2016.1169701

Marshall, M. G. (2018). Polity IV Project: Dataset Users' Manual. In.

McKay, A. (2002). Defining and Measuring Inequality. *Inequality Briefing*. Retrieved from
https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3804.pdf

Midtbø, T. (2007). *Regresjonsanalyse for samfunnsvitere : med eksempler i SPSS*. Oslo:
Universitetsforlaget.

Morris, S. D., & Klesner, J. L. (2010). Corruption and Trust: Theoretical Considerations and Evidence From
Mexico. *Comparative Political Studies, 43*(10), 1258-1285.

Mungiu-Pippidi, A. (2011). Contextual choices in fighting corruption: Lessons learned.

Newton, K. (2007). Social and Political Trust. In *The Oxford Handbook of Political Behavior*.

OECD. (2012). *International Drivers of Corruption: A Tool for Analysis*. In. Retrieved from
http://dx.doi.org/10.1787/9789264167513-en

OECD. (2019). Gross national income (indicator). Retrieved from
https://data.oecd.org/natincome/gross-national-income.htm

Piketty, T. (2006). The Kuznets Curve: Yesterday and Tomorrow. In Abhijit Vinayak Banerjee, Roland
Bénabou, & Dilip Mookherjee (Eds.), *Understanding Poverty* (pp. 63-72).

Piketty, T. (2014). *Capital in the twenty-first century* (A. Goldhammer, Trans.).

Ringdal, K. (2013). *Enhet og mangfold : samfunnsvitenskapelig forskning og kvantitativ metode*. Bergen:
Fagbokforlaget.

Rock, M. T. (2009). Corruption and Democracy. *The Journal of Development Studies, 45*(1), 55-75.
doi:10.1080/00220380802468579

Rose-Ackerman, S. (2004). *The Challenge of Poor Governance and Corruption*. Retrieved from
Copenhagen Consensus: https://www.copenhagenconsensus.com/sites/default/files/CP+-
+Corruption+FINISHED.pdf

Rose-Ackerman, S. (2001). Trust and Honesty in Post-Socialist Societies. *Kyklos, 54*(2-3), 415-443.
doi:doi:10.1111/1467-6435.00161

Rothstein, B. (2011). *The quality of government corruption, social trust, and inequality in international
perspective*. Chicago, Ill.; London: The University of Chicago Press.

Rothstein, B. (2013). Corruption and social trust: Why the fish rots from the head down. *social research,
80*(4), 1009-1032.

Rothstein, B., & Uslaner, E. M. (2005). All for All: Equality, Corruption, and Social Trust. *World Politics,
58*(1), 41-72.

Sanz-Barbero, B., Vives-Cases, C., Otero-García, L., Muntaner, C., Torrubiano-Domínguez, J., & O'Campo,
y. P. (2015). Intimate partner violence among women in Spain: the impact of regional-level male
unemployment and income inequality. *The European Journal of Public Health, 25*(6), 1105-1111.

Smerdon, D., & Blauw, S. (2017). Trust and Inequality: Just Bad luck? , 40.

Søreide, T. (2014). *Drivers of Corruption : A Brief Review*. Washington, DC: World Bank: World Bank
Publications.

Stephany, F. (2017). Who are Your Joneses? Socio-Specific Income Inequality and Trust. *Social Indicators
Research, 134*(3), 877-898. doi:10.1007/s11205-016-1460-9

Teorell, J., Dahlberg, S., Holmberg, S., Rothstein, B., Pachon, N. A., & Svensson, R. (2018). *The Quality of
Government Standard Dataset, version Jan18*. Retrieved from: http://www.qog.pol.gu.se

Teorell, J., Dahlberg, S., Holmberg, S., Rothstein, B., Pachon, N. A., & Svensson, R. (2019). *The Quality of
Government Standard Dataset, version Jan19*. Retrieved from: http://www.qog.pol.gu.se

Teorell, J., Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Anna Khomenko, & Richard Svensson.
(2016). *The Quality of Government Standard Dataset, version Jan16.* Retrieved from:
http://www.qog.pol.gu.se

Transparency International. (2017). *Global Corruption Barometer 2017*.
    https://www.transparency.org/news/feature/global_corruption_barometer_citizens_voices_fro
    m_around_the_world: Transparency International.

Transparency International. (2018a). *Corruption Perception Index 2017* (24-Jan-2018 ed.).
    https://www.transparency.org/news/feature/corruption_perceptions_index_2017:
    Transparency International.

Transparency International. (2018b). Corruption Perceptions Index 2017: Frequently Asked Questions.

Transparency International. (2018c). What is Corruption? . Retrieved from
    https://www.transparency.org/what-is-corruption#define

Treisman, D. (2000). The causes of corruption: a cross-national study. *Journal of Public Economics, 76*(3),
    399-457. doi:https://doi.org/10.1016/S0047-2727(99)00092-4

Treisman, D. (2007). What Have We Learned About the Causes of Corruption from Ten Years of Cross-
    National Empirical Research? *Annual Review of Political Science, 10*(1), 211-244.
    doi:10.1146/annurev.polisci.10.081205.095418

Treisman, D. (2014). What does cross-national empirical research reveal about the causes of corruption?
    In *Routledge handbook of political corruption* (pp. 109-123): Routledge.

UNDP. (2018). Human Development Indices and Indicators: 2018 Statistical Update. In *Technical Notes*.

Ünver, M., & Koyuncu, J. Y. (2016). The Impact of Poverty on Corruption. *Journal of Economic Library,
    3*(4), 11. doi:10.1453/jel.v3i4.1096

Uslaner, E. M. (2002). The Moral Foundation of Trust. *SSRN Electronic Journal*. doi:10.2139/ssrn.824504

Uslaner, E. M. (2005). *The bulging pocket and the rule of law: Corruption, inequality, and trust.* Paper
    presented at the The Quality of Government: What It Is, How to Get It, Why It Matters, The
    Quality of Government Institute, Department of Political Science, Göteborg University,
    Göteborg, Sweden.

Uslaner, E. M. (2013a). Trust and corruption revisited: how and why trust and corruption shape each
    other. *Quality & Quantity, 47*(6), 3603-3608. doi:10.1007/s11135-012-9742-z

Uslaner, E. M. (2013b). Trust as an alternative to risk. *Public Choice, 157*(3/4), 629-639.

Wood, A. (2006). Correlating violence and socio-economic inequality: an empirical analysis. *Attacking
    the root causes of torture, poverty, inequality and violence*, 23-93.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. In (pp. 741). Retrieved
    from https://jrvargas.files.wordpress.com/2011/01/wooldridge_j-
    _2002_econometric_analysis_of_cross_section_and_panel_data.pdf

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach*. Australia; Mason, Ohio:
    South-Western Cengage Learning.

World Bank. (2016). *Poverty and Shared Prosperity 2016: Taking on Inequality*. Washington, DC: World
    Bank.

You, J.-S. (2006). *A comparative study of corruption, inequality, and social trust.* (3217945 Ph.D.),
    Harvard University, Ann Arbor. Retrieved from https://search-proquest-
    com.pva.uib.no/docview/305338679?accountid=8579 ProQuest Central; ProQuest Dissertations
    & Theses Global database.

You, J.-S., & Khagram, S. (2005). *A Comparative Study of Inequality and Corruption* (Vol. 70).

You, J. S. (2012). Social Trust: Fairness Matters More Than Homogeneity. *Political Psychology, 33*(5), 701-
    721. doi:10.1111/j.1467-9221.2012.00893.x

Zhang, Y., Cao, L., & Vaughn, M. S. (2009). Social Support and Corruption: Structural Determinants of
    Corruption in the World. *Australian & New Zealand Journal of Criminology, 42*(2), 204-217.
    doi:10.1375/acri.42.2.204