DYNAMICS OF ECONOMIC CORRUPTION

A DISSERTATION IN
Economics
and
Mathematics

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
the requirements for the degree

DOCTOR OF PHILOSOPHY

by
RAFED AMIN AL-HUQ

B.S., California Institute of Technology, 2000
M.A., University of Missouri-Kansas City, 2012

Kansas City, Missouri
2022

DYNAMICS OF ECONOMIC CORRUPTION

Rafed Amin Al-Huq, Candidate for the Doctor of Philosophy Degree

University of Missouri-Kansas City, 2022

ABSTRACT

Corruption, especially in developing countries, has been shown to have large persistent effects on inequality and economic growth. Efforts to eradicate or suppress corruption over the last two decades have floundered. This dissertation argues that urban population density is a major underlying factor of economic corruption that previous studies have never considered before. Increased population density communicate bribe success information rapidly and resistance to it becomes muted as more success is observed within a population. Urban density has not been considered before because while data for urban populations exist, data for urban extents did not. This dissertation uses satellite night-time lights data to get an estimate of urban extent and uses it to calculate urban population density. This research shows that an increase in urban density by 100 people per square kilometer has the same effect on corruption as a decrease in the per capita GDP by $172 (2005 dollars). Using spatial effect models we show why a small, smooth change in behavior parameters lead to large scale, sudden cascading effects in corruption across the population. Also using network theory models we show that most real worlds are small in the number of links required to traverse a population and it gets easier to find successful links between a briber and the bribed as urban density increases.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Graduate Studies have examined a dissertation titled "Dynamics of Economic Corruption," presented by Rafed Amin Al-Huq, candidate for the Doctor of Philosophy degree, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Mathew Forstater, Ph.D., Committee Chair
Department of Economics

Erik Olsen, Ph.D.
Department of Economics

Scott Fullwiler, Ph.D.
Department of Economics

Linwood Tauheed, Ph.D.
Social Science Consortium

David Spade, Ph.D.
Department of Mathematics and Statistics

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Overview

Corruption – often defined as the misuse of public office for private gain – is pervasive and even more so in developing countries. It includes acts of bribery, nepotism, fraud, and at the extreme, state capture. It reinforces inequality and stalls economic development. The World Bank (Kaufmann, 2005) estimates that corruption costs at least 5% of global GDP $(\sim 2.5$ trillion US dollars$)$ with over a trillion dollars paid in bribes. Political scandals regarding bribe taking have sparked enough public outrage that many governments have been forced out of office.

Against this backdrop, economists and political scientists have analyzed corruption by trying to explain what causes it and hoping that if they can explain it they may be able to uproot it. Yet the last two decades of anti-corruption efforts have rarely succeeded. We have a surfeit of data of measured corruption perceptions and new papers attempting to explain corruption and/or attempting to calculate its cost appear nearly every year. Yet corruption keeps on persisting if not growing in certain parts of the world.

This dissertation argues that urban population density is an important factor in affecting corruption that has been ignored by the literature so far. One of the major reasons for that is while most nations have an accurate measure of their urban population they rarely have an accurate measure of their urban extent. This is because most economic and sociological studies have rarely been interested in incorporating population density (at the national level) in their models. The few times they do include it, they take the overall population density of the nation as the appropriate measure. This is a mistake because in such scenarios the real effective density is highly underestimated for countries whose total populations agglomerate in a fraction of the countries' area (e.g. Canada, Australia, Egypt, Brazil, etc.). Therefore, one has to incorporate in ones models the appropriate urban population density of a nation.

In this dissertation, we take the urban populations numbers from the various nations statistical data (as collated and standardized by the World Bank) and the urban extent data from calculations derived from satellite imagery of night-time lights data. Satellite imagery data has in the last decade taken off as a good measure of land-use and population estimates and we consider it as a critical tool in attempting to answer the questions in this paper.

We demonstrate the effect of urban population density on corruption by regressing corruption data from Transparency International's Corruption Perception Index (TI-CPI) against urban population density measures along with per capita GDP and other control variables. We show that an increase in urban population density leads to increased corruption. Whereas an increase in per capita GDP decreases corruption. Specifically, an increase in urban density by 100 people per square kilometer has the same effect on increasing corruption as decreasing the per capita GDP by $172 (2005 dollars, $243 in 2021 dollars).

The dissertation will then continue by showing how urban density can have a spatial effect in propagating corruption across the nation. One aspect that surprises nations (and governments) in how suddenly and quickly corruption propagates in the country. If urban density increases, people are more likely to see what others are doing (and successfully doing with corruption). So as density increases, peoples propensity (or preference or desire or threshold - we use all these terminology cause that is what we will explore) will increase. In our models we show, that if such parameters change, that after crossing a certain parameter value corruption easily propagates throughout the system. We show with various spatial effect models how this can possibly come about and how sensitive it can be to various parameters

Now as urban density increases, the number of connections to other individuals increases. This dissertation argues that this implies that the set of all possible paths between a briber and a potential bribed bureaucrat increases. But it does not necessarily mean that an individual can actually discover such a path, or at least find someone who can help them find the connection. We show using network theory that as the number of connections increases it actually gets easier to find the mutual connection.

Structure of the Dissertation

Chapter 2 will discuss the theoretical reasons as to why urban population density should have an effect on corruption and we will then present the data that we will work with. The chapter will then continue to run regressions on the data and show that there is a substantial effect of urban density on corruption.

Chapter 3 will then explore the effect of corruption on the variables that have a long term effect on economic growth. Given the fact that we now additionally need to consider urban density in our models, we run regression to see what the effect will be along with the effect of corruption.

In chapter 4 we will explore three different spatial effect models. The Granovetter threshold model, the Schelling dynamic segregation model, and the Renormalization group percolation model. We show how these models can reflect corruption propagation in a population and how their parameters can drive the dynamics.

In chapter 5 we will explore three different network theory models of connection. The Erdős-Renýi random network model, the Watts-Strogatz small world model, and the Barbasi-Albert network model. We show how various assumptions of structure between links in a network has different implications about the ability to find a link between a briber and the bribed.

Chapter 6 concludes the dissertation with a big picture analysis of what was done and ideas for possible further work.

CHAPTER 2

THE EFFECT OF URBAN POPULATION DENSITY ON CORRUPTION

Theoretical Justification for the Effect of Population Density

Is there a reason why population density should matter in affecting corruption? Well one can appeal to the model of the waiting line. In third world countries, one can observe in almost every bureaucratic office a long line to get things done. The bureaucrat can request a bribe and the person he requests it from can refuse it. But if the line is long enough and more importantly visible enough – someone in the line will pay the bribe and more importantly his job will be visibly solved (either he jumps the line or his papers/job gets delivered). Population density is a proxy for such kind of line behavior. Just having a large population is not enough – there has to be visible proximity where one can see and know about the actions of others. This is why if you want anything to be done you have to move to the capital city where everyone lives in close proximity.

Another aspect is kin networks. A different way to describe corruption is that it favors one individual over another (for bureaucratic services) not because that individual paid a bribe but that the individual was part of the same kin network as the bureaucrat. The more population dense a region is, the more likely one shares a kin network with someone in the bureaucracy and can thus be tapped upon for help. In fact – more importantly – out of kin people seek out in-kin people and bribe them instead so that they will put in good word (and usually share the bribe through informal channels of family) with the government official. The more dense the population the more likely one can seek and find such a go-between.

The kin network is really an example of the power of weak ties. The core argument in this chapter is that corruption propagates itself through weak ties and that the number of weak ties an individual possesses scales with urban population density.

What is a weak tie? Granovetter (1973) argued in his classic paper “The Strength of Weak Ties” that when it came to finding a job, the weak social ties (acquaintances, friends of friends) were more useful in getting them a job then the strong social ties (family, close friends, roommates). The reasoning is that strong ties move in the same social circle as us and thus have mostly the same information as us. Weak ties on the other hand are our bridge to the outside world and communicate what is available and thus enable us to tap a much larger world of opportunities.

That paper opened up a floodgate of research in exploring weak ties in innovation, operations research, criminology, community development, etc. to the point that it is one of the top 10 most cited papers in sociology.

So how do weak ties work in corruption? For an individual the problem is that he seeks to obtain influence with a bureaucrat. He cannot directly ask for a favor. Therefore, he has to ask around his social network to see if anyone can offer an introduction for him. The more weak ties he has the more likely he will find someone in his social network to arrange for a meeting between him and the bureaucrat. Strong ties can work too – but the odds are better with the weaker bridging connections. This approach to corruption has been quite well researched in the literature (Kranton, 1996; Ledeneva, 1998; Oldenburg, 1987; Shleifer, 1997).

Thus we can now argue why kin networks matter. A bureaucrat might not be able to take a direct bribe but his kin can (and then share with the bureaucrat through the informal means of family). The total number of the bureaucrats' kin does *not* increase with increasing population density. However, the requirement of finding a weak network connection to one person (the bureaucrat) becomes the easier requirement of finding a weak network connection to any one of the kin-network. And as population density increases the probability of finding a successful network connection increases.

The argument now is to show that total number of weak (or strong for that matter) ties increases with population density. Here we refer to Pan et al. (2013) paper "Urban characteristics attributable to density-driven tie formation". Their research develops a mathematical model that demonstrates that social tie density (or total number of social ties in a unit area) scales proportional to (urban density * ln (urban density)). Or:

$$T(\phi) \ \sim \ \rho \, ln\rho$$

where ρ is the population density of the city and $T(\phi)$ is the social-tie density.

They then show that this super-linear scaling matches the empirical evidence of simple diffusion models (here they use AIDS/HIV infection prevalence data in US metropolitan areas) and also information diffusion models (here they use Euro Total GDP per square kilometer data of a set of European cities).

In both cases they find a good fit with the empirical data. In fact, one of the major reasons for their research is to argue that their functional form is a better match that the power law (with parameter tuning) that is being done currently as they can dispense with the parameter tuning.

They also show that their model is a good match for cell phone call volume data versus population (Calabrese et al., 2011). Their super-linear scaling is also confirmed in total number of contacts and total communication activity in cell phone data from UK and Portugal (Schläpfer et al.,2014).

Literature on Corruption that however does not study the effect of Population Density

In almost all the research in the literature the strongest correlation of corruption is poverty. The lower the per capita GDP the higher the corruption. Everybody agrees on this significant result – very few agree on why or even on the direction of causality.

Mauro (1995) was one of the earliest econometric based studies on the impact of corruption. Using *Business International* (BI) data from 1980-1983 Mauro investigated the link between corruption and bureaucratic efficiency and political stability and argued that the more politically unstable the nation is, the less bureaucratic efficient it is and less financial investment and educational investment are made leading to more corruption. Mauro researched the question of the direction of causality (whether corruption affected growth or the other way round) by using ethnolinguistic fractionalization as an instrument for perceived corruption and showed corruption affected growth. Controversy rages however as to whether the instrument is valid (Easterly & Levine, 1997) since ethnic divisions can effect growth in many ways like political instability and ethnically redistributive politics without affecting corruption directly.

Triesman (2000) instrumented for the countries' economic development using the countries distance from the equator to show causality of higher development causes lower corruption. However Glaser et al. (2004) critiques that malaria and other tropical diseases are more a reflection in such instruments and that these causes of bad health will affect the pace of development.

Bardhan (1997) looking at long term historical data argued that corruption had a negative effect on growth. However, the correlation was on past $(\sim 30$ year$)$ growth data on current (then 1996) corruption data. Wei et al. (1999) however using contemporary data argued that corruption had an adverse effect on economic growth by reducing domestic investment, decreasing foreign direct investment, increasing government spending, and especially distorting the composition of government spending.

There are also people who argue the flip side that corruption is a good idea. The classic being Samuel Huntington's infamous quote (Huntington,1968) "In terms of economic growth, the only thing worse than a society with a rigid, over-centralized, dishonest bureaucracy is one with a rigid, over-centralized and honesty bureaucracy."

Leff (1964) had argued that corruption will actually speed through bureaucratic delays and increase investment and also increases certainty of investment opportunities paying off – a 'hedge' against the full losses of bad economic policy as he preferred to frame it. He also argued that this would enable a subset of bureaucrats to work harder where bribes would act as a piece rate.

Channeling Huntington; Meon and Weill (2010) found that corruption could provide a 'greasing of the wheels' instead of 'putting sand in them'. Their work 'repeatedly found that corruption is less detrimental in countries where the institutional framework is weaker'. Aidt (2003) finds similar weakly positive findings for corruption. However most scholars disregard such findings and usually accept that corruption in a negative phenomenon

Maddison (2003) went to data going back to 1500s to show that current levels of perceived corruption correlate highly with the estimated per capita income of the last few hundred years. He demonstrated that the development level of countries (around 40 countries or so in his data set) nearly 200 years ago correlate very well with current corruption – however it just might be that it reflects the fact that rich countries then are rich countries now and that the industrial revolution took off in such countries earlier.

Montinola and Jackman (2002) have demonstrated that democracy has a quadratic relationship on perceived corruption. Their argument is that in the short run democracy will engage in graft to maintain power, but the longer democracy runs corruption decreases since institutions like free press and accountability develops and takes hold.

This however contrast with Ugo (2001) who shows that directly elected presidential systems are significantly perceived as more corrupt. Kunicova and Rose-Ackerman (2005) argues and shows that increase in state federalism leads to a significant increase in perceived corruption.

La Porta et al. (1999) trying to explain factors of good government hit upon the aspect that countries with Protestant background, and civil law systems (Britain and its colonies) had higher quality governments (which correlates with lower perceived corruption). Treisman (2000) confirmed this conclusion and extended it to countries having democracy for decades (more than 30 years with the implicit argument that it takes time to develop institutions sufficiently strong enough to influence people). This aspect of having democracy for decades also led Treisman to argue that colonies that retained common law for long periods (and did not revert to pre-colonial laws) were more likely to be less corrupt than colonies that did otherwise.

This analysis reflects the debate in the early 2000s between those who wished to argue that countries with good institutions developed faster (Acemoglu et al., 2001) against the flip argument that good economic development lead to better institutions (Boix & Stokes, 2003).

Ades and Di Tella (1999) show that openness to foreign trade slightly reduces corruption (although the direction of causality in their research is ambiguous). They also showed (this one was more clearer) that countries that have large natural resources rents are more likely to be corrupt (with the argument that in such countries officials who allocate rights to such rents are more likely exploit their position).

Braun and Di Tella (2004) show that corruption increases as inflation become high and variable. They argue that the reason is that unpredictable inflation makes it harder to monitor public spending and check on public contracts allowing corruption to flourish.

All these papers focus on different aspects in trying to explain corruption. However, none of them has considered urban population density as a potential factor which this paper attempts to incorporate.

Data on Corruption and Urban Extent

There have been two main sources of data the last 20 years for perceived corruption. The Corruption Perceptions Index (CPI) issued by the Berlin based NGO Transparency International (TI) (Transparency International, 2012) and the Control of Corruption Index (CCI) issued by the World Bank (WB) (Kaufmann et al., 2009) as part of their Governance indicators. The CCI takes values from -2.5 to 2.5 and the CPI from 0 to 10 (both implying a higher index indicates lower corruption).

Both of these indices are 'polls of polls' in that they come from aggregating information from a variety of sources like domestic and international business people surveys, country risk ratings by consultancies, and polls of country inhabitants. TI includes in its calculations at least a minimum of three surveys (back when it started) but now includes almost more than seven to up to a dozen for most countries. TI ratings have been available since 1995 (starting with 54 countries) and by their 2016 release (CPI 2016) included 168 countries. The World Banks CCI used to be biannual (1996 to 2002) but has gone annual since 2003 starting with 54 countries to 215 in the last survey in 2016. Even though the methodologies and sources are different, the two ratings are highly correlated. Treisman (2007) has calculated correlation of $r = 0.96$ in 2002 data and $r = 0.98$ in 2004 data.

Both of these are 'perceived corruption' data in that they use opinions and not conviction rates or total amount of bribes collected. There are many reasons for this. Amounts given as bribes data is impossible to collect, usually because admitting to giving a bribe can implicate one for criminal prosecution. Much safer in surveys to claim that if one does not give a bribe one cannot get things done instead of stating an amount. Secondly, criminal prosecution is more of a function of government funding which is extremely variable depending on the political climate of the country. In addition, criminal prosecutions in many third world countries are used to prosecute supporters of political rivals. Even if these supporters are bribe givers (or takers) the systematic bias of the data makes it unusable.

Because these two data sources are so similar, in this paper we will use the data from Corruption Perceptions Index (CPI) by Transparency International (TI).

For the data in our time frame, the Corruption Perception Index (CPI) has values of 0 to 10, 0 indicating most corrupt and 10 indicating least corrupt. We will work with inverted TI-CPI scores that makes 10 most corrupt and 0 least corrupt. We invert it so the higher number indicates more corruption; it makes our analysis and especially our graphical analysis more understandable.

The urban extent data come from night-time lights that is measured by satellites of the United States Air Force Defense Meteorological Satellite Program (DMSP). These satellites circle the earth 14 times a day and record the intensity of Earth-based lights with their Operational Linescan System since 1970. The digital archive of their data is from 1992 onwards. Scientists at the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center (NGDC) process the raw data to remove moonlight, aurora activity, and forest fires; leaving mostly man-made lights. This data is finally averaged over all valid nights to produce a satellite-year dataset. The night lights data has a range (called digital number (DN)) from 0 to 63. We code the areal cells that have a DN value greater than 12 as urban and consider them as part of the urban areal extent (Small et al., 2011; Zhou et al., 2015).

Henderson et al. (2012) took the data set from 1992 to 2008 and calibrated it to get total urban areal extent for those years. They included not only extent data but also total intensity with which they attempted to develop a proxy measure for GDP growth rate. We ignore the intensity data but take the extent data and using the urban population data from the World Bank World Development Indicators (WDI) (World Bank, 2015) calculate the urban population density. The final data set runs from 1995 (with 54 countries, when TI began calculating corruption data) to 2008 (when the light areal extent data ends). By 2008, 148 countries are in the data set. Appendix A has the list of all countries in the data set sorted by continent.

Other Data and their Sources

Almost all of the data that will be used as independent variables (GDP, GDP per capita, total population, total urban population, natural resource rents, remittances, inflation, and government debt) come from the World Development Indicators (WDI) issued by the World Bank every year. Preliminary data analysis was done on the International Political Economy Data Resource (Graham et al., 2016). The final analysis was done on collected WDI data. This data comes from the statistical divisions of the various countries and collated, corrected, and standardized by the World Bank with supplements of data coming from the International Monetary Fund (IMF). In all the corruption literature WDI data is the usual source for analysis and there is little controversy regarding it. We will be using collated data issued in 2015 (World Bank, 2015).

The government debt data is complemented with data from the IMF's Global Debt

Database (Mbaye et al., 2018). The Gini data comes from the World Bank's WDI and complemented with data from the World Inequality Database (Alvaredo et al., 2020)

The democracy duration data comes from Boix et al. (2018) ”Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2015”. It is a dataset that provides very clear data on countries histories of democratic government. It is very popular in political science literature. Since the data set starts from 1800, it codes UK and USA democracy at a maximum of 209 years for the year 2008.

The religion data of the percentage of Christian, Muslim, and Buddhist populations comes from Brown and James (2015) ”Religious Characteristics of States Dataset Project”. It is a very detailed dataset and commonly used in political science literature. We do not use data for the percentage Hindu, Jewish, or Shinto populations since it ends up being a dummy variable for India, Israel, and Japan respectively.

Table 2.1 shows the variables used in this chapter and their units. Table 2.2 has the summary statistics for every variable.

Colonial Data

One of the things we wish to control for in our analysis is the colonial past of the countries in our dataset. Treisman (2000) had shown that former British colonies had significantly lower levels of corruption. Therefore we need to incorporate the colonial status of every country in our data set. We use the very important paper by Lange and Dawson (2009) ‘Dividing and Ruling the World? A Statistical Test of the Effects of Colonialism on Postcolonial Civil Violence’ to catalog our countries. They allocate a country a colonial status if it was a colony for at least 10 years.

Another important feature of their analysis is they consider important the idea of an internal colony. That is former empires or unions that broke apart into independent

Table 2.1: Units of the Data Variables

	(1) Units
Corruption	unit less number between 0 and 10
GDP per capita	constant 2005 US $
Urban Density	people per square kilometer
Natural Resource Rents	% of GDP
Remittances	% of GDP
Government Debt	% of GDP
Inflation	percentage
Gini	unit less number between 0 and 100
Democracy Duration	years
% Christian	% of total population
% Muslim	% of total population
% Buddhist	% of total population

Table 2.2: Summary Statistics of the Variables Data ($n = 1340$)

	(1) Mean	(2) Std. Dev.	(3) Min	(4) Max
Corruption	5.60	2.25	0	9.6
GDP per capita	11447	15127	144	67805
Urban Density	2120	2685	68	14892
Natural Resource Rents	9.17	14.97	0	85.26
Remittances	3.18	5.69	0	49.29
Government Debt	50.17	34.37	0.47	260.9
Inflation	8.17	16.40	-8.53	325
Gini	39.07	9.08	21.8	66.3
Democracy Duration	38.12	37.28	1	209
% Christian	54.48	35.17	0.09	97.77
% Muslim	22.09	34.13	0.01	99.65
% Buddhist	3.57	14.10	0	87.84

Table 2.3: Percentage of the data of the 1340 observations that belong to each colonized subset and the number of countries in each colonized subset.

	(1) Percentage of the Observations	(2) Number of Countries
Never Colonized	18.36	22
Colony of Britain	29.25	44
Colony of France	10.00	24
Colony of Spain	15.89	18
Colony of Portugal	2.39	4
Colony of Netherlands	2.09	3
Colony of Others	4.40	9
Internal Colony of Austro-Hungarian Empire	5.75	6
Internal Colony of Ottoman Empire	2.99	4
Internal Colony of Soviet Union	8.88	14
Sum	100	148

nations; like the Soviet Union, the Austro-Hungarian Empire, and the Ottoman Empire. Appendix B has a list of the countries according to their colonial status. Table 2.3 shows the colonial powers that existed, the number of countries within each colonial power, and the percentage of the data observations that fall within each colonial status.

Graphical Analysis

Let us take a look at a small sample of the data to get a feel for the trends that could exist. We will take the corruption data from 2008 and plot it against the various independent variables. All of these graphs are in Appendix C.

In Figure C.1 we plot corruption against urban population density. As expected, we see as urban population density (in persons per square kilometer) increases corruption increases (the higher the number the more corrupt the country (here we just invert TI-CPI numbers to make it appear more understandable)). Most of the developed countries have a score of six or below.

To see the effect more clearly, in Figure C.2 we plot the same data with the urban population density plotted on a logarithmic scale. We use a logarithm of base 2 (and not base 10) to show how for every doubling of population density corruption increases. It gives us a more intuitive and wider dynamic range that base 10 would have provided. The lower bound of urban population density is $2^6 = 64$ and the upper bound is $2^{14} = 16384$. The whole dynamic range is about 8 doublings.

Similarly, in Figure C.3 we plot corruption against GDP per capita (in constant 2005 US dollars). We can see very clearly as income increases corruption decreases. To see the effect more clearly at lower incomes, in Figure C.4 we plot the same data with the GDP per capita plotted on a logarithmic (base 2) scale. The lower bound of GDP per capita is $2^7 = \$128$ and the upper bound is a bit over $2^{16} = \$65536$. Corruption really does not appear to decrease until we cross $2^{12} = \$4096$.

In Figure C.5 we plot corruption against the percentage of GDP that comes from natural resource rents (like crude oil, natural gas, coal). We observe that only a few countries, as their natural resource rents go above 20% of GDP, can avoid the curse of corruption.

In Figure C.6 we plot corruption against inflation. In the low inflation scenarios there is not much pattern to be discerned. But when the country has inflation more than 12% more often than not the country is considered among the corrupt.

In Figure C.6 we plot corruption against inflation. In the low inflation scenarios there is not much pattern to be discerned. But when the country has inflation more than 12% more often than not the country is considered among the corrupt.

In Figure C.7 we plot corruption against the percentage of GDP that come from personal remittances. The graph clearly shows that for countries for which remittances make more than 5% of GDP it is usually considered to be corrupt.

In Figure C.8 we plot corruption against the Gini index. Here the implications are not very clear. For low inequality countries (Gini index $<$ 35) the countries are considered not as corrupt (CPI $<$ 6). But the same is true for very unequal countries (Gini index $>$ 55). It is the countries that are in between (35 $<$ Gini index $<$ 55) that most of them are considered corrupt.

Finally in Figure C.9 we plot corruption against the government debt to GDP ratio. Here again it is hard to discern a pattern. For countries with a debt to GDP ratio of less than 50% (a majority of them), its seems evenly split between corrupt and non corrupt. But that maybe more of a consequence of the fact that most nations it is really difficult to run up a high amount of government debt in the first place. Again we need to remind the reader that all these graphs are for the year 2008. Patterns might be different for other cross-sectional years.

Regression Analysis

We can now run a regression on our data. One issue we need to be careful about when analyzing regression models is that the coefficients of the independent variables might be statistically significant but not necessarily significant in magnitude (Ziliak et al., 2008). Therefore in the following analysis we make sure to discuss the size of the magnitude of the coefficients and their appropriate real world effects' significance.

Our dependent variable, as previously mentioned, will be corruption as measured by the inverted TI-CPI scores. Inverted because TI-CPI makes 0 most corrupt and 10 least corrupt - we invert it so 10 is most corrupt so that it appears more understandable. The independent variables will be all the other variables in Table 2.1.

The data consists of 148 countries from 1995 to 2008 and is unbalanced (that is some countries have more data years than others). There are a total of 1340 data points. We are using a pooled OLS (Ordinary Least Squares) model and the regression is corrected for heteroscedasticity by using robust standard errors. All regressions are calculated using the software STATA. Table 2.4 and Table 2.5 have the regression results.

Model 1 in the table does not incorporate the colonial dummy variables, but instead just calculates a constant (intercept term). Model 2 does incorporate the colonial dummy variables and removes the constant (intercept term) so as to avoid perfect multicollinearity. Incorporating the colonial dummy variables does not change model 2 much from model 1. The coefficient signs remain the same and almost all of them are still significant. There are just two major changes. The coefficient for the Gini variable increases by an order of magnitude and goes from not significant to significant at the 1% level. And the coefficient of the % Christian population variable goes from significant at the

Table 2.4: Regression Output of the Corruption Models

	(Model 1) Corruption	(Model 2) Corruption
Urban Density	0.000105***	0.000138***
	(0.000008)	(0.000009)
GDP per capita	−0.0000895***	−0.0000804***
	(0.0000036)	(0.0000038)
Natural resource rents	0.0298***	0.0317***
	(0.0018)	(0.0018)
Remittances	0.0258***	0.0210***
	(0.0044)	(0.0039)
Inflation	0.0047**	0.0045**
	(0.0021)	(0.0021)
Government Debt	0.0022***	0.0047***
	(0.0008)	(0.0008)
Gini	0.0029	0.0224***
	(0.0039)	(0.0049)
Democracy Duration	−0.0107***	−0.0082***
	(0.0015)	(0.0015)
% Christian	0.0059***	0.0012
	(0.0015)	(0.0017)
% Muslim	0.0075***	0.0059***
	(0.0015)	(0.0016)
% Buddhist	0.0105***	0.0095***
	(0.0022)	(0.0024)
Constant	5.669***	
	(0.189)	
Observations	1340	1340
Adjusted R^2	0.7958	0.8089

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: Colony Dummy Variables Coefficients - Regression Output of the Corruption Models

	(Model 1) Corruption	(Model 2) Corruption
Never Colonized		4.669***
		(0.244)
Colony of Britain		4.453***
		(0.275)
Colony of France		4.578***
		(0.274)
Colony of Spain		5.061***
		(0.298)
Colony of Portugal		4.396***
		(0.308)
Colony of Netherlands		5.396***
		(0.265)
Colony of Others		4.749***
		(0.292)
Internal Colony of Austro-Hungarian Empire		5.370***
		(0.205)
Internal Colony of Ottoman Empire		5.579***
		(0.242)
Internal Colony of Soviet Union		5.435***
		(0.236)
Observations	1340	1340
Adjusted R^2	0.7958	0.8089

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1% level to not significant at all. It also goes smaller by a factor of 5. We will only work and interpret with the results from model 2.

One crucial thing to be aware of when we compare the two models is the change in the p-values of the coefficients. Appendix D shows the two regression models with the exact p-values. Looking at the specific change of the p-value of a coefficient can be useful in understanding the two differently specified models.

Looking at the results from model 2, we see it does a good job of explaining the variation (adjusted R-square of 80.89%) and more importantly for our purposes urban population density is statistically significant at the 1% level (p value < 0.01). The coefficient is 0.000138 and it says that for every thousand increase in urban population density the Corruption Perception Index (CPI) goes up by 0.138. Given that many of the most corrupt countries (like Bangladesh, Ethiopia, etc.) have a urban population density of $\sim$ 10,000 persons or more per square kilometer, we are looking at easily a 1.4 or more change in the CPI just due to population density. For the corrupt countries, the perception range usually ranges from 6 to 9, so nearly 15% to 23% of the perceived corruption just comes as an effect of the urban population density. This is quite a large effect.

As a comparison to get the same effect of approximately 1.4 change in the CPI the per capita GDP has to increase by approximately $17,400 to counter the effect (using the coefficient of -0.0000804 for the per capita GDP).

This negative coefficient for the per capita GDP is telling us that as the country get wealthier corruption decreases. And more strongly for every $1000 increase in per capita GDP, corruption goes down by 0.08 in CPI.

Therefore, we can argue that for every 100 increase in urban population density the GDP per capita has to increase by approximately $172 (in 2005 dollars, $243 in 2021 dollars) to counter it. For such poor countries (Bangladesh 2008 per capita GDP was \$493), it is an almost impossible amount and we can see the why most attempts to eradicate corruption have failed. They are fighting an uphill battle against a structure that prevents their success and they are not even aware of it. For many countries their population is increasing faster than their GDP and corruption in such cases will be almost impossible to counter.

The other variables are also interesting. The coefficient of Natural Resource rents (also significant at the 1% level) says that for every 10% of its contribution to total GDP, corruption increases by 0.317. So for a country like Saudi Arabia (CPI of 6.5) for whom 65% of GDP (in 2008) come from crude oil rents; approximately 32% of its corruption perception comes from this.

Remittances are also important. The coefficient of Remittances (significant at the 1% level) says that for every 10% contribution in remittances as a percentage of GDP brings about a 0.21 increase in CPI. So in 2008, for Nepal (CPI of 7.3) or Honduras (CPI of 7.5) for whom 22% of their GDP comes from remittances from abroad, approximately 6% of its corruption perception comes from this.

These four variables are the most important variables in affecting corruption. Our next variable Inflation has a coefficient of 0.0045 or for every 10% of inflation brings about an increase of 0.045 in CPI which is very negligible. It is also only statistically significant at the 5% level. In the 2008 dataset, Venezuela (CPI of 8.1) has the highest inflation at 31.5% but that only contributes 1.75% to its corruption perception. So inflation does not have a large effect on corruption.

For the government debt to GDP ratio variable, its coefficient is 0.0047 or a 10 percentage point increase in debt increases the CPI by 0.047. In the 2008 dataset, Lebanon (CPI of 7) has a debt to GDP ratio of 160% which contributes 10.7% to its corruption perception.

The Gini variable is much more interesting. In model 1 it was not statistically significant at all. In model 2 with the colonial dummy variables it is statistically significant at the 1% level with a coefficient of 0.0224 or a 10 point Gini will contribute 0.224 to the CPI. In the 2008 data set, South Africa (CPI of 5.1) has one of the highest Gini coefficient of 59.4 which contributes 26% to its corruption perception.

The Democracy Duration variable's coefficient is negative which says the longer a country is a democracy the less likely the society will be corrupt. It is considered to be an important control variable in the literature but it has a small coefficient of -0.0082. In 2008 India (CPI of 6.6) has a democratic duration of 59 years which decreases its CPI by 0.48 or by approximately 7%.

As control variables both models incorporate the percentage of the population that is Christian, Muslim, and Buddhist respectively (Hinduism and Judaism are not added cause they effectively end up being a dummy variable for India and Israel). They are common in the literature as control variables for ethnic harmonization. The more the country is religiously homogeneous, the more likely an individual would be willing to help a fellow co-religionist in effecting an bribe. Or a better way to restate it - if a population is divided over religious lines, it is unlikely an individual will help someone of another religion on actions of dubious legality.

The variable for the percentage Christian population is significant in model 1 but not model 2 when we include the colonial dummies. This is most likely the colony dummies capture the effect of Christian religious dominance in certain colonial empires (e.g. Spanish and Portuguese empires). Both the percentages of Muslim and Buddhist populations are significant at the 1% level and they have small but positive coefficients. If a nation was 100% Muslim it would have its CPI increase by 0.59 and if a nation was 100% Buddhist its CPI would increase by 0.95.

We next look at regression coefficients of the colonial dummy variables in Table 2.5. These coefficients tell us the baseline corruption of the countries based on their colonial status. As we can see the British colonies have a very low baseline corruption, but so do the colonies of France, Portugal, and those never colonized. Spanish colonies are the first major subset of countries whose baseline corruption is statistically higher than that of the British colonies.

The difference is the baseline between Spanish and British colonies is 0.608 in CPI. But we have to be careful since these are random variables and we need to account for their variances and covariances. Using data from the variance-covariance matrix (in Appendix E) the standard error of the difference is 0.117 (using the formula of the sum of the two variances minus twice the covariance of these two variables, followed by the square root). Or the 95% confidence interval of the difference is (0.379, 0.837) which is different from zero. Therefore we can say statistically (at the 5% level) that Spanish colonies have a baseline corruption higher than British colonies.

This calculation can be done for any pair of colonies and it can be clearly seen for colonies whose baseline CPI is greater than 5. An interesting observation is that all internal colonies (Austro-Hungarian, Ottoman, and former Soviet Union) have a baseline corruption level much higher (> 0.9) than non-internal colonies. This angle is worth exploring as to why this is true and what institutional structures lead to such a scenario happening.

Countries that are Different than the Model Predicts

One of the consequences of developing a model like this is that it enables us to find which countries are more (or less) corrupt than usual. Figure 2.1 has a plot of the true Corruption (the measured CPI values in 2008) against the fitted values of the corruption that the regression model would predict (this is model 2). Countries that fall on the straight line (45 degree line) the measured data matches the model predictions. But the more interesting countries are the ones off the 45 degree line.

To see it more clearly, in figure 2.2 we plot the measured Corruption CPI against the standardized residuals. Thus we can clearly see which countries are 2 or more standard deviations less corrupt than the model predicts. Two countries stand out; Chile and Uruguay are more than 2 standard deviations less corrupts that the model suggests they should be. They are the major outliers. If we wish to study a few more less corrupt countries (more than 1.5 standard deviations less corrupt) we can also include New Zealand, Estonia, Barbados, Dominica, Oman, Bhutan, Ghana, and Mauritania in our list.

The benefit in having a list like this is that it suggest to us which countries to study further and explore their institutional structure to understand why they are less corrupt that usual. For example (this is a personal communication from a friend who is a resident of Jamaica) Barbados is less corrupt compared to Jamaica cause the colonial overlords moved to Barbados and thus had an incentive to develop the nation and not strip extract the value of the island which was what happened in Jamaica. It is question about institutional structures like these that are worth studying and this model tells us on which countries to concentrate our attention on.

Similarly, on the flip side, we can find which countries are more corrupt than expected. Five countries are two standard deviations more corrupt that the model predicts they should be – Equatorial Guinea, Belarus, Italy, United States, and Norway. Depends on your viewpoint it can be obvious or surprising. Italy has the problem of syndicated crime being in collaboration with the government. Norway is a bit more surprising, but it has been a common perception of Norwegians that the state is misusing their massive sovereign wealth fund and it has been a huge issue in that nation.

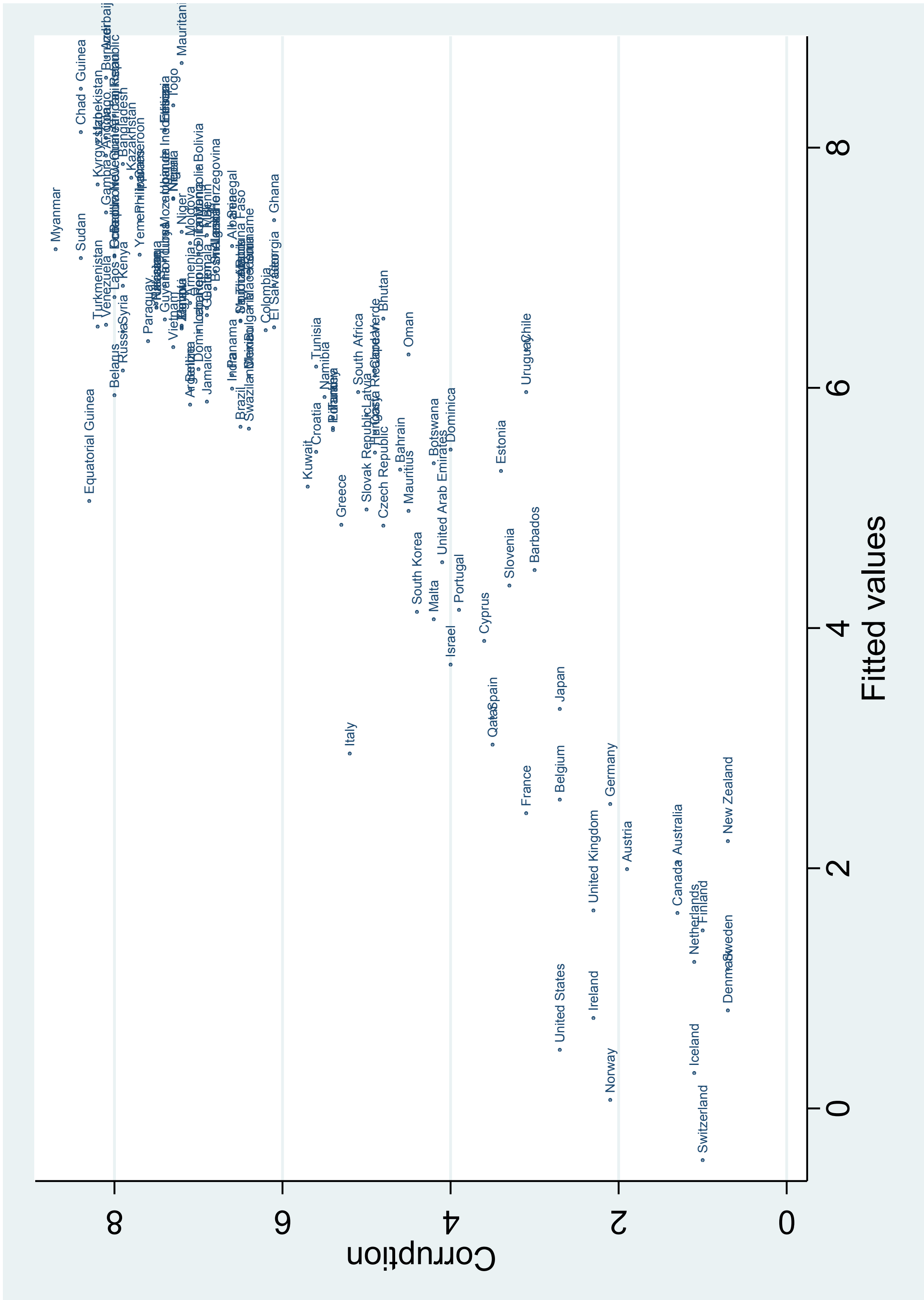

Figure 2.1: Plot of Corruption against Fitted Values of Corruption on 2008 data

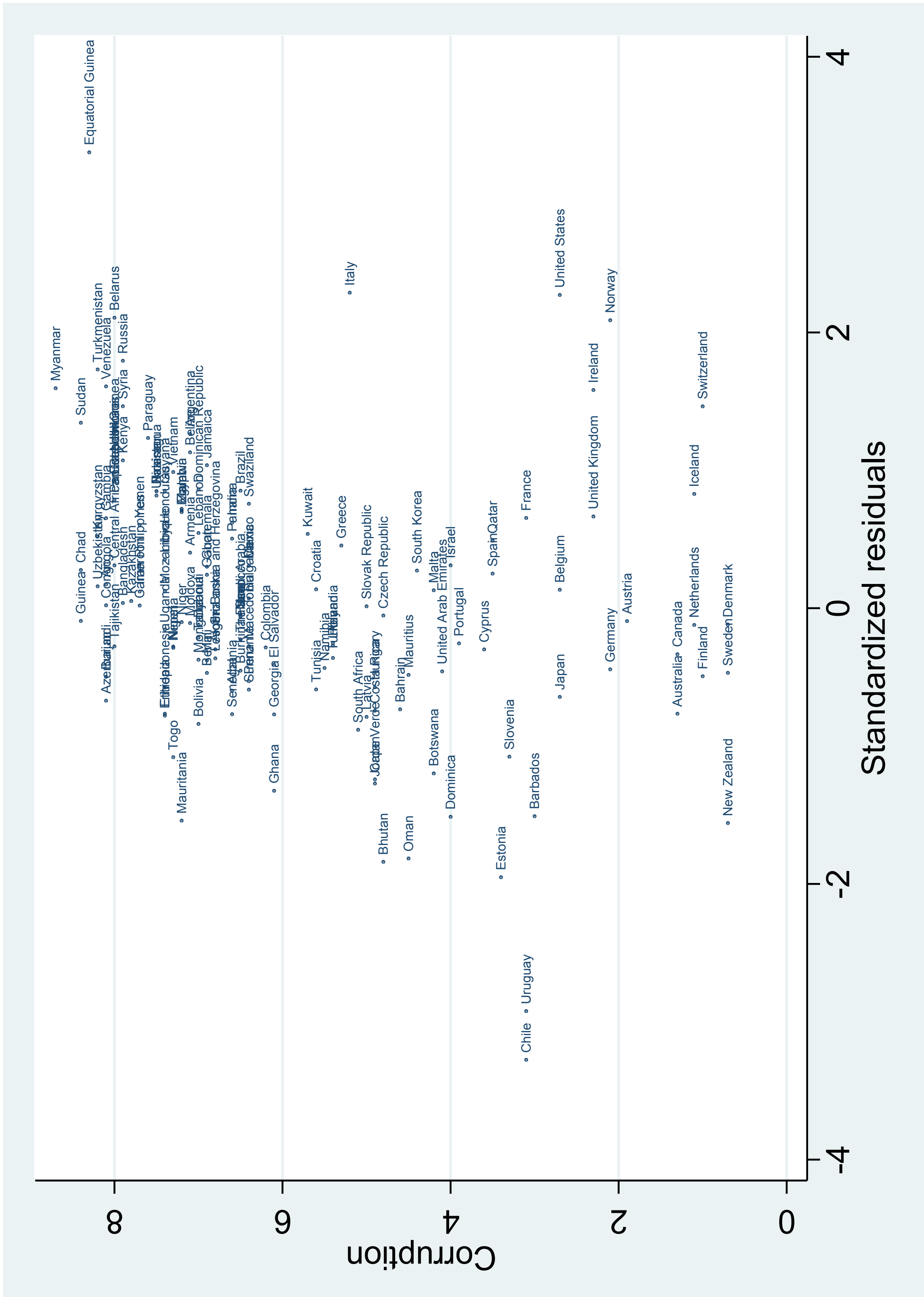

Figure 2.2: Plot of Corruption against Standardized Residuals on 2008 data

As for the United States – there could be many reasons – but for its given per capita GDP, urban density, and other variables the expectations is that its CPI should be closer to that of Canada's CPI of 1.3 (year 2008) rather than the 2.7 that it actually is – it is more than a doubling.

If we want to study a few additional more corrupt countries (more than 1.5 standard deviations more corrupt) we can also include Switzerland, Ireland, Argentina, Belize, Paraguay, Kenya, Syria, Russia, Venezuela, Turkmenistan, Sudan, and Myanmar. Thus we definitely have a list of countries whose institutional structure is worth studying in why they are more corrupt than expected.

Conclusions

In this chapter, we have tried to argue that attempts to explain economic corruption need to take into account urban population density in their model specifications. It has quite a large effect in the poor and corrupt countries and ignoring it will be a major mistake. An increase in urban population density by 100 people per square kilometer has the same effects on corruption as decreasing the per capita GDP by $172 (in 2005 dollars, $243 in 2021 dollars). This is a substantial effect and has never been considered before when develop anti-corruption programs.

Other major variables of interest we show are Natural Resources rents (for every 10% of its contribution to total GDP, corruption increases by 0.317) and remittances (for every 10% of its contribution to total GDP, corruption increases by 0.21).

This chapter also explored the effect of corruption due to the colonial status of countries. We show that countries that were Spanish colonies have their corruption perception index higher by 0.608 compared to British colonies (95% confidence interval of the difference is (0.379, 0.837)). Also internal colonies (Austro-Hungarian, Ottoman and former Soviet Union states) have a baseline corruption level much higher (> 0.9) than non-internal colonies.

We also show, in this chapter, countries that are outliers in their corruption perceptions than what the model predicts. Chile and Uruguay are more than 2 standard deviation less corrupt than the model suggest they should be. New Zealand, Estonia, Barbados, Dominica, Oman, Bhutan, Ghana, and Mauritania are all more than 1.5 standard deviation less corrupt. This gives us a list of countries to study further and explore their institutional structure as to why they are less corrupt than usual.

On the flip side we also find countries that are more corrupt that the model predicts. Five countries are more than 2 standard deviations more corrupt that the model suggests they should be – Equatorial Guinea, Belarus, Italy, United States, and Norway. Thus we have a list of countries whose institutional structure and historical evolution is worth exploring.

CHAPTER 3

ECONOMIC EFFECT OF CORRUPTION

Introduction

In the previous chapter we have introduced urban density as an important variable in studying corruption. In this chapter we incorporate urban density along with corruption to see its effect on a few different but critical economic variables.

The four new variables we will explore are the tax revenue collected by a country, the net inflow of foreign direct investment (FDI) into a country, government spending, and government spending on education. These four variables were chosen cause they have a large effect on the economic growth of a country and corruption can have an effect on them.

Data

We will be using the same corruption data (from Transparency International) and the same independent variables that we used in chapter 2. The four new dependent variables that we will explore comes from the World Development Indicators (WDI) as issued by the World Bank. We will be using the collated data issued in 2015 (World Bank, 2015). Table 3.1 shows the summary statistics of these new dependent variables and the number of data points we have for them.

Table 3.1: Summary Statistics of the Dependent Variables

	(1) Mean	(2) Std. Dev.	(3) Min	(4) Max	(5) n
Tax Revenue (% of GDP)	17.57	7.85	0.88	65.90	915
Foreign Direct Investment (FDI) net inflow (% of GDP)	4.31	5.03	-8.40	50.97	1329
Government Spending (% of GDP)	15.50	5.19	2.74	42.51	1309
Government Spending on Education (% of GDP)	4.59	1.61	0	14.79	914

Graphical Analysis

Let us take a look at a small sample of the data to get a feel for the trends that could exist. We will take the corruption data from 2008 and plot the new dependent variables against it to see what kind of effect corruption has. All these graphs are in Appendix C.

In Figure C.10 we plot the tax revenue collected (as % of GDP) against corruption. We can observe a pattern of a small downward trend, as corruption increase less tax revenue is collected.

Similarly in Figure C.11 we plot Foreign Direct Investment (FDI) net inflow (as % of GDP) against corruption. Here the downward trend is even harder to detect, but what is much clearer is that there is a large variance in net FDI inflow as corruption increases.

In Figure C.12 we plot total government spending (as % of GDP) against corruption. Here the downward trend is much more clearer. We observe that as corruption increases, total government spending decreases.

Finally in Figure C.13 we plot total government spending on education (as % of GDP) against corruption. Here we can clearly see that as corruption increases, total education spending decreases. However there is a large variance in education spending among the highly corrupt countries.

Regression Analysis

We can now run regressions on our data. Corruption as measured by the inverted TI-CPI scores will be now considered as an independent variable in these models. Except for the four new dependent variables, the rest of the variables will be the same independent variables we used in Chapter 2.

The data consists of 148 countries from 1995 to 2008 and is unbalanced (some countries have more data years than others). We have four different regression and they all have different number of observations from a maximum of 1329 to a minimum of 914. We are using a pooled OLS (Ordinary Least Squares) model like we used in Chapter 2 and the regressions are corrected for heteroscedasticity by using robust standard errors. All the regression models incorporate the colonial dummy variables and thus the constant (intercept term) is removed to avoid perfect multicollinearity.

All regressions are calculated using STATA. Tables 3.2 to 3.5 have the regression results.

One possible issue with these regressions is the possibility of multicollinearity. We have previously showed that corruption is affected by a multitude of independent variables. Now in these models we are including corruption as an extra independent variable along with the other independent variables. To see if we might have any issues, we calculate the variance inflation factor (VIF) of these regression models. Table F.1 in

Appendix F has all the VIF value for these four regression models.

Wooldridge (2015) has argued that multicollinearity is not an issue for any VIF less than 10. As we can see in Table F.1 no VIF is near 10. GDP per capita is the variable that has the highest average VIF value of ~ 6 followed by the Corruption variable with an average VIF value of ~ 5.5. Thus multicollinearity is not an issue for our models.

Regression Results of Tax Revenue Collected

In Table 3.2 the first model regresses the Tax revenue collected (as % of GDP) against the other variables. The first coefficient is for corruption which states that for every 1 unit increases in CPI (Corruption Perception Index) the total tax revenue collected decreases by 1.595 percentage points. This is quite a large effect and is statistically significant at the 1% level. This shows that corruption has a major and persistent effect in the economic activity of a country and worth considering in ways to fix it.

The second coefficient is for the variable of Urban Density (also significant at the 1% level) and says that for every 1000 people per square kilometer increase in density, tax revenue collected will go up by 0.6% points. This is not that much of a huge effect. The GDP per capita coefficient is negative (and only statistically significant at the 10% level so a weak effect) and says that for every $10,000 increase in GDP per capita the tax revenue collected will go down by 0.6% points.

Natural resource rents have no effect on tax revenue collected. However, Remittances does have a large effect, for every 1% point increase in remittances a nation receives the tax revenue goes up by 0.27% points (both variables are measured in % of GDP) So one can argue that almost 25% of all remittances flow in a nation is collected as taxes, which could explain why nation states are very eager to control and encourage remittance flow.

Table 3.2: Regression Output of Tax Revenue Model and FDI Inflow Model

	(1) Tax Revenue	(2) FDI Inflow
Corruption	−1.595***	−0.549***
	(0.206)	(0.123)
Urban Density	0.000607***	−0.00006
	(0.000092)	(0.00005)
GDP per capita	−0.000062*	0.000003
	(0.000037)	(0.000022)
Natural resource rents	−0.00002	0.0613***
	(0.03252)	(0.0157)
Remittances	0.2707***	0.1296***
	(0.0843)	(0.0213)
Inflation	0.0415**	−0.003
	(0.0183)	(0.008)
Government Debt	0.0197*	0.0144***
	(0.0119)	(0.0050)
Gini	−0.159***	−0.022
	(0.057)	(0.021)
Democracy Duration	−0.039***	−0.005
	(0.012)	(0.004)
% Christian	0.071***	0.011
	(0.010)	(0.007)
% Muslim	−0.017	0.004
	(0.012)	(0.006)
% Buddhist	−0.002	0.013
	(0.017)	(0.009)
Observations	915	1329
Adjusted R^2	0.3654	0.4945

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Colony Dummy Variables Coefficients - Regression Output of Tax Revenue Model and FDI Inflow Model

	(1) Tax Revenue	(2) FDI Inflow
Never Colonized	27.46***	4.43***
	(2.52)	(1.17)
Colony of Britain	32.79***	5.82***
	(3.11)	(1.18)
Colony of France	35.00***	6.40***
	(3.26)	(1.13)
Colony of Spain	25.02***	5.91***
	(2.84)	(1.28)
Colony of Portugal	33.10***	6.83***
	(3.45)	(1.48)
Colony of Netherlands	34.19***	6.28***
	(2.89)	(2.03)
Colony of Others	28.17***	3.71***
	(2.47)	(1.17)
Internal Colony of	26.57***	8.34***
Austro-Hungarian Empire	(2.00)	(1.52)
Internal Colony of	28.12***	7.31***
Ottoman Empire	(2.35)	(1.69)
Internal Colony of	25.72***	9.50***
Soviet Union	(2.32)	(1.32)
Observations	915	1329
Adjusted R^2	0.3654	0.4945

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Inflation has a not so large effect (and also statistically significant only at the 5% level), a 10% increases in inflation just increases tax revenue by 0.415% points (ignoring of course the entire seigniorage effect). But it does suggest that inflation does not really hurt economies in at least taxes collected.

Similarly Government debt does not have that large an effect (statistically significant at the 10% level). A 10% point increase in debt to GDP ratio has only a 0.197% point increase in tax revenue. This of course ignores entirely the multiplier effect of increased debt and/or the fact if the debt has really been monetized by the nations' central bank.

Inequality does have a major effect (statistically significant at the 1% level). A 10 point change in the Gini coefficient decreases the tax revenue collected by 1.59% points. So a highly unequal nation does have a populace paying less tax than average.

The coefficient of democracy duration is negative (and significant at the 1% level) and says for every 10 years of democracy a state will collect 0.39% points less in taxes, but this seems more a reflection of the fact that long duration democratic countries are more richer than average anyway (due to economic stability). The religion variables are just the usual control variables.

The more interesting coefficients are the colonial dummies that we can see in Table 3.3 (model 1 for tax revenue). It shows a clear demarcation of two groups. One group of colonies (of Britain, France, Portugal, and Netherlands) collected between 32% to 35% of GDP as tax revenue. The remainder (especial the Spanish colonies and the internal colonies) collected between 25% to 28% of GDP as tax revenue.

These differences are strongly statistically significant, for example the difference between Britain and Spain is 7.76 and the standard error of the difference is 0.8. So it is worth exploring the institutional structure of the colonies (power dynamics, rule of law, etc.) to understand why such a difference occurs.

Regression Results of Foreign Direct Investment (FDI)

In Table 3.2, the second model regresses the Foreign Direct Investment (FDI) inflow (as % of GDP) against the other variables. . The first coefficient is for corruption which states that for every 1 unit increases in CPI (Corruption Perception Index) the FDI inflow decreases by 0.549 percentage points (this is statistically significant at the 1% level). Given that the average FDI inflow for all the countries in the dataset is 4.31% this is a very large effect. A one unit increase in CPI has a nearly 13% decrease in FDI inflow. This shows that corruption has a major and persistent effect in the net foreign investment of a country and can have a large braking effect on potential economic growth.

Urban Density and GDP per capita has no effect on FDI inflow. But Natural Resource Rents does, which makes sense since if the country is making a lot of money for exporting its raw materials. A lot of foreign investment money will flow in to make that extraction easier. The coefficient says for every 10% point increase in natural resource rent FDI inflow increases by 0.613% points. Again a very large effect.

Remittances also have a large effect. For every 10% point increase we observe a 1.296% point increase in FDI inflow. Among all the variables, remittances have the largest effect. An argument could be made that since a substantial fraction of the labor force is abroad they are able to convince foreigners to try to invest in the labors home country.

Government debt has a large statistical significance at the 1% level but it not of a large magnitude. For every 10% point increase in debt to GDP ratio, FDI inflow increases by 0.144% points. All the other remaining variables do not have much of an effect on FDI inflow.

We now look at the colonial dummies coefficient in Table 3.3 (model 2 for FDI inflow). One thing that pop outs very clearly is that internal colonies have a much large FDI inflow that external colonies Internal colonies on average have an 8% FDI inflow compared to an average 6% FDI inflow for external colonies. This makes sense if we view internal colonies still being economically strongly attached to their former colonial overlords.

This also makes sense when we see that among external colonies, French colonies have a higher FDI inflow. Since the French central bank backs the CFA Franc, it has a high level of economic influence in those countries (Benin, Burkina Faso, Côte d'Ivoire, Guinea-Bissau, Mali, Niger, Senegal, Togo, Cameroon, Central African Republic, Republic of the Congo, Gabon, Equatorial Guinea, and Chad) who together have approximately 20% of GDP in sub-Saharan Africa. Thus there is more FDI inflow in such countries, though not necessarily benefitting them.

Regression Results of Total Government Spending

In Table 3.4, the first model regresses Government Spending (as % of GDP) against the other variables. The first coefficient is for corruption which states that for every 1 unit increases in CPI total government spending decreases by 1.452% points, which is a large effect (and statistically significant at the 1% level). Since in this data set the mean government spending is 15.5% a unit change in CPI has a nearly 10% decrease in government spending. So corruption does have a substantial effect in the ability of the state to make a difference to the populace.

Urban Density (also significant at the 1% level) has an effect. For every increase of 1000 people per square kilometer, government spending goes down by 0.216 percentage points. One reason for the negative effect is that since people are so crowed together one does not have to spend too much to affect many people. An infrastructure project serves way more people than average in a densely populated location.

GDP per capita has a small but substantial effect (significant at the 1% level). For every $1000 increase, government spending decreases by 0.088 percentage points. One reason for the negative effect might be diminishing marginal utility. As a country gets richer (and collects more money as taxes in absolute terms) and spends more money on the populace it ends up having diminishing effect on the well-being of the population. So the state might end up not spending as much. Also a rich populace might not demand or need too many government services.

Inflation has a very small effect (and only statistically significant at the 10% level). For every 1% point increase in inflation, government spending goes up by 0.027% points.

Government debt also has a small magnitude effect (but statistically significant at the 1% level). For every 10% point increase in debt to GDP ratio, government spending goes up by 0.28% points. This might seem very small, but we need to remember that debt is a stock and spending is as flow and thus the ratio between them will appear small. Government spending is usually flat year after year (and by definition cannot be more than GDP) whereas debt is usually increasing. For example in 2008, US debt to GDP ratio was 67% but government spending 20.2% of GDP (ratio between them was 0.3). And in 2019, US debt to GDP ratio was 105% and government spending was 17.4% of GDP (ratio between them was 0.17).

Inequality does have a substantial effect (statistically significant at 1% level). A 10 point change in the Gini coefficient increases government spending by 1.2% points. This makes sense since a country that has a high level of inequality will require the state to provide for the poor who are comparatively more plentiful.

All the other remaining variables (Natural Resource Rents, Remittances, and

Table 3.4: Regression Output of Government Spending Model and Government Spending on Education Model

	(1) Govt. Spending	(2) Govt. Spending on Edu.
Corruption	−1.452***	−0.291***
	(0.108)	(0.039)
Urban Density	−0.000216***	−0.00012***
	(0.000054)	(0.00002)
GDP per capita	−0.000088***	−0.000020***
	(0.000015)	(0.000007)
Natural resource rents	0.009	−0.002
	(0.010)	(0.005)
Remittances	0.069	0.051**
	(0.045)	(0.022)
Inflation	0.027*	−0.002
	(0.015)	(0.003)
Government Debt	0.028***	−0.0054***
	(0.004)	(0.0017)
Gini	0.12***	0.003
	(0.02)	(0.009)
Democracy Duration	−0.0016	0.002
	(0.0044)	(0.002)
% Christian	0.025***	0.013***
	(0.006)	(0.003)
% Muslim	−0.003	0.0002
	(0.006)	(0.0028)
% Buddhist	−0.039***	−0.003
	(0.008)	(0.003)
Observations	1309	914
Adjusted R^2	0.4709	0.3717

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Colony Dummy Variables Coefficients - Regression Output of Government Spending Model and Government Spending on Education Model

	(1) Govt. Spending	(2) Govt. Spending on Edu.
Never Colonized	19.21***	6.10***
	(1.14)	(0.44)
Colony of Britain	16.53***	6.27***
	(1.35)	(0.49)
Colony of France	16.68***	6.61***
	(1.38)	(0.51)
Colony of Spain	12.21***	4.46***
	(1.39)	(0.50)
Colony of Portugal	19.14***	5.81***
	(1.60)	(0.56)
Colony of Netherlands	17.22***	5.75***
	(1.65)	(0.54)
Colony of Others	20.82***	6.54***
	(1.21)	(0.45)
Internal Colony of	22.30***	5.67***
Austro-Hungarian Empire	(1.04)	(0.36)
Internal Colony of	18.37***	4.67***
Ottoman Empire	(1.32)	(0.44)
Internal Colony of	19.81***	5.73***
Soviet Union	(1.15)	(0.42)
Observations	1309	914
Adjusted R^2	0.4709	0.3717

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Democracy Duration) are not statistically significant. The religion variables are just the usual control variables.

We now look at the colonial dummies coefficient in Table 3.5 (model 1 for Government Spending). The first that that pops out is that former Spanish colonies have much lower government spending (approximately 12% of GDP) compared to the higher average of the other European colonies (greater than 16.5%) . The other observation that sticks out is that Austro-Hungarian colonies have the highest government spending (22.30% of GDP) in this dataset. Other than that most countries seem to be within the narrow band of 16% to 20% of GDP as a fraction of their annual government spending. This may be a consequence of modern IMF policies which demand a limit on government spending on nations who wish to get cheap loans or aid for their economies. But this is not a clear answer and much more work would be needed to understand this.

Regression Results of Government Spending on Education

In Table 3.4, the second model regresses the Government Spending on Education (as % of GDP) against the other variables. The first coefficient is for corruption which states that for every 1 unit increase in CPI (Corruption Perception Index) the education spending goes down by 0.291 percentage points (this is statistically significant at 1% level. Given that the average state spending on education for all the countries in the dataset is 4.59%, this is quite a decent effect. A one unit increase in CPI has a 6.3% decreases in educational spending.

So (in 2008 year data) the difference between a very corrupt country with a CPI of 8.4 (Sudan) and a semi corrupt country with a CPI of 6.4 (Mexico), would have a 12.6% difference in educational spending. This if consistent on the long run will have a huge difference on human capital buildup and therefore economic growth.

Urban Density has an effect (significant at the 1% level). The coefficient says that for every 1000 people per square kilometer increases, education spending by the state goes down by 0.12% points. This is half the effect that density has on total government spending. The negative effect makes sense since expenditures in more dense areas affect more individuals and so not much spending is required.

GDP per capita has a small but substantial effect (significant at the 1% level). For every $1000 increase, state education spending decreases by 0.02 percentage points. This is four times smaller than the effect that a change in per capita GDP has on total government spending. Which makes sense, since as countries get richer (and if they are reducing government spending), education spending will fall slower compared to the fall in total government spending.

Remittances in observed to increases government spending on education, but by a small amount (statistically significant at the 5% level). For every 10% point increases in remittances, education spending by the state goes up 0.51% points.

Government debt also has a small magnitude effect but in the decreasing direction (statistically significant at the 1% level). For every 10% point increase in debt to GDP ratio, state education spending goes down by 0.054% point. A very small effect.

All the other remaining variables (Natural Resource Rents, Inflation, Gini, Democracy Duration) are not statistically significant. The religion variables are the usual control variables.

We now look at the colonial dummies coefficient in Table 3.5 (model 2 for Government Spending on Education). Like the previous model on total government spending, the same thing pops out. Former Spanish colonies and have much lower government spending on education (approximately 4.5% of GDP) . The same is true for internal colonies of the Ottoman Empire. All the other countries spend more than 5.5% of their

GDP on education with French colonies spending the most at 6.61% of GDP.

Conclusions

In this chapter, we have attempted to show the effect of corruption on economic variables that have an effect on the populace. We have focused on variables that have effect both in terms of economic growth and development of human potential. We also developed the models by incorporating urban density to control for the effect of that variable, so that we can see the effect of corruption much more clearly.

We studied four new variables: tax revenue, foreign direct investment (FDI), total government spending, and government spending on education. All the variables were measured in % of GDP.

The tax revenue decreased by 1.58% points for every 1 unit increase in CPI (Corruption Perceptions Index) and is a major effect statistically. Remittances had a large effect. A 1% increases in remittances increased tax revenue by 0.25%. Inequality also had a large effect. A 10 point change in the Gini coefficient decreased the tax revenue collected by 1.59% points. Long term colonial effects also exist. Spanish colonies were observed to collected less tax revenue (7.76% points less than British colonies and very statistically significant).

Foreign direct investment (FDI) decreased by 0.549% points for every 1 unit increase in CPI which is a decent effect. Natural resource rents and remittances have a large effect. FDI inflow increases by 0.613% points for every 10% point increase in natural resources rents. And FDI inflow increases by 1.296% points for every 10% point increase in remittances. Colonial effects are major. Internal colonies on average have an 8% FDI inflow compared to an average 6% FDI inflow for external colonies. Also, among the major European powers, French colonies have the largest FDI inflow.

Corruption has a large effect on total government spending. A 1 unit increase in CPI has a 1.452% point decrease in government spending. Since in this data set the mean government spending is 15.5%, a unit change in CPI has a nearly 10% decrease in government spending which is a substantial effect. Urban density has a statistical significant effect at the 1% level. For every increase of 1000 people per square kilometer, government spending goes down by 0.216 percentage points. This suggests more efficiency in spending due to density effects. GDP per capital has a small but substantial effect (at the 1% level). For every $1000 increase in GDP per capita, government spending decreases by 0.088% points.

Inequality has the other large effect. A 10 point change in the Gini coefficient increases government spending by 1.2% points. Which makes sense, since a country with a high level of inequality will require more state services to provide for the more plentiful poor. The only major colonial effect noticed is that former Spanish colonies have much lower government spending (approximately 12% of GDP) compared to the higher average of the other European colonies (greater than 16.5% of GDP).

Finally, corruption decreases government spending on education. A one unit increase in CPI has a 0.291% point decrease in education spending. Given that the average spending on education among all countries in the dataset is 4.59%, this is quite a decent effect. A one unit increase in CPI has a 6.3% decreases in state educational spending.

The only other variable that has a major effect is remittances. A 10% point increase in remittances increases the state education spending by 0.51% points. The major colonial effect observed is that former Spanish colonies spend less government money (4.46% of GDP) on education that other former European colonies (greater than 5.7% of GDP).

CHAPTER 4

MODELS OF SPATIAL EFFECT

Overview

In the previous chapters we have shown that urban population density has a major effect on corruption. In the next two chapters we will show why that is true using a few different mathematical models. In this chapter we will focus on spatial effect models. In these models the population is located in space and everyone can see whether others are corrupt or not. Depending up how their nearest (or even further) neighbors behave, an individual might end up acting on it and change their behavior from honest to corrupt or vice versa.

What we will show in this chapter is given that people affect each other, sometimes a very small change in the parameters of the populations can have a very large change in the gross properties (or explicit external perceptions) of a population. A change in parameters causes a change in dynamics and positive feedback effects can lead to cascade effects and shift a population to a very different direction.

There are three models that we will explore in this chapter. The Granovetter threshold model, the Schelling dynamic segregation model, and the Renormalization Group percolation model.

Granovetter Threshold Model

The first model that we explore is Granovetter (1978) threshold model of collective behavior.

In the simplest framework, the threshold model consists of a population of *n* individuals, each of whom has a choice to make: whether or not to participate in a certain movement. In Granovetter's 1978 paper, the population was 100 individuals milling around a city square and the decision was to whether or not to join a riot. Each individual has a 'threshold' or a level of participation of the surround population before they would join. The thresholds are randomly assigned from a given probability distribution. We will extend the model by redefining the population, instead of a population waiting to riot it is a population waiting to engage in corruption. This population does need to be able to see if all the others have or have not engaged in corruption.

The simplest probability distribution for thresholds is the uniform distribution. In this scenario there is one individual with threshold 0, one with threshold 1, one with threshold 2 and so on up to the last individual with threshold 99. So for example, the individual with threshold 2 will not engage in corruption until two other people have already started in being corrupt. For this distribution, the effect is very clear. There will always be a situation where the entire population is corrupt. Where everybody participate due the 'domino' or 'bandwagon' effect. The person with threshold 0 (also called the 'instigator' or 'initiator') will always engage in corrupt behavior (say taking or offering a bribe (in the Granovetter riot model it was breaking a window). This will activate the person with threshold 1. The actions of these two people will thus activate the individual with threshold 2, and so on, until all 100 individuals have joined in being corrupt. The equilibrium point in this scenario is always 100.

The uniform distribution of threshold is also very sensitive to perturbations (or accidents). If we remove the one individual with threshold 2 and replace them with an individual with threshold 3, then corruption does not cascade throughout the population. The initiator (threshold 0) turns corrupt, the individual with threshold 1 turns corrupt, but there is no individual with threshold 2 available to turn corrupt (the next individual is of threshold 3), so corruption ends at that point with just two corrupt individuals. For Granovetter this is an important insight, two populations that are essentially identically (very similar means and variance), have very different outcomes. So one cannot infer the individual dispositions of the population from aggregate outcomes. One cannot make claims that a corrupt population has underlying dishonest population of individuals versus a community that did not engage in corruption. This implication in true only if the underlying threshold distribution is uniform.

We can mathematically model how one goes from a frequency distribution of thresholds to an equilibrium outcome. We will denote the thresholds by $x,$ the frequency distribution of the thresholds by $f(x)$ and the cumulative distribution function (CDF) by $F(x)$. The CDF tells us the fraction of the population that has a threshold less than or equal to x. Now, by discrete time interval t, the proportion of the population that has become corrupt can be represented by $r(t)$. Assume for some time t we know what $r(t)$ is. For example at time $t = 2$, 40% of the population has joined in being corrupt. Then what fraction of the population will have joined the dishonest at time $t = 3$? By using the definition of thresholds, it must be exactly the fraction of the crowd whose thresholds are less than or equal to 40%. Therefore we can describe this process by the difference equation $r(t + 1) = F[r(t)]$.

Therefore for a given frequency distribution, we can solve the solve the difference equation to get an expression for $r(t)$ for any given t. Then, by setting $r(t + 1) = r(t)$

we can find the equilibrium outcome. However an explicit solution is only possible when the functional form is simple. For a more complicated function form (like a Normal Distribution) we can still compute the equilibrium by forward recursion. However, there is an easier and more insightful way to go about this issue. It can actually be done diagrammatically (Granovetter, 1978; Robertson, 2014; Wiedermann et al., 2020).

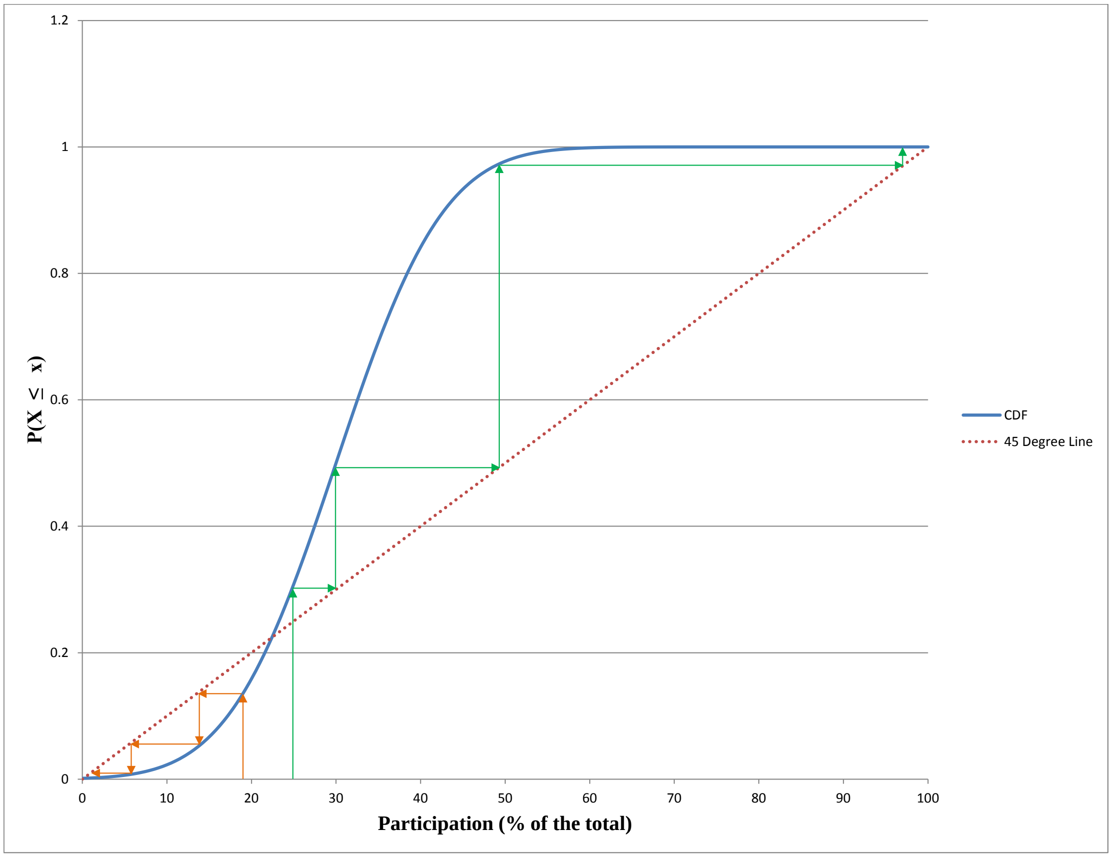

Figure 4.1: The evolution of Corruption participation for initial starting conditions of $p_0 = 19$ (orange lines) and $p_0 = 25$ (green lines). Underlying population is Normal(30,10).

In Figure 4.1, we graph the CDF $F(x)$ against the thresholds x. We use a normal frequency distribution (rather than a uniform distribution) since it is more realistic to describe thresholds of populations. In this particular graph (the blue curve), the normal distribution has a mean (μ) of 30, a standard deviation (σ) of 10, and a population (n) of a 100 individuals. The red dotted line is the $45°$ line. This $45°$ line is to be understood as a line of equilibrium values, since by definition for any point on this line a particular threshold value exactly matches the percentage of the population with that threshold.

We can now build a cobweb plot and find the fixed points. Let us say at time $t = 0$, the population starts being corrupt with 25 individuals participating or $p_0 = 25$. If we follow the first green arrow to the CDF we find we are strictly on the part of the curve that is above the $45°$ line. This means that the percentage of people in this population whose threshold is 25 is more than 25%, it is about 31%. Therefore more individuals will be 'activated' to join the dishonest which will grow to 31 individuals or $p_1 = 31$ for time $t = 1$. If we continue to follow the green arrow and observe the CDF when participation is 31% the number of people with that threshold is more than 31, it is about 51%. Therefore at time $t = 2$, the dishonest will grow again to $p_2 = 51$. This process continues (the green arrows) until the entire population of 100 people have joined the dishonest. This is the equilibrium point ($p_{final} = 100$) and a very counterintuitive result. For any starting $p_0 > 22$ (the intersection point of the CDF with the $45°$ line) the dishonest will keep on expanding until it goes to the maximum of 100 individuals in this particular population.

On the other hand, if we start at $p_0 = 19$, and follow the orange arrow to the CDF, we see that only about 14% of the population has a threshold of 19 or less. This means there are at least 5 individuals who are below their threshold and will want to leave in being dishonest. In other words we are now in the section of the curve that is strictly below the $45°$ line, so individuals will start to abandon the corrupt population and it will begin to shrink. In the next time iteration the number of individuals among the dishonest will be 14, but only 5% of the population has 14 as their threshold. So 9 people will leave. The process of people leaving will continue until all the individuals will abandon being corrupt and it will all end. This is the other equilibrium point ($p_{final} = 0$). For any starting $p_0 < 22$ corruption will keep on shrinking until it goes to 0.

$p_0 = 22$ is also an equilibrium point (or fixed point). If the dishonest population starts with exactly 22 individuals, it will neither shrink nor grow but stay exactly at 22. However this equilibrium point is unstable and any perturbation will cause the equilibrium to shift and move to either of the two other stable equilibrium points.

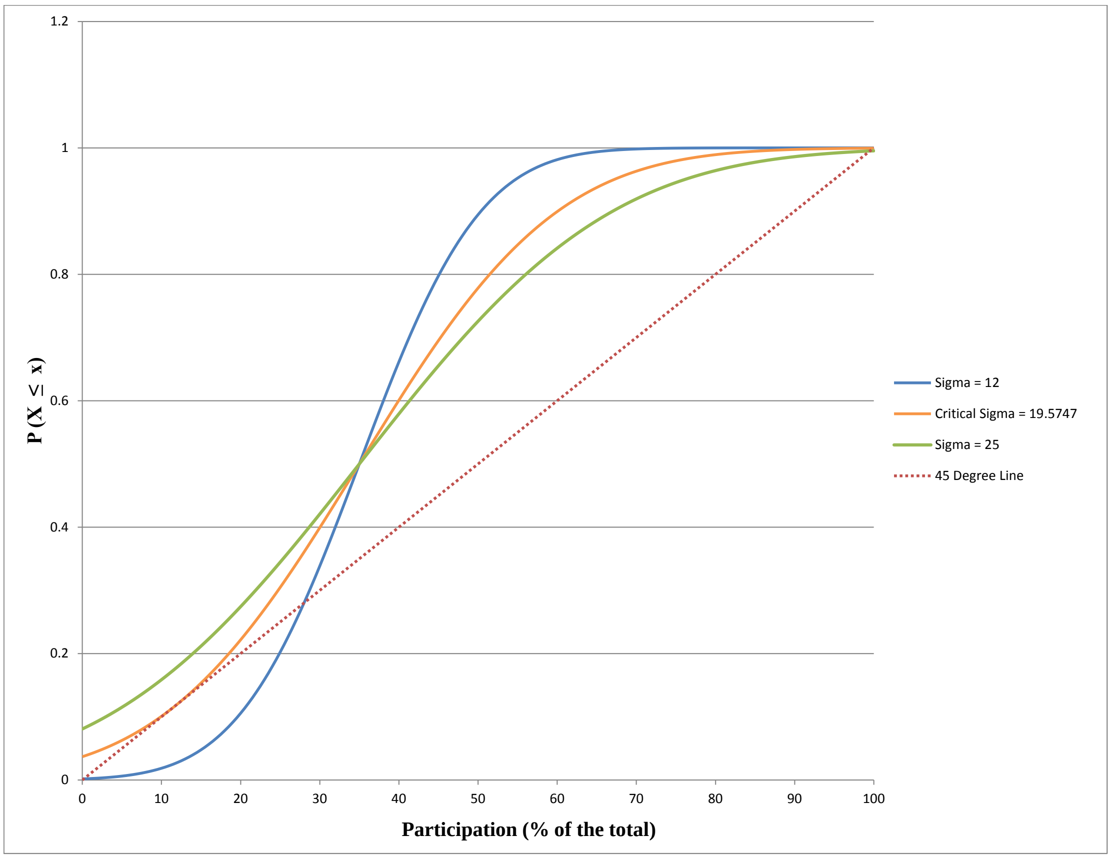

Figure 4.2: CDFs for $\mu = 35$ and three different values of σ

Figure 4.2 shows CDFs for a few distributions of $\mu = 35$ and different values of σ. One of the facts of the normal distribution is that the CDF intersects the $45°$ line either one, two, or three times. So given that the $45°$ contains all the equilibrium points possible, it implies that only one, two, or three equilibria are possible for any given normal distribution. When $\sigma = 12$ (the blue curve) there are three equilibrium points at $p = 0$, $p = 28$, and $p = 100$. This was the scenario that we describe above. When $\sigma = 25$ (the green curve) there is only equilibrium point at $p = 100$. If we look at the curve we will notice that the entire CDF curve is above the $45°$ line until it intersects at $p = 100$. This implies that for any starting p_0 the population of the dishonest will expand until the entire population participates in corruption.

Given these two observations and the fact that he CDF will intersects the $45°$ line either one, two, or three times; we can define a critical value σ (for a given μ) for which there will be exactly two equilibria. In Figure 4.2 this is the orange curve in the middle and for $\mu = 35$ the value of the critical σ^* is 19.5747. Above this value there is only one equilibrium where there is 100% participation and below this value there are three equilibria (two stable and one unstable). Thus σ^* represents an interesting cutoff behavior. When $\sigma > \sigma^*$, any initial starting participation will lead to the same equilibrium of 100% corruption participation.

To calculate the critical σ^* for any given μ two conditions needs to be satisfied. First, the CDF must intersect with the $45°$ line, which is the equation $y = 0.01x$. Second, the CDF must be tangent to the $45°$ line. Now the slope of the CDF is given by the PDF (in this case the Normal Distribution), so we can set the PDF equal to 0.01 which is the slope of the $45°$ line. The equations are

$$0.01 = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{4.1}$$

$$0.01x = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]. \tag{4.2}$$

Here erf is the error function. This is a nonlinear simultaneous equation model. The easiest way to actually solve this is using the Solve package in Excel for every value of μ. This is given in more detail in Appendix G.

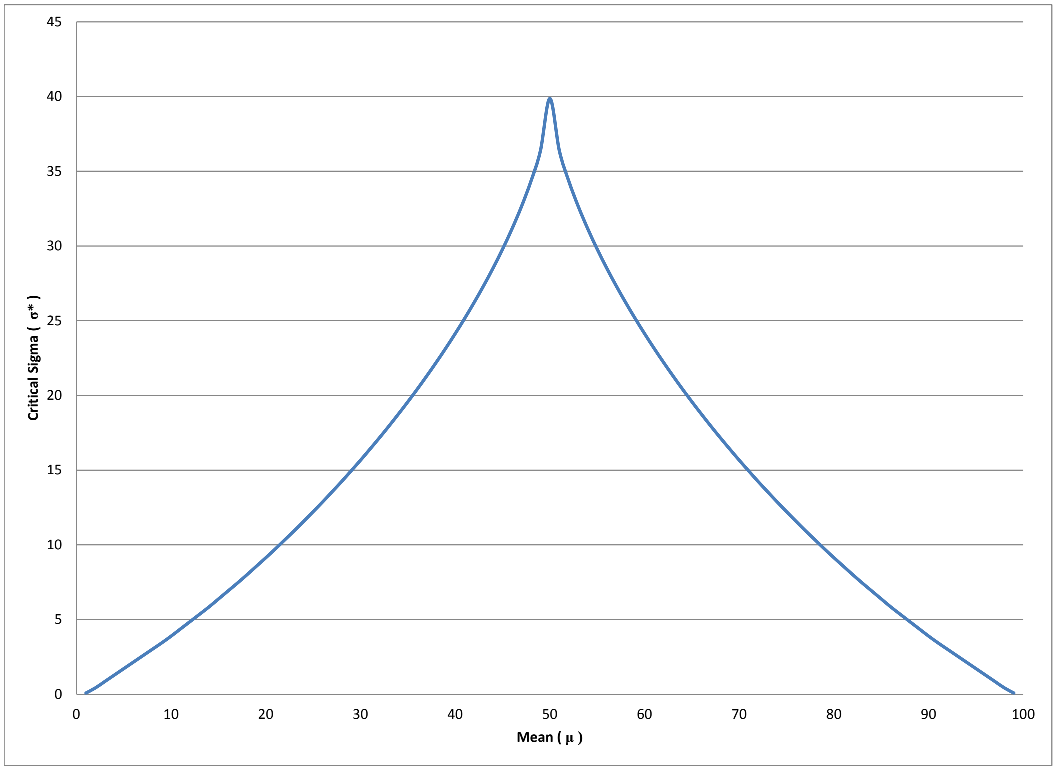

Figure 4.3: Critical Sigma σ^* for μ between 1 to 99. Underlying population distribution is Normal.

Using the values calculated (as shown in Appendix G), Figure 4.3 plots the value of critical sigma σ^* for values of μ from 1 to 99. As we can see the graph is symmetric and the critical sigma σ^* is very low for small values of μ and then increases rapidly until it peaks at a value of 40 (when $\mu = 50$) and then goes down again. What this means is, for values of $\mu < 50$ if the value of sigma is greater than critical sigma, then corruption will cascade through the community and the population will stabilize at 100% dishonest. And for values of $\mu > 50$ if the value of sigma is greater than critical sigma, then honesty will cascade through the community and the population will stabilize at 100% honest. For corruption analysis we are more interest in the left side of the graph.

A crucial thing to observe in the graph is that as the mean decreases, the critical sigma decreases faster. For very small values of μ like 10, the critical sigma σ^* is something like 4. This implies that it is very possible for a nation to be overall honest year after year while having a not so high threshold of corruption perceptions and assume things are safe and stable when all of a sudden it appears that corruption has taken over the population even though the mean perception (threshold) has not changed. This is because the variance could have been decreasing year after year without it being noticed until one year it crossed the critical sigma value, which leads to a sudden and unexpected predominance of corruption within the overall population.

This is the advantage of the Granovetter model in that it predicts the possibility of sudden transitions in the population's behavior. Figure 4.4 is a plot of the standard deviation of the corruption perceptions of every country on the y-axis against the mean corruption perception of every country (10 years of data on average per country) on the x-axis. What we observe is that very corrupt countries (CPI $>$ 7) and very honest countries (CPI $<$ 3) have a low standard deviation $($between 0.1 to $0.5)$ in their corruption perceptions (with the caveat that there are more corrupt countries and their variance is wider that honest countries).

We can use the CPI standard deviation as a potential proxy for the populations' unobserved threshold standard deviations (or σ). So the corrupt countries may have a low σ but they also have a low threshold (a low μ) so their low σ is actually above the critical sigma value which leads to corruptions affecting the whole of the nation. Similarly, in the

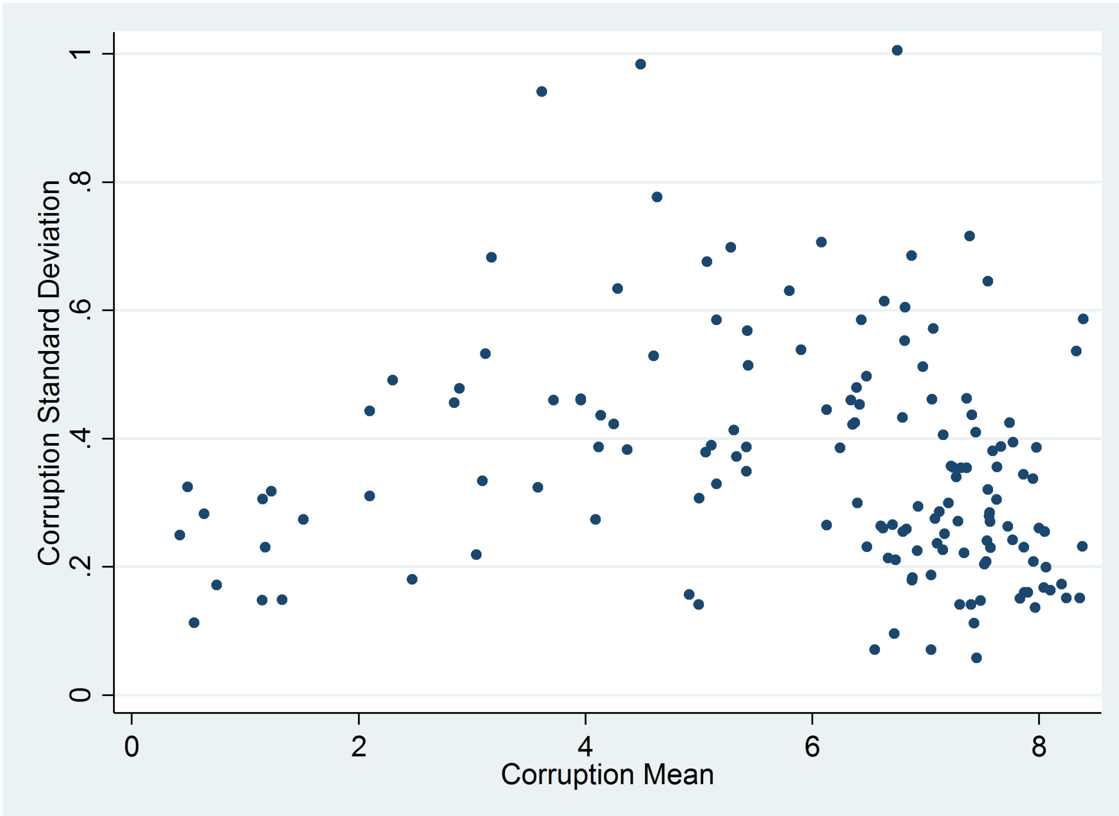

Figure 4.4: Plot of the Corruption Perceptions Standard Deviation against the Corruption Perceptions Mean (data over 10 years for most countries).

opposite direction for very honest countries. It is the middle countries that are the most interesting. The have a much higher variance than we would imagine otherwise. Since they are not really corrupt they have a high corruption threshold (or high μ), but they also have a much larger σ than expected.

One implication is that in these countries the opinion is still evolving among the various individuals as to whether it is worth moving to being corrupt or not (their threshold values are changing), which is reflected in the high variance in the corruption perceptions of the population when they are polled. The dilemma is that the mean could stay stable and high (thus giving the impression to the countries populace and its state authorities) that corruption will not take hold in the country. But if the standard deviation goes high enough and crosses the critical sigma, it could lead very quickly to the entire country to engage in corruption. This Granovetter model is thus very important in recognizing how quickly corruption can take over a country without anyone realizing it or thinking they have enough time to impose institutional structures to prevent it.

Thus one would consider countries whose corruption mean is between 3 to 7 (the not-so-corrupt countries) and pay attention to their standard deviations. Any country with a standard deviation greater than 0.5 is worth looking at more carefully. Figure 4.5 shows the same graph of the corruption perceptions standard deviation against the corruption perception means along with the country names, so that one can observe which countries can be potentially at risk. Since the country labels might be hard to view, Figure C.14 and C.15 in Appendix C have a more detailed view. Figure C.14 presenting for countries with corruption mean between 0 and 6 and the C.15 presenting for countries with corruption mean between 6 and 10.

There are quite a few countries in this subset (corruption mean between 3 to 7 and standard deviation of 0.5) and they are not exactly the same as the subset of countries that are more corrupt that the model predicts which we explored in chapter 2. Thus this gives us a bigger subset of countries to explore and consider what is different regarding their institutional characteristics.

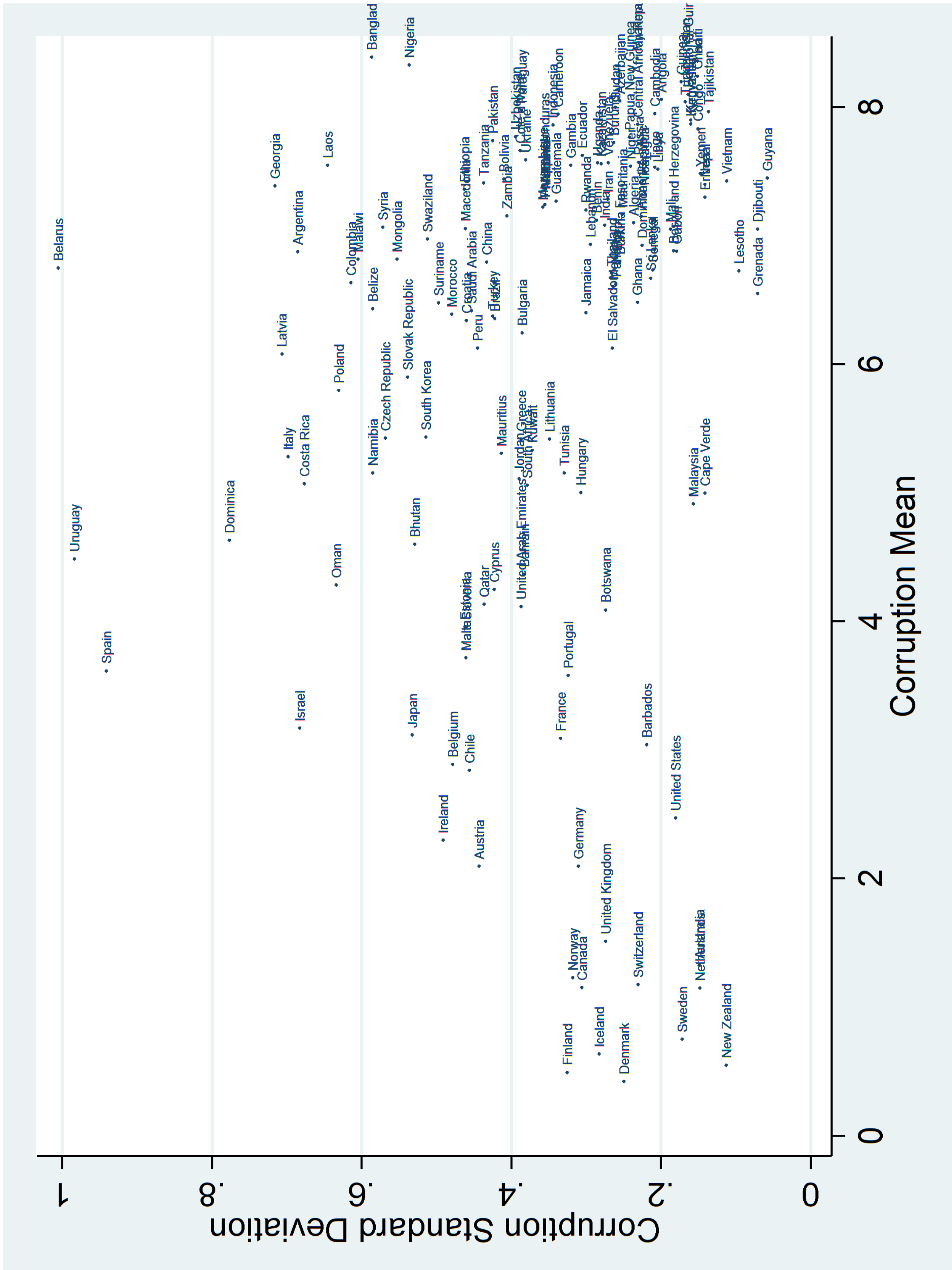

Figure 4.5: Plot of the Corruption Perceptions Standard Deviation against the Corruption Perceptions Mean Showing the Country Names

Schelling Dynamic Segregation Model

The second model we will explore is the Schelling Segregation Model. Schelling (1971) developed this mathematical model to show how using just individual preferences and free movement, the model could explain why urban areas ending up segregating itself by race. We can use the same model to show how corruption ends up taking over an area.

The model starts with two major assumptions. Everyone can see everyone else's behavior within a region – whether they are honest or corrupt. And everyone is free to move into or out of the area.

Every individual has an innate, unchanging preference for honest or corrupt neighbours that is called tolerance. By tolerance we imply the mix of individuals that a particular individual is willing to interact regularly with. For example if an individual is honest - it is obvious that they do not wish to interact with the corrupt. On the other hand, for the corrupt, while they might not object to interact with someone who is honest, there is no payoff in that interaction. So they would prefer to interact with more of their own kind (corrupt) that would ideally lead to more opportunities for payoff.

Consider an example scenario of a straight-line (linear) distribution of tolerance. This is the simplest (and quite insightful) distribution that we can work with. Let the total number of corrupt and honest individuals be 100 each. In Figure 4.6 the horizontal axis shows the total number of corrupt individuals. The vertical axis is the 'tolerance' ratio for honest individuals by the corrupt; it shows the ratio of honest to corrupt that represents the upper bound of the tolerance that a corrupt individual has.

In this particular graph, the median corrupt individual is willing to regularly interact with honest individuals half the time. Or in other words they will abide an honest to corrupt ratio of 1.0 or greater. The most tolerant corrupt individual can abide an honest-

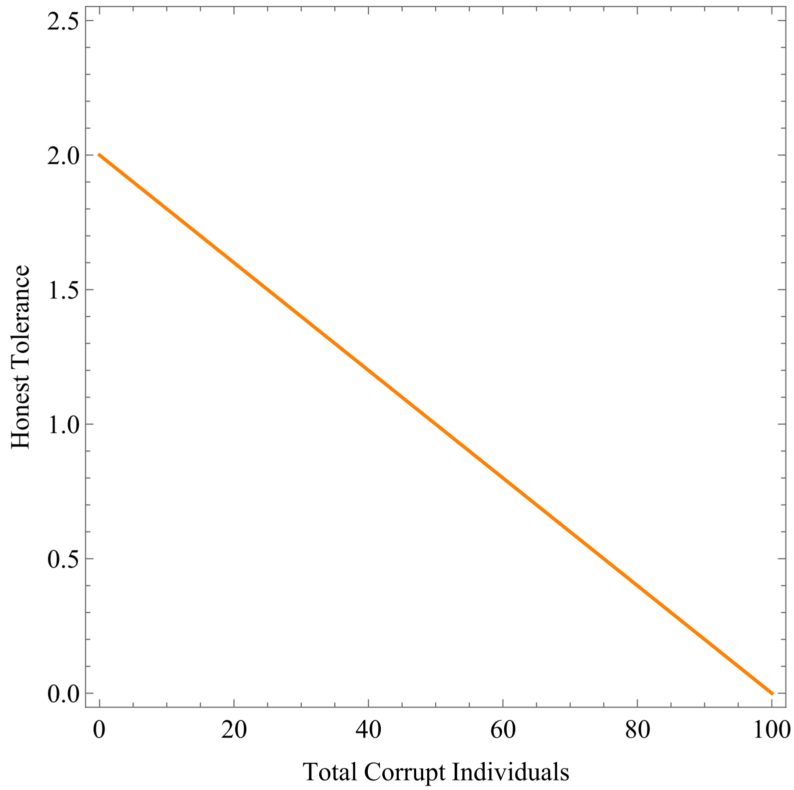

Figure 4.6: Straight Line Distribution of ‘Tolerance’ for Honest Individuals by Corrupt Individuals

corrupt ratio of 2.0, or they are willing to be in a one-third minority. Thus we can see the graph has a vertical intercept of 2. The least tolerant corrupt individual cannot accept any honest individuals to interact with, so their tolerance is zero. Thus the straight line crosses the horizontal axis at 100. In mathematical terms, the tolerance can be described by the equation

$$H_{tolerance} = a - m \times C \tag{4.3}$$

$$m = \frac{a}{p} \tag{4.4}$$

Where a is the vertical intercept (in this example $a = 2$) and m is the gradient of the tolerance curve. m is defined as (a/p) or the maximum tolerance divided by the total population of the subgroup. In this example the subgroup is the corrupt, $p = 100, a = 2$, therefore $m = 0.02$. The gradient is defined in this way so that when the linear curve hits the value of $C = 100, H_{tolerance} = 0$ (in other words the least tolerant corrupt individual has a tolerance of 0). A similar equation (and graph) can be given for the other subgroup (the honest) in this model.

Now given a tolerance distribution we can find how many of each subgroup are willing to interact with each other in total. So we have to convert the tolerance schedule into an absolute numbers curve. So the total number of honest (H) individuals that the corrupt (C) are willing to accept are

$$H = H_{tolerance} \times C = (a - mC)C$$

or

$$H = aC - mC^2. \tag{4.5}$$

We can plot this equation in Figure 4.7. The orange parabola shows the total number of honest individual in the community the corrupt are willing to interact with as the size of their population increases. And similarly the blue parabola for the total number of corrupt individuals the honest are willing to allow into their communal interaction. We can observe an area of overlap of both the parabolas. Any point within that area of overlap indicates a statically viable combination of honest and corrupt. That is there are that many corrupt who are willing to engage with the honest and that many honest who are willing to engage with the corrupt.

Any point to the right of the blue curve but beneath the orange curve indicates a mixture of corrupt and honest in which all the corrupt are satisfied but not all the honest

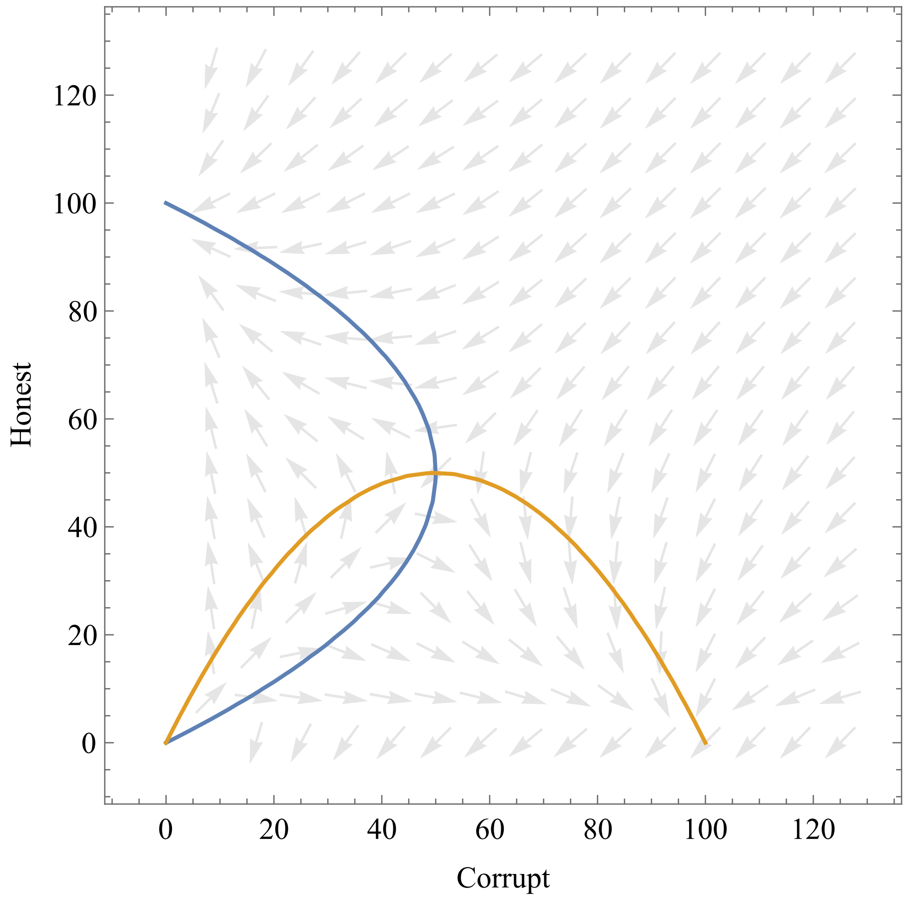

Figure 4.7: The total number of Honest that the Corrupt are willing interact with (Orange curve) and the total number of Corrupt the Honest are willing to interact with (Blue curve)

(some of the honest will be satisfied, just not all of them). A similar case (with the roles flipped) applies to below the blue curve and to the left of the orange curve. Any point that lies outside both curves (the upper right quadrant) indicate a mixture of both corrupt and honest where there exists individuals of both parties who are unsatisfied and would prefer not be in such a mixture.

Not all combinations of tolerance greater than zero leads to a statically viable combination (or area of overlap). In Figure 4.8 the tolerance of both parties is 1.0 ($a = 2, m = 0.02$) or the most tolerant corrupt (and similarly honest) individual can abide a honest-corrupt ratio of 1.0, or they are willing to be in a 50:50 population mix. In such a case we see that the curves have no overlap at all. There exists no statically viable population mixture where both the corrupt and the honest are satisfied. Under the orange curve, everyone corrupt is satisfied but not all honest are. And a similar analysis can be applied under the blue curve with the roles reversed. Under the blue curve, everyone honest is satisfied but not all the corrupt are.

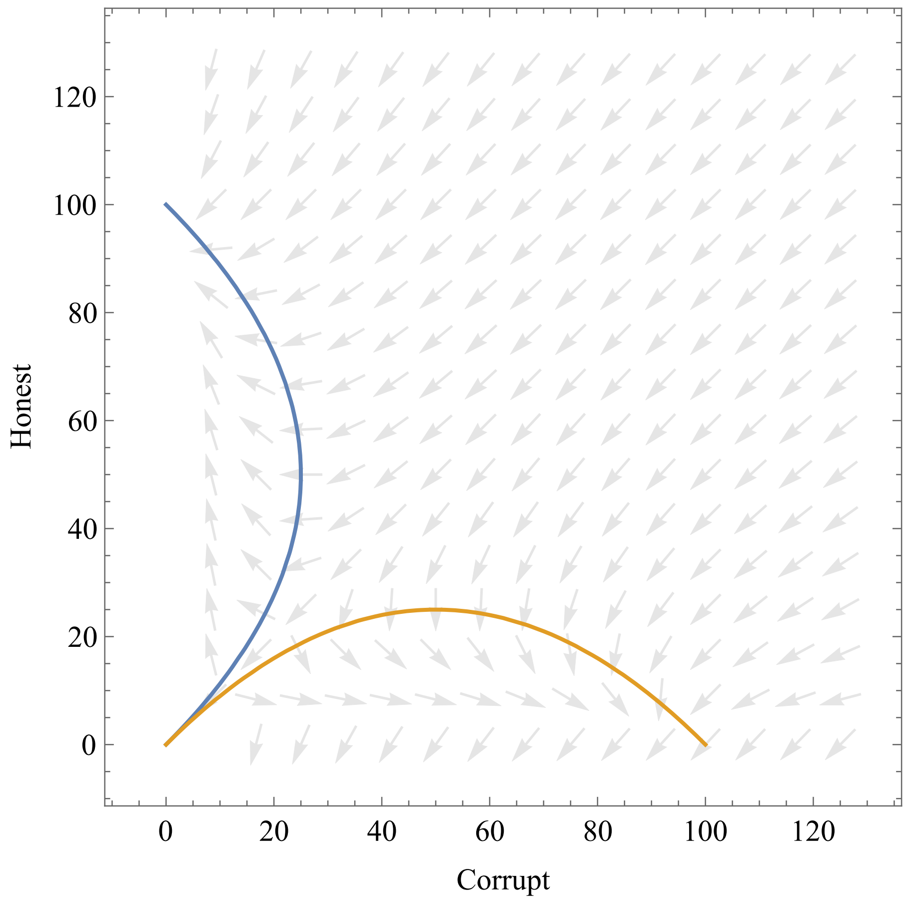

Figure 4.8: Population configuration when tolerance of both parties is 1.0

So far everything we have been discussing is static viable combination. But the major insight of Schelling (1971) was that no matter the static configuration, it is the dynamics that has a much bigger effect. The simplest dynamic would be if all the corrupt in a particular region are content and those outside the region would be content if they moved into the region. Then we will observe movement into the region as long as those inside are still content and there exists some outside who would prefer to move.

In contrast, if all the corrupt within a particular region are not satisfied, then some will leave in the order of their dissatisfaction. Those that remain will be more tolerant than those that left, and when the ratio of the honest to the remaining corrupt rise such that the remaining corrupt are all satisfied, then no more will leave. One can generate a similar rule that controls the exit and entry of the honest individuals.

Given such a rule, we can now plot for every point in the population diagram the vector of population change (Montgomery, 2009). These are the grey arrows in the graphs we have seen so far. What we observe is that within the overlapping region of the two curves the numbers of honest and corrupt are both increasing and the approach the equilibrium point of (50 Corrupt, 50 Honest). However this is *not* a stable equilibrium point. Any perturbation leads to the population vector to move away from it (the gray arrows point away, except for that that come from the overlap region). For any point within the corrupt (orange) curve but outside the honest (blue) curve the honest will be leaving and the corrupt will be coming in leading to the equilibrium point of (100 Corrupt, 0 Honest). Similarly within the honest (blue) curve but outside the corrupt (orange) curve the corrupt will be leaving and the honest will be coming in leading to the equilibrium point of (0 Corrupt, 100 Honest). These two points are the only two *stable* equilibrium points.

Thus even though the tolerance was high enough, such that a mixed population of honest and corrupt could interact and coexist (the overlap region), in the long run the two populations separate themselves due to i) any minor perturbation ii) the influx of individuals from outside who prefer the existing ratio in the shared space but by the very fact of moving in changes the ratio and thus sets of a cascade that leads to the population separation. This shows that a tolerance of 2.0 by both parties is not high enough.

Is there a tolerance value where we can have a stable equilibrium of both populations? Yes, we can increase the tolerance to 5.0 for both parties for each other. This means that the most tolerant corrupt is willing to accept a ratio of 5 honest to corrupt (or willing to be an $\approx 17\%$ minority in any region). This also means that the median corrupt is willing to tolerate a ratio of 2.5 honest to corrupt (or willing to be an $\approx 29\%$ minority). A similar analysis applies for the honest.

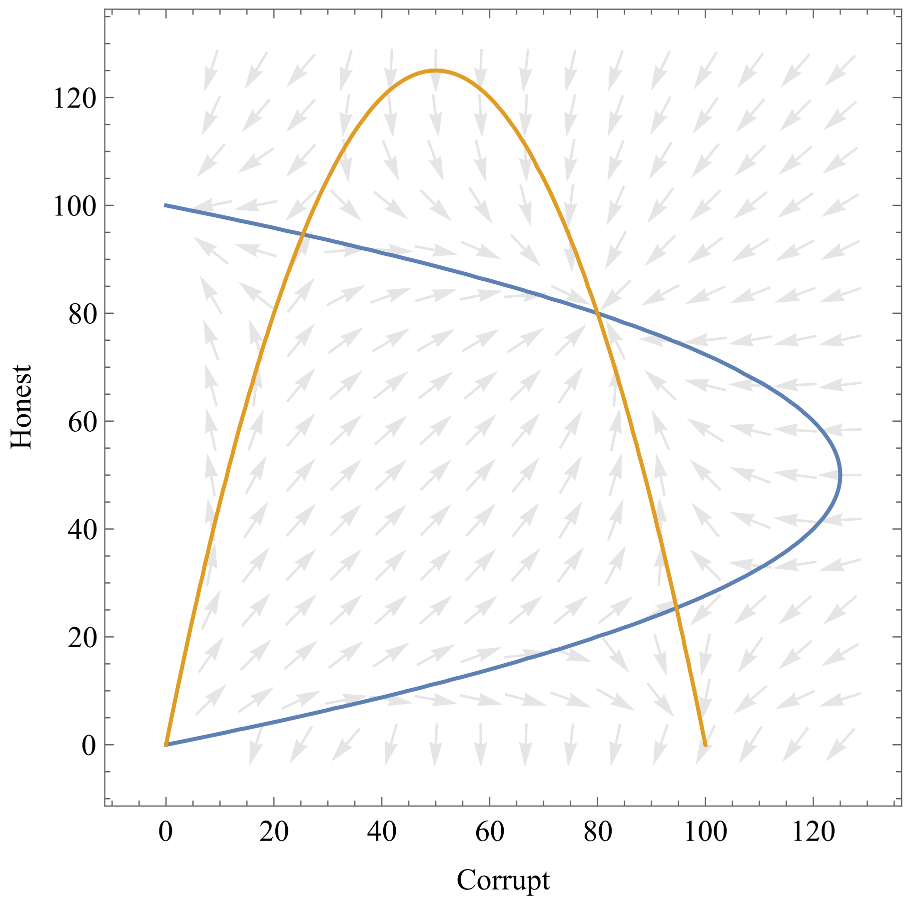

Figure 4.9: Population configuration when maximum linear tolerance of both parties is 5.0 and a stable equilibrium point at (80,80)

Figure 4.9 shows what happens when we plot the population curves. We now have three points of intersection of the two curves. Two of the intersection points (25.4 Corrupt, 94.6 Honest) in the upper left quadrant and (94.6 Corrupt, 25.4 Honest) in the lower right quadrant are unstable equilibrium points, as we can see the grey arrows move away for it. But one point, (80 corrupt, 80 honest) is a stable equilibrium point as we can see all the grey arrows around it converge to it. There are also two other stable equilibrium points. (100 Corrupt, 0 Honest) and (0 Honest, 100 Corrupt) are equilibria where the population only consists of one type of individual. But while these two are stable equilibria, only very few starting population states will end up moving to those two points. Most of the initial starting states, as we can see from the grey arrows, actually converge to the mixed population (80 Corrupt, 80 Honest) through movement of the respective sub-populations.

As long as slightly more than 40% of both types exist, the dynamics of entry and exit will always lead to a the stable mixture of (80 Corrupt, 80 Honest). This is even true if the starting population is very small as we can see in the lower left quadrant (which is part of the overlap region). Also (0 Corrupt, 0 Honesty) is not an equilibrium point, the total population is always greater than zero. And given how large the overlap region is and how large the attraction domain of (80 Corrupt, 80 Honest) is as we can see from the grey vector arrows, we can also state that this equilibrium point is very stable against fairly large perturbations.

A major implication of the Schelling model is that total segregation happens when the maximum tolerance level is 3 or below. This was shown by Schelling (1971) and also can be simulated numerically. If a linear tolerance model has a maximum of 3, then the median tolerance is 1.5. Or in other words the honest (and corrupt) are willing to be at most a 40 % minority in society (1 honest / (1 honest + 1.5 corrupt) = 0.4).

When the tolerance level increases, it implies a willingness to be an even smaller minority in the society. Here is the crucial aspect. As long as the median honest want to live in a 40% or more honest society, the honest and corrupt population are stably separated from each other and do not interact. But once the median honest are satisfied in living in less than a 40% honest society, then suddenly the stable population is 50% corrupt and 50% honest. The different populations are interacting with each other.

This suddenness is the main point. Societies are surprised when corruption suddenly seems to be accepted in their communities. The societies' mental expectation has always been that if corruption does happen, it should have grown slowly in a population such that it could be detected and the society could use its institutional power to stop it or control it. What the Schelling model shows is a population is either totally segregated or suddenly in 50/50 mode (in the time frame of people changing interaction behavior which is much shorter than a generation and definitely in the average span of how often people change residences (lets say approximately 7 years)).

It is in its ability to explain this suddenness of population composition change that the Schelling model is worth using. It complements the Granovetter model. Also note that no matter how high the tolerance gets (of the willingness to be a minority smaller than 40% of the total population), the stable population is always 50/50 of both honest and corrupt.

There is also a different issue to consider. Let us say the population is in 50/50 integrated mode. Since the median honest has a tolerance of 3 and higher (or willing to be a minority of 40% or lower in the population), they are satisfied since 40% is not that much lower than 50% and thus they can imagine that things are going alright. The problem is the longer the population stays in such a situation it enables the following generation to think this is the normal state of affairs and then they will increase their tolerance even further (accept being a smaller and smaller honest minority in the nation) and thus in the long run can lead to corruption being harder and harder to reduce in the nation.

Renormalization Group Percolation Model

The last model we are going to explore is the Renormalization Group Percolation model. Renormalization group was an idea that arose in particle/statistical physics and was taken up in other fields. Kadanoff (1966) was crucial in developing it in such a way that made it tractable to use in different domains.

Percolation theory began in fluid flow physics but quickly became a subject in probability theory. Percolation is one of the simplest models in probability theory (Grimmett, 1999) that can show critical phenomena. This means that there exists a parameter in the model for which the observed global behavior in the model drastically changes. Many methods are used to find this critical parameter for which the behavior change in the model is observed.

In the late 1970's the ideas of renormalization group was brought into percolation theory to give a different way to find the critical parameter for simple two dimensional models. A few decades later these ideas get more exposure and popularity when Stauffer & Aharony (1991) publish their famous book 'Introduction to Percolation Theory'. A large part of the understanding of the modelling framework here are adapted from Stauffer & Aharony (1991) and Sayama (2015).

In this model framework we begin with individuals distributed over a two dimensional lattice and are neighbors to each other. The model will argue that neighbors affect each other. If a neighbor is not corrupt, then all the neighbors around them will also not be corrupt by definition. But if a neighbor is corrupt, then the neighbor next to them might also become corrupt by some probability (or propensity) amount.

For a particular population (in a particular snapshot in time), this probability parameter is fixed, it does not evolve and it does not need to change for the model to work. Given that probability parameter, then there is a likelihood that one neighbor affects the next neighbor, which affects the next neighbor, and so on, until the whole area is corrupt. So one can view the probability parameter as the propensity of an individual to turn corrupt given its neighbors has already turned corrupt. The question this modelling framework wants to solve is what is the minimum probability amount, beyond which it ends up converting the entire neighborhood to corruption.

Thus the analogy to the forest fire. If one tree is burning and it can affect other trees, how resistant do the trees have to be to burning to prevent the whole forest from being ablaze. A simple 2 dimensional lattice is used because in a one dimensional lattice even one tree not burning can stop the forest fire. But in a two dimensional lattice there is no guarantee that a single non-burning tree can stop the trees after it from burning, since many more paths are available to just go around the non-burning tree.

A similar argument can be made with human neighbours. Even if one neighbor is adamant in not engaging in being corrupt, there are links around them that are still possible and thus have an effect on further located individuals. So for different lattice dimension what would be the probability parameter?

In a 1-dimension (1-d) lattice (one neighbor after another in a straight line),the answer is simple. The probability has to be 1. Even if one individual chooses to resist corruption, it will cause everyone after them on the 1-d line to continue to stay honest. This is because an individual has one direct neighbor it can be affected by. Lets say in a neighborhood line (lattice) running West to East (corruption starting at the West-most point), an individual can only be affected by the neighbor to the West (left) and can only affect its neighbor to the East (right). So the only way the 1-d neighborhood goes entirely corrupt is when the propensity of corruption for every individual is only 1.

The problem arises in 2-d (2-dimensional) models (and of course even in higher dimensional models - but those models are intractable and we will not consider them). In a 2-d model an individual has more than one direct neighbor that it can be affected by (and also affect in return). It can be affected by the neighbor to its West. But also by the neighbor to the North, South and East. Even if the neighbor to the West is not corrupt, the individual can be affected by longer chains that avoid (go around) the Honest neighbor of the West. There is more than 1 possible connection link to the corrupt neighbor that is 2 steps (houses) West of the individual. Let us try to develop the model more precisely and mathematically.

We begin with the idea that the entire population can be modelled as a lattice of 1x1 cells on a finite 2-dimensional plane. Each cell contains an individual that can be corrupt with a given probability q. This probability is the same for all individuals in the lattice. Each cell has 8 neighbours, 4 direct and 4 on the diagonals. The question we are asking is if there exists a corrupt individual at one edge of the lattice, does the corruption propagate to the entire lattice (reach the other edge).

If we view the lowest level of the lattice, we have a 1x1 cell (or a single individual). Now the probability that this cell will propagate corruption is just the probability that the cell (or individual) is corrupt, so $p_1 = q$. Now, we need to find what happens at the next level of the 2x2 block (or a collection of 4 individuals next to each other).

We have to use the 1x1 block as the smaller sized building block and p_1 as the fundamental property of that block. Next we enumerate all the possible arrangements of the 1x1 building block that will propagate corruption across the 2x2 block. This can clearly be seen in Figure 4.10 where the grey squares have turned corrupt and the white

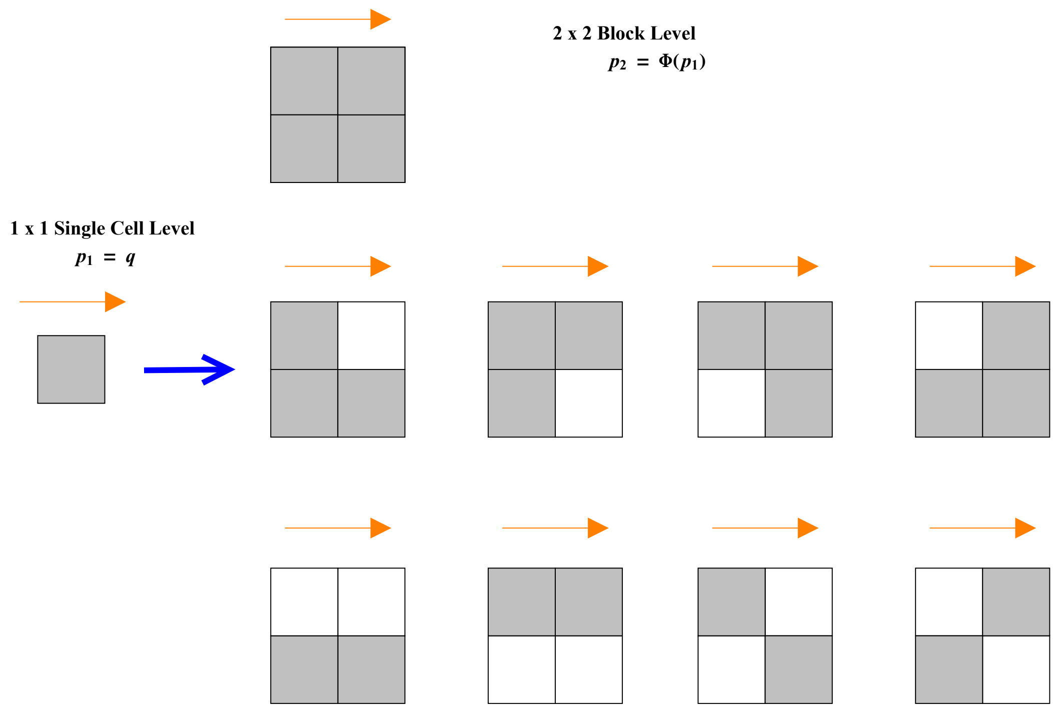

Figure 4.10: Estimating the corruption propagation probability p_2 at the 2x2 bock level given the probability p_1 at the 1x1 single cell level.

squares are not.

At the top of the figure, if all four cells (individuals) are corrupt, then corruption will definitely be propagated to the other side of this sub-lattice (this is schematically represented by the orange arrow). All that is needed is at least one cell in each row and column to be corrupt for corruption to be able to propagate to the next 2x2 block. If only 3 cells only or 2 cells only are corrupt, there are still four different ways each for corruption to be propagated.

But if only one cell in the 2x2 block in corrupt, there is no possible way to propagate corruption. And of course that is also true if none of the cells in the 2x2 block is corrupt. So we do not have to consider those cases. As we can see in Figure 4.10, there are only 9 possible configurations (out of 16 total configurations) that would allow for the 2x2 block to be corrupt enough such that they can propagate corruption to other 2x2 blocks. We can now calculate the probability for each sub-situation (a p_1 for each gray square and a $(1 - p_1)$ for each white square) for the 9 blocks and sum them up to get the following relationship between p_2 and p_1:

$$p_2 = \Phi(p_1) = p_1^4 + 4p_1^3(1 - p_1) + 4p_1^2(1 - p_1)^2 \tag{4.6}$$

At the next step is where we make the renormalization approximation (Stauffer & Aharony, 1991). The key idea of renormalization is of self-similarity at the critical point on different scales. That is what is true at the smaller scale is also true at the larger scale and thus the properties at the larger scale can be approximated (renormalized) with the same properties of the smaller scale. Different lattice sizes are similar to each other and share the same mathematical properties.

In Figure 4.11 we can see the 2x2 block being used as a building block for the

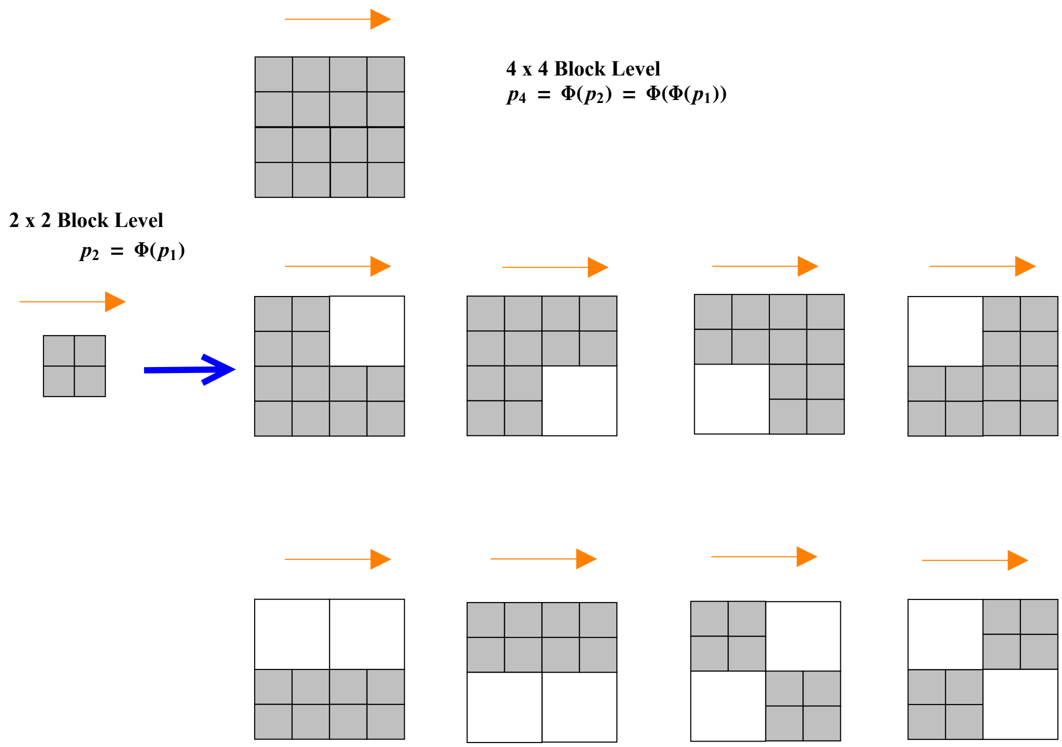

Figure 4.11: Estimating the corruption propagation probability p_4 at the 4x4 bock level given the probability p_2 at the 2x2 block level.

4x4 block level. We can now claim according to renormalization group theory that the 4x4 block can easily be consider as a 2x2 block, just larger. But with the same dynamics. Similarly, we can keep on applying this relationship to 8x8 blocks, 16x16 blocks, and so on onto the entire neighborhood lattice. So that for all s:

$$p_{s+1} = \Phi(p_s) = p_s^4 + 4p_s^3(1 - p_s) + 4p_s^2(1 - p_s)^2 \tag{4.7}$$

Once we have this equation we can draw a cobweb plot for this equation. Figure 4.12 shows the cobweb plot. The blue curve is the plot of equation 4.7 and the red line is the $45°$ line or $y = x$ line.

To find the critical corruption propagation probability we set $p_c = \Phi(p_c)$ and find the roots.

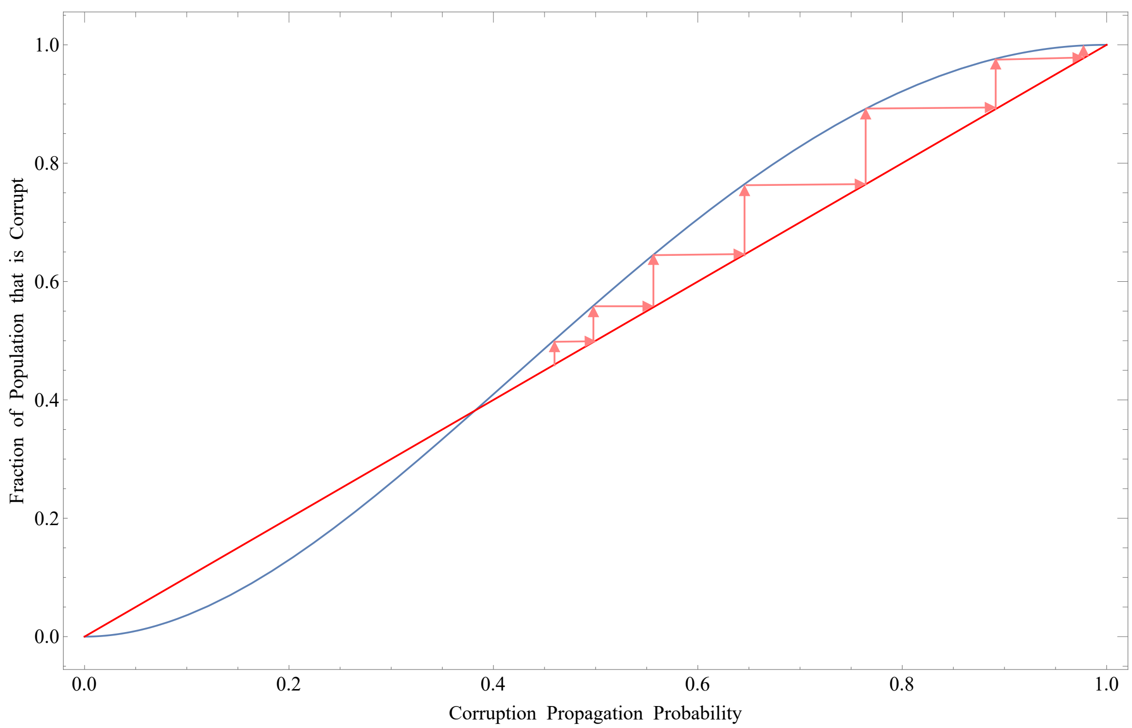

Figure 4.12: Cobweb plot showing the three fixed points of the critical value of the corruption propagation probability.

$$
\begin{aligned}
p_c = \Phi(p_c) &= p_c^4 + 4p_c^3(1-p_c) + 4p_c^2(1-p_c)^2 \\
0 &= p_c^4 + 4p_c^3(1-p_c) + 4p_c^2(1-p_c)^2 - p_c \\
0 &= p_c(p_c^3 + 4p_c^2(1-p_c) + 4p_c(1-p_c)^2 - 1) \\
0 &= p_c((p_c^3 - 1^3) + 4p_c^2(1-p_c) + 4p_c(1-p_c)^2) \\
0 &= p_c((p_c-1)(p_c^2+p_c+1) + 4p_c^2(1-p_c) + 4p_c(1-p_c)^2) \\
0 &= p_c(p_c-1)((p_c^2+p_c+1) - 4p_c^2 - 4p_c(1-p_c)) \\
0 &= p_c(p_c-1)(p_c^2 - 3p_c + 1)
\end{aligned}
\tag{4.8}
$$

Therefore the 4 roots of this equation are:

$$
p_c = 0, 1, \frac{3-\sqrt{5}}{2}, \frac{3+\sqrt{5}}{2}. \tag{4.9}
$$

The last root is greater than 1 and makes no sense as a probability and we can therefore discard it. The first two roots are the trivial roots 0 and 1. It implies that if the corruption propagation probability is 0 then of course there exists the trivial equilibrium that there is no corruption in the population. And if the corruption propagation probability is 1 then there exists the trivial equilibrium that everyone in the population is corrupt.

It is the third root, $p_c = (3-\sqrt{5})/2 = 0.38197$ that is important. It tells us that for any corruption propagation probability less than 38% , corruption does not spread through the population and instead exist in small, isolated pockets (if they don't die out already). On the other hand, for any corruption propagation probability greater than 38%, corruption percolates throughout the population such that almost everyone ends up being corrupt. Not only that, because a large fraction of the population in now corrupt, it is easier for an individual to find a chain of connections (network links) to a potential bureaucrat to bribe.

This suggest (similar to the Granovetter model) that a population might be complacent in thinking that corruption cannot take over their population cause they have such a small propensity to turn corrupt, and such a propensity is way below 50%. But this propensity can slowly shift upward over time, and once it crosses 38% it will be seen that suddenly (in opposition to expectation) that corruption has pervaded throughout the population.

Conclusions

In this chapter we explored three kinds of spatial effect models. Spatial effect models are models where the population are located in space and can observe their near (and sometimes far) neighbors. The key aspect in these models is that individuals might change their behavior (or threshold parameters) by observing their neighbors. And as a population gets more dense, it is easier to observe the change in behavior (and their outcomes) of ones neighbors. But the very act of the very small individual behavior change can change the large scale observed characteristics of the population.

In the Granovetter model every individual had a threshold of corruption and the entire nation's population had a threshold distribution with a given mean and standard deviation. We demonstrated that it is very possible for a sudden transition to happen such that the entire population could shift to fully corrupt if the standard deviation of the population threshold distribution went over a critical sigma (σ^*). A population could have a stable mean but a shifting standard deviation over the years and thus can suddenly become corrupt without the nation realizing such a quick shift could be possible. This also implies that governments can be blindsided into how much corruption has pervaded their society if they fixate on the wrong measure.

The Schelling dynamic segregation model shows that while preferences (for living in an honest/corrupt society) can change smoothly, the emergent equilibrium outcome of the population distribution is not smooth but sudden. If the median honest individual desires to live in a 40% or more honest society, the honest and corrupt population are stably separated from each other and do not interact. However if the median honest individual is satisfied living in a less than 40% honest society; then suddenly the emergent, equilibrium, stable population is 50% corrupt and 50% honest. There is no intermediate stable equilibrium. It is this ability to explain the suddenness of the population composition change that is the strength of the Schelling model, and which also makes it a good complement to the Granovetter model.

The Renormalization group percolation model shows how the average propensity of an individual to corruption can affect its neighbours across a two dimensional lattice. The spatial aspect is key, since we can move around a honest (uncorrupted) individual and still affect the individuals beyond. The model then shows that if the propensity of individuals in a location is below 38% then corruption in the lattice remains local and isolated. But if the propensity (or probability) is above 38% then corruption cascades throughout the system (or nation). This explanation of suddenness of propagation in a useful property of the model and is a good complement to the Granovetter model. In the Renormalization group model we consider the possible shifting change in propensity to corruption while in contrast in the Granovetter model we consider the shifting standard deviation.

CHAPTER 5

NETWORK THEORY MODELS OF CONNECTION

Overview

One aspect of the urban density effect on corruption is that as urban population density increases, the potential number of connections between a briber and bureaucrat increases through the strength of weak ties. But just because a potential connection exists does not mean that an individual can discover it. In this chapter we will use tools and ideas from network (or graph) theory to show that as the number and type of connections between individuals increases, the length of the link between the briber and bureaucrat gets shorter and also the discovery of the particular needed connections (links) gets easier.

Historically network theory started as a subject in mathematics and was called graph theory. Euler started the work in 1736 when he solved the Bridges of Königsberg problem (Biggs et. al., 1986), but the big push happened in 1959 (Newman, 2018) when Paul Erdős and Alfred Rényi began to develop and publish their work on random graphs.

Around the same time in the 1950s, in the field of sociology at MIT, de Sola Pool and Kochen were developing ideas using social networks and trying to quantify the idea of distances between individuals through chain of connections. The ideas went unpublished for a long time but were circulating among many communities and were finally published in 1970s (de Sola Pool & Kochen, 1978).

What however kickstarted the idea of social networks into high gear was when

Stanley Milgram came into the scene in the late 1960s and tried to empirically quantify the typical distance between individual within a social network and showed that it was small (Milgram, 1967; Travers & Milgram, 1969). In his most famous experiment, Milgram chose 296 people randomly in Omaha, Nebraska and Boston, Massachusetts and given them each a package which they were to attempt to send to a specific individual in Boston, Massachusetts. They were given the target individuals name, address, job (stockbroker) and they could only forward the package to someone they knew on a first-name basis who they believed could be closer to the target individual.

64 packages reached their target successfully. Boston starting chains were shorter than Nebraska starting chains and the mean number of intermediaries between the starters and the target were 5.2 (Travers & Milgram, 1969). This is one of the origins of the six degrees of separation myth where everyone in the world is connected through just 6 hops. But what Milgram did show successfully was that it was possible to view the world as small in some way.

Network Terminology

In this section we will introduce some useful terminology that will be used to describe our networks.

Network: A network is simply a collection of points connected by lines. Or more formally a collection of nodes connected by links. It is an abstraction where the complexity of the nodes and their interaction is simplified and instead the underlying structure or topology is focused on. Also called a graph in the literature.

Node: The fundamental unit in a network. In this work it will refer to an individual, which is the unit we analyze with. However as long as we remember that it is a mathematical abstraction we can use it to refer to organizations, or countries, or even pages in the world wide web. Every node has a label $i \in \{1, 2, \ldots, n\}$, where n is the total number of nodes in the network. We will refer to n as the size of the graph. Other terms used for node in the literature are vertex, agent, and site.

Link: A line between two nodes that is used to represent a connection or relationship between the two nodes. For example a friendship or a weak tie or a blood relation between two individuals in that society represented by the network. Links can be directed, in the sense that there is no symmetry or reciprocation in the link. For example node B is son of node A, but the reverse is not true. Undirected links imply symmetry, node B is a friend of node A also implies node A is a friend of node B. In this work we will only work with undirected links and networks. Other terms used for link in the literature are edge, tie, or bond.

Degree: The degree k_i of a node i in the number of links incident to it. So if three other nodes connect to node i its degree will be three. In this work there is only one link possible between any two nodes.

Mean degree: The average degree of the nodes in the network. That is the average number of connections per node (or individual in this work). This is also the first moment of the degree distribution of the network. We are going to use angle brackets $\langle k \rangle$ rather than the expectation operator $\mathbb{E}(k)$ to represent the mean in this work to match the literature The formula to compute the mean degree of the entire network is:

$$\langle k \rangle = \frac{1}{n} \sum_{i=1}^{n} k_i \tag{5.1}$$

Degree distribution: The degree distribution for a given network is the probability p_k that a node selected at random has a degree k. This is important because, as we will see, different network models will have different degree distributions; implying different properties of networks. By definition, since p_k is a probability, it sums to 1. This thus allows us to rewrite the mean degree as:

$$\langle k \rangle = \sum_{k=0}^{\infty} k p_k \tag{5.2}$$

Clustering: This refers to the propensity for a set of three nodes (triples) to be fully connected. Compared to random networks, real world network usually have very high levels of clustering. For example, if node A is a friend of Node B and node B is a friend of Node C, then it is highly likely that node A is a friend of Node C too. Given the existence of clustering, there needs to be measures of it which will be discussed later.

Erdős-Rényi Random Networks

A *random network* is a network in which certain properties are fixed while the others are allowed to be random. The simplest way to define a random network is to fix the number of labeled nodes n and the number of links L that randomly connect the nodes. This is the $G(n, L)$ model and was the approach used by the mathematicians Paul Erdős and Alfred Rényi (Erdős & Rényi, 1960) from the late 1950's to late 1960's. They were the original pioneers of the field and all work in network theory (or graph theory as mathematicians call it) follows in their footsteps.

The other approach is to start with a fixed number of unconnected (isolated) labeled nodes n and then for every pair of nodes in the network we connect them with probability p. This is the $G(n, p)$ model and was introduced by E.N Gilbert (1959). This is the definition of random networks most commonly used in the literature and also what we are going to use in this work. Thus we can define a random network as a network of n nodes where each pair of nodes is connected with probability p.

The steps now to develop the random network are very simple. They are:

i) Begin with n isolated nodes.

ii) Select a pair of nodes and generate a random number between 0 and 1. If this number is greater than p, connect the node pair with a link. Otherwise leave them disconnected.

iii) Repeat the second step for each of the $n(n-1)/2$ node pairs.

Since the size of the network is n, every time we create a link for a node there are $(n-1)$ other nodes that it can be connected to. Therefore the average degree $\langle k \rangle$of a random network is the product of the probability p that two nodes are connected and $(n-1)$ which is the maximum number of links any node can have in a n sized network. Or:

$$\langle k \rangle = p(n-1). \tag{5.3}$$

This implies as p increases the random network becomes denser. Given these facts we can now find the degree distribution of the network. But we have to be careful. A direct analysis will indicate that for a random network the degree distribution will have a binomial distribution. However in a real world, almost all networks are sparse. Or the number of links k that each nodes can have is much smaller than the total number of nodes n available $(k \ll n)$. Because of this the degree distribution of a random network in the real world is a Poisson distribution. Appendix H shows the derivation of this particular distribution and equation 5.4 shows the Poisson degree distribution for the random network.

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \tag{5.4}$$

The major benefit of using the Poisson formulation other than its analytical tractability is that the key network characteristics like the mean and the variance just depends on a single parameter $\langle k \rangle$. The peak of the distribution is at $\langle k \rangle$, by definition the mean number of links an individual has. As p increases, the network becomes denser and thus increases $\langle k \rangle$ and moves the center of the distribution to the right.

The width of the distribution (or its dispersion) of this Poisson distribution is $\sigma_k = \langle k \rangle^{1/2}$. This implies that as the network gets denser, the wider the degree distribution. Hence we will observe more and more nodes with an extreme number of links. However we have to be careful with the word extreme, it does not mean huge. For example if the network has a mean degree distribution of 9 $(\langle k \rangle = 9)$, then 95% of the nodes will have between 3 links to 15 links $(\langle k \rangle \pm 2\sigma_k = 9 \pm 2(\sqrt{9}) = 9 \pm 6)$.

The other key thing to observe is that this Poisson distribution does not depend on the number of nodes n. Therefore different Poisson random networks of different sizes but the same mean degree $\langle k \rangle$ will have the same properties, and thus can be considered indistinguishable from each other. This is important, when we try to compare nations of different population sizes.

Deriving the Diameter of a Random Network

The diameter of a network is the 'longest shortest path' in the network. There is always a shortest path between any two nodes, and this path can be made longer by backtracking and taking alternative longer routes. By 'longest shortest path' what we mean is that among all the shortest paths that exists between all the nodes in the network, the diameter is the longest. The diameter tells us the upper bound of the number of hops (of intermediate connections or friends or weak ties) that one has to take so that two randomly selected individuals in the network are connected.

This measure is important in this research because it gives us the upper bound on the connection length in the network between the briber and the bribed. Why do we choose the upper bound? Because by showing the upper bound is a small number, we make it possible to argue that even in a long path scenario, a briber and the bribed can be connected. Also the mean path distance is difficult to calculate for most networks and it does not communicate much information about the network.

The diameter of a random network is easy to calculate (Newman, 2018; Barabasi & Posfai, 2016). Let us consider a random network of average degree $\langle k\rangle$. Therefore on average there are:

$\langle k\rangle$ nodes at a distance of one $(d = 1)$,

$\langle k\rangle^2$ nodes at a distance of two $(d = 2)$,

$\langle k\rangle^3$ nodes at a distance of three $(d = 3)$.

...

$\langle k\rangle^d$ nodes at distance d.

Therefore we can calculate the expected number of nodes that might exist in a network at a distance d from an arbitrary starting node (including the starting node). It can be calculated (using the sum of a geometric series) as

$$N(d) = 1 + \langle k\rangle^1 + \langle k\rangle^2 + \ldots + \langle k\rangle^d = \frac{\langle k\rangle^{d+1} - 1}{\langle k\rangle - 1} \tag{5.5}$$

However $N(d)$ cannot be more than n, which is the total number of nodes in the network. Therefore

$$N(d_{max}) = n. \tag{5.6}$$

Or

$$\frac{\langle k\rangle^{d_{max}+1} - 1}{\langle k\rangle - 1} = n. \tag{5.7}$$

Now for most random networks we can safely assume $\langle k \rangle \gg 1$, so we can neglect the -1 terms, thus

$$\frac{\langle k \rangle^{d_{max}+1}}{\langle k \rangle} \approx n, \tag{5.8}$$

or

$$\langle k \rangle^{d_{max}} \approx n. \tag{5.9}$$

Taking logs on both sides gives

$$d_{max} ln \langle k \rangle \approx ln(n). \tag{5.10}$$

Therefore the diameter of a random network is

$$d_{max} \approx \frac{ln(n)}{ln \langle k \rangle}. \tag{5.11}$$

Thus, we see that the diameter of a random network (d_{max}) scales with the natural logarithm of the size of the network (n). Albert and Barabasi (2002) heuristically argued that this calculated value of d_{max} is actually a better approximation for the average distance between two randomly chosen nodes $\langle d \rangle$, which by definition is smaller than d_{max}.

Let us say the whole world has 7.5 billion individuals, this is our n. And the average individual knows about a 100 other people, this is our $\langle k \rangle$. Therefore

$$d_{max} \approx \frac{ln(n)}{ln \langle k \rangle} = \frac{ln(7.5 \times 10^9)}{ln\ 100} = 4.94 \tag{5.12}$$

Even if we consider that 100 is too large a number for the number of acquaintances, we can change $\langle k \rangle$ to just 10, and then $d_{max} = 9.86$, it just doubles. An order of magnitude change in mean connections only halves the average distance between any two randomly selected individuals. The logarithm dominates the equation.

Thus the popular idea that the whole world is separated by 6 degrees of separation. The upper bound in the 'longest shortest distance' measure suggests it. This of course does not take care of clustering issues and other density issues.

Also we are working with the huge population of the world. A more realistic upper bound assumption would be a city of 20 million population (approximately the New York metropolitan area) if we wish to be aggressive in selecting for cities with large population. Even if the average individual know at most 20 others, the

$$d_{max} \approx \frac{ln(n)}{ln\langle k \rangle} = \frac{ln(20 \times 10^6)}{ln\, 20} = 5.6. \tag{5.13}$$

So at most 6 links can possibly separate a briber and the potential bribed in a huge city. For a more median sized city like the Kansas City metropolitan area of approximately 2.2 million people and a median acquaintance rate of 40, the number of links required to traverse the network drops to 4.

Thus between 4 to 6 links potentially separate a briber and the bribed. And all this comes from the assumption of pure randomness of the network linkages which is not necessarily a good assumption.

Watts-Strogatz Small-World Networks

While the Erdős-Rényi random network can be useful and provide some insight, its greatest shortcoming is that the connections between nodes are random and have no structure. In the real world networks are neither entirely random nor are they completely ordered. Because network can fall somewhere in the spectrum of order and randomness, a different way to think about it might be necessary.

In 1998, Duncan Watts and Steven Strogatz in their *Nature* paper (Watts & Strogatz, 1998) introduced the small-world network, which incorporated both structure and randomness in their network. And it was done in a very simple way through the ability to tune a single parameter that allows the network to continuously interpolate between ordered to random topologies. This network model has become one of the most influential in the field and has inspired a lot of research on it.

The steps to develop the network are very simple. They are:

i) Arrange the n nodes in a ring.

ii) Join each node to its $2k$ nearest neighbours (k nearest neighbours clockwise and k nearest neighbours anti-clockwise) within a given range such that all the nodes have the same degree. For example in top network in Figure 5.1, every node is connected to its four nearest neighbours $(k = 2)$, thus giving all the nodes the same degree of 4.

iii) Rewire a fraction p of the links. This is achieved by randomly selection a node from the whole network, drop one of its links randomly and then creating a new link with a new node randomly chosen from the entire network. For example in the middle network in Figure 5.1, we can observe a few of the links has been disconnected from their near nodes and randomly attached to farther nodes. The rewiring fraction in this particular example network is $p = 0.1$.

That is all that is required to create a small-world network. The value that the parameter p takes totally determines the complexity of the resulting network. In the first step, where $p = 0$, the network is absolutely ordered, such that when one node (individual) wishes to communicate to a farther node (individual), it has to communicate with all the intermediate nodes. The ordered network has a long path length to traverse the

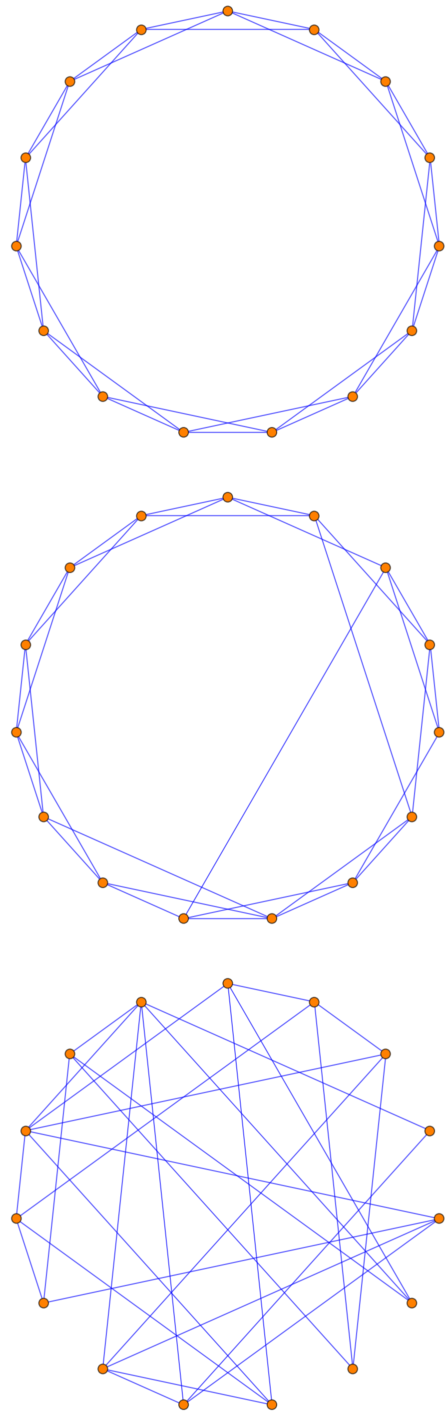

Figure 5.1: The development of the Watts-Strogatz small-world network model. All the three networks have 15 nodes. The top network has all its nodes connected to its 4 nearest neighbours. The middle network has a few of its links randomly rewired at $p = 0.1$ thus creating a small-world network. The bottom network is rewired at $p = 1$ degenerating into an Erdős-Rényi random network.

network. In the corruption model parlance, an individual has to find many successful intermediate ties between themselves and the corrupt bureaucrat for a successful connection to occur.

But in the second network, where $p = 0.1$, what happens is that the rewiring creates 'shortcuts' in the network. A few nodes (individuals) instead of having a link to their nearest neighbor, have a link to a node much farther away. This link ends up being used as a shortcut to jump across the network. Now instead of having to communicate with many intermediate nodes, one can jump across the network and get to the desired destination node with fewer hops. The path length to traverse this same sized network has reduced. The world has become 'small' in a way.

There is however a caveat. If too many rewiring of links takes place, the order in the network breaks down and it degrades to being a Erdős-Rényi random network. In the third network, where $p = 1$, most of the nodes are linked to other far nodes and very few to nearest neighbours. So while the path length to traverse the whole network is still small, it might be the case that to connect to an immediate neighbor way too many hops might have to be takes. In other words, the global path length (for network traversal) is low but the local path length has become high.

To understand this critical 'shrinking distance' features of small-world networks, Watts & Strogatz (1998) had to develop some way to measure these features. They created two measures of clustering coefficients.

Local clustering coefficient: This measures clustering at the local level. They argued that suppose a node i has k_i neighbours. Then at most $k_i(k_i - 1)/2$ links can exist between them (this will occur when every neighbor of node i in connected to every other neighbour of i). Therefore the local clustering coefficient is defined as the total number of links (L_i) that do exist to node i divided by the maximum number of possible links.

Or:

$$C_i = \frac{2L_i}{k_i(k_i - 1)}. \tag{5.14}$$

Note that C_i has a value between 0 and 1. If $C_i = 0$, then none of the neighbours of node i link to each other. If $C_i = 1$, then all of the neighbours of node i connect to each other. They form a complete network or clique. And if $C_i = 0.5$ then there is a 50% chance that two neighbours of a node are linked to each other. Another way to say it is that when $C_i = 0.5$, half of your total group of friends are friends with each other.

Thus C_i measures the local link density of the network. The more densely interconnected the local neighborhood, the higher the local clustering coefficient

Mean clustering coefficient: This measures the degree of clustering of the whole network. The formula for a network of n nodes is:

$$\langle C \rangle = \frac{1}{n} \sum_{i=1}^{n} C_i \tag{5.15}$$

$\langle C \rangle$ can be interpreted as the average cliquishness of the entire population that we are interested in. It can also be said the as a country gets more denser in population, $\langle C \rangle$ also increases in value.

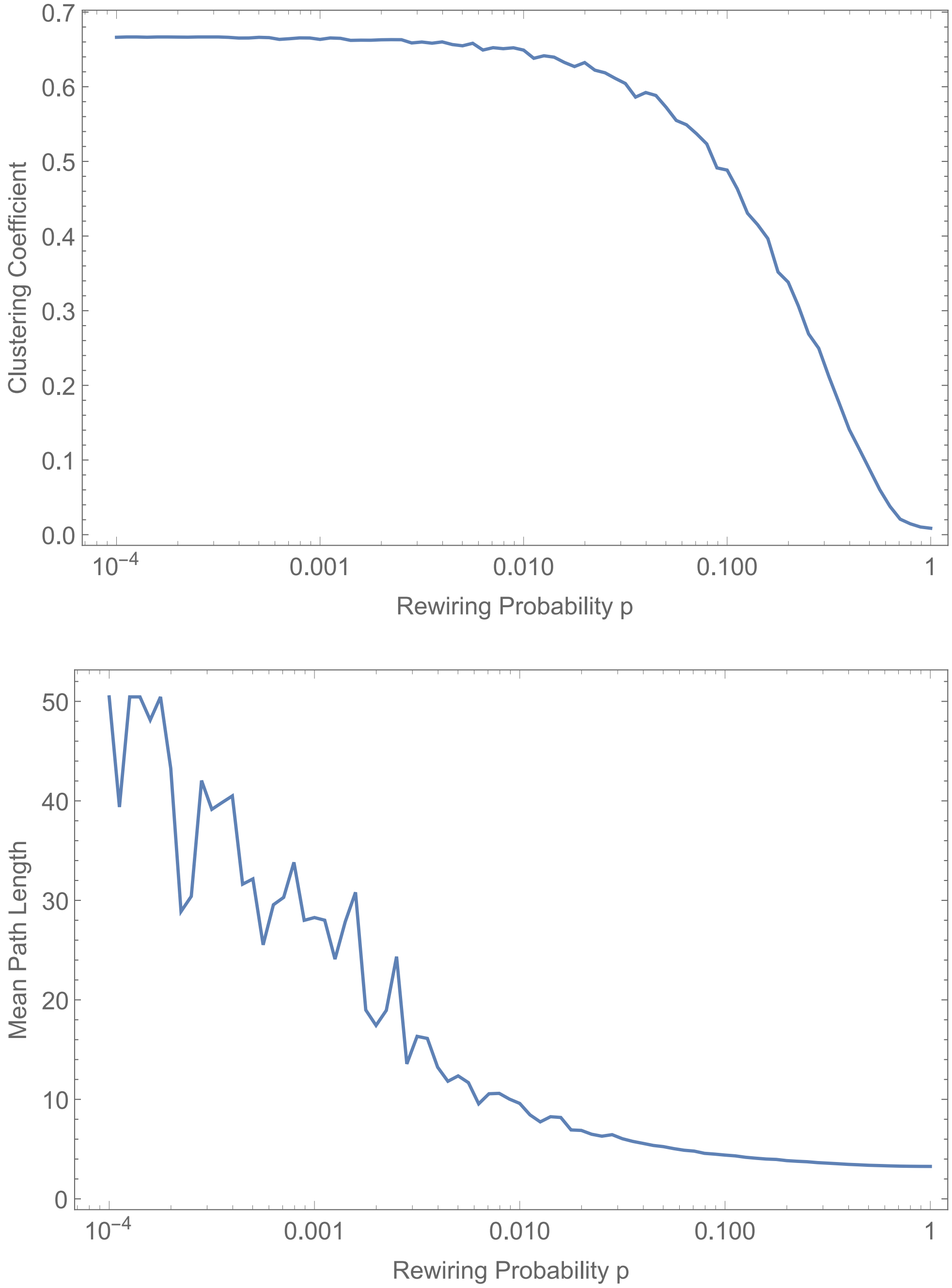

Figure 5.2: Plot of Mean Clustering Coefficient against Rewiring Probability p (top) and Plot of Mean Path Length against Rewiring Probability p (bottom). Network is 1000 nodes with starting configuration of links to 10 nearest neighbours. Note: x-axis is a logarithmic scale and the same on both graphs.

Using this, Watts and Strogatz then ran simulations (analytical solutions are intractable) to show how the mean path length (the average length of all the shortest paths between nodes in the network) varied as rewiring probability changed. This is more clearly seen in the next figure which we simulated in *Mathematica*.

In Figure 5.2, the top graph shows how the clustering coefficient decreases as the rewiring probability increases. The network starts with a 1000 nodes, each connected to its 10 nearest neighbours. At rewiring probability of 0.0001 the clustering coefficient starts at $\approx$ 0.65, implying that for every clique 65% of the friends are friends with each other. The bottom graph shows that at that rewiring probability of 0.0001 the mean path length (the average length of all the shortest paths in the network) is 50. Or 50 hops to traverse the network.

As the rewiring probability increases all the way to $p = 0.01$ (or 1 in every 100 nodes/individuals have a link that is not connected to their nearest neighbor but to some other random further node), the clustering coefficient is still stable at $\approx$ 0.65, but the mean path length has massively decreased to $\approx$ 8. This is a decrease of almost a factor of 6.5. By just introducing a very small perturbation we have changed the network distance characteristic massively.

What does this imply for corruption? The Erdos-Renyi random network model says that almost all networks are small, there are very few steps between bribers and bribed. The dilemma is that it is hard for an individual to find the appropriate short path among all the random links since there is no structure to follow. What Watts-Strogatz did was show that even in very structured networks, the world is also very small (very few steps between bribers and bribed in our parlance). And all it takes for a structured network to become small is a very small fraction of randomly rewired links (approximately 1% in our example and it is very close to a typical small world random network).

But it is the structure of the Watts-Strogatz network that is very important and useful. It lets the individual find the appropriate short path easily. All one has to do is ask around (seek) in the cliques that one is part of and very soon the weak ties (or strong ties) that has a far off link that one requires will be found.

The entire point of this analysis is to show that very structured network with a very small randomness component is sufficient to make the world a small-world. This thus enables us to argue that as population density increases, the probability of cliques having the random component increases. It is more likely in dense population for people to have a connection to someone in a far off clique and these kinds of connections increases as density of interaction increases.

Barabasi-Albert Network Model

Even though the Watts-Strogatz network model gave us a lot of realistic and useful features, some things are still missing. One is that the number of nodes n are fixed (this is also true of the random network model). There was no possibility of network growth, whereas in the real world networks (and populations) are always growing. The other problem is that there are no such things as hubs or centralized nodes or popular nodes. In the real world there are individuals and organizations which are more connected to others (have way more links to other nodes than average). Albert and Barabasi (2002) argued that the existence of such hubs can potentially change the properties and dynamics of a network that possess them.

The steps to develop the network are:

i) We start with m_0 nodes with arbitrary links randomly placed between them as long as each node has at least one link.

ii) This is the growth step. At each step in time we add one new node with

m ($\leq m_0$) links that connects the new node to m different nodes that already exist in the network.

iii) This is the preferential attachment step. The probability $\Pi(k)$ that the new node connects to node i is dependent on the degree k_i of node i and the sum of the degrees over all the pre-existing nodes j:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \tag{5.16}$$

The preferential attachment is a probabilistic mechanism. While a new node is free to connect to any other node in the network, this equation implies that if a new node has a choice to link between a degree-two or degree-four node, it is twice as likely to connect to the degree-four node. This is a reflection of the real world. People who have lots of friends to begin with also end up making up more new friends than average. Or even more appropriately, if an individual wishes to connect to a network they are more likely on average to befriend someone who has lots of connections/friends in the first place.

Therefore heavily linked nodes or 'hubs' will arise that tend to quickly accumulate even more links while nodes that have only a few links will most likely stay lowly linked. Thus we can say that new nodes have a 'preference' for attaching themselves to heavily linked nodes. After t time-steps the network generated will have $t+m_0$ nodes and m_0+mt links.

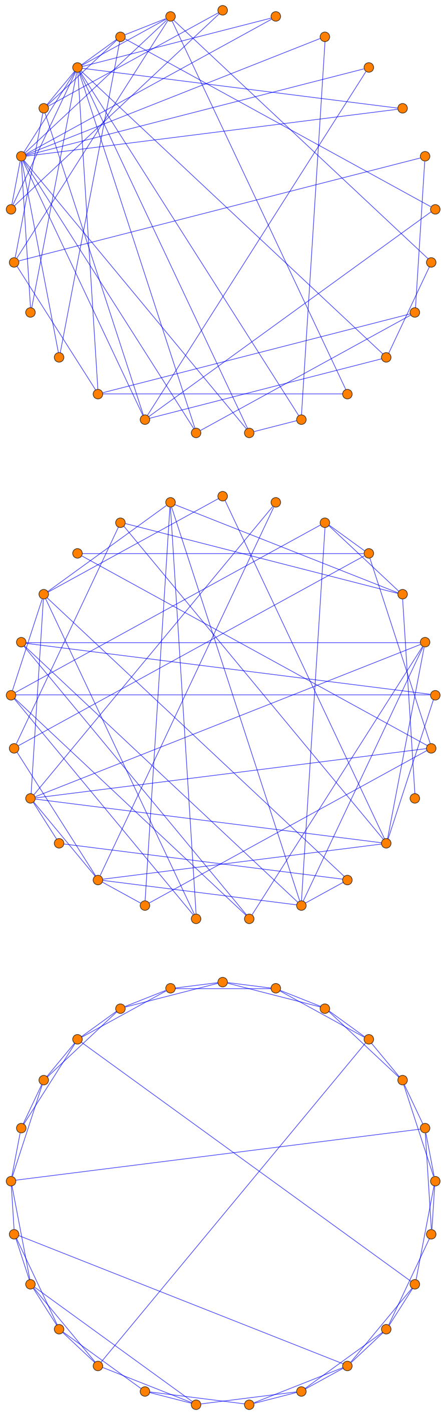

Figure 5.3: The Barabasi-Albert Network (top) compared to the Erdős-Rényi Random Network (middle) and the Watts-Strogatz Small-Worlds Network (bottom). All the networks have 25 nodes. One can observe the hub nodes with more than average links in the top-left quadrant of the Barabasi-Albert network.

In Figure 5.3, the top network is a Barabasi-Albert network and we can observe that the upper-left quadrant has nodes with more than average links. These would be the ‘hubs’ of the network population. In comparison the middle Erdős-Rényi random network has no preferential nodes and it is hard to discern any structure at all. In the bottom is the Watts-Strogatz small world network and we can see how much structure it comparatively has.

The other aspect of the Barabasi-Albert network is that its diameter (the maximum number of hops to traverse the network) grows slower than a random network (Bollobas & Riordan, 2004; Cohen & Havlin, 2003). A random network grows $d_{max} \sim ln(n)$ whereas the Barabasi-Albert network grows:

$$d_{max} \sim \frac{ln(n)}{ln\ ln(n)}. \tag{5.17}$$

Where n is the total population. So when the population n is more than $10,000$ individuals/nodes the Barabasi-Albert network has the smaller diameter.

Thus for most populations the Barabasi-Albert network model shows that hubs end up dominating. This gives a network that has a much smaller size required to traverse to link the briber and the bribed. And also more importantly it gives the network more structure so that the briber can find the appropriate person to link to more easily. One just has to link to an individual that behaves as a hub, and since they are a hub more people know about them and thus finding them is much easier.

Conclusions

In this chapter we explored three kinds of network connection models. These models tell us how the structure of interconnection between individuals has an effect on how far, on average, everyone is maximally connected to each other and how easy it might be to find a desired chain of links within the network. The key aspect in these models is how very simple rules (how links are established) in developing the network leads to very far reaching consequences in the ability of arbitrary nodes to connect across the network.

The Erdős-Rényi random network model shows that than even in a totally randomly connected world, the maximum traversed distance between a potential briber and bribed falls between 4 and 6 links. This upper bound requires no structure at all in the network and tells us that the traversed value can only decrease if structure can be introduced.

The Watts-Strogatz model shows that unlike Erdős-Rényi, we do not need to a lot of individuals to have far off links. In fact the usual structures that exist in normal human populations are actually useful. Most people belong to groups (or cliques) and it is the property of cliques that most of the individuals know each other (the friends of my friend are also usually my friends). Given this clique structure across a network population, only a small addition of randomness in linking to non-clique members leads to a massive change in the traversal structure of the network.

Watts & Strogatz shows with numerical simulation that a 1% rewiring of links is sufficient to shirk the traversal length of the network by a factor of ~ 6. A Watts-Strogatz small world network has approximate the same diameter (total number of hops to traverse the network) as a random network but it is so structured in intra-clique connections that finding an appropriate link to connect the briber across the network to a bureaucrat becomes very easy.

In contrast the Barabasi-Albert preferential attachment network shows is that even if we assume no structure like that of the Watts-Strogatz model and we revert to random linking model, a very conservative assumption that hubs (people who have more friends or connections than average) exist leads to network with a very small diameter. And in fact the diameter is smaller than Erdős-Rényi models once we are above a population of $\sim$ 10k. The usefulness of hubs is that by definition more than an average number of people know them and they are usually the first ones newcomers prefer to connect to. This leads to hubs being the easiest was to find the appropriate link between briber and bribed.

The real world is small and it does not need much structure to make it so. This enables corruption to propagate much easily than one would naively expect. These three modes gives us a useful bound in how hard it would be to find a link between a briber and the bribed while at the same time telling us how simple the population network structure needs to be to make the world small and discoverable.

CHAPTER 6

CONCLUSIONS

Effect of Adding Urban Population Density to Corruption Models

In this dissertation we have argued that attempts to explain economic corruption need to take into account urban population density in their model specifications. It has quite a large effect in the poor and corrupt countries and ignoring it will be a major mistake.

One of the reasons urban population density is not used is that it is hard to calculate. Most nations while they know their urban population, they are not precisely sure of their urban extent. We use data from night-time lights to delineate urban extent area and calculate urban density.

Our regressions show that an increase in urban population density by 100 people per square kilometer has the same effects on corruption as decreasing the per capita GDP by $172 (in 2005 dollars, $243 in 2021 dollars). This is a substantial effect and has never been considered before when develop anti-corruption programs.

Other major variables of interest we show are Natural Resources rents (for every 10% of its contribution to total GDP, corruption increases by 0.317) and remittances (for every 10% of its contribution to total GDP, corruption increases by 0.21).

Another major idea we explored is the effect on corruption due to the colonial status of countries. We show that countries that were Spanish colonies have their corruption perception index higher by 0.608 compared to British colonies (95% confidence interval of the difference is (0.379, 0.837)). Also internal colonies (Austro-Hungarian, Ottoman and former Soviet Union states) have a baseline corruption level much higher (> 0.9) than non-internal colonies.

We also show many countries that are outliers in their corruption perceptions than what the model predicts. Chile and Uruguay are more than 2 standard deviation less corrupt than the model suggest they should be. New Zealand, Estonia, Barbados, Dominica, Oman, Bhutan, Ghana, and Mauritania are all more than 1.5 standard deviation less corrupt. This gives us a list of countries to study further and explore their institutional structure as to why they are less corrupt than usual.

On the flip side we also find countries that are more corrupt that the model predicts. Five countries are more than 2 standard deviations more corrupt that the model suggests they should be - Equatorial Guinea, Belarus, Italy, United States, and Norway. Thus we have a list of countries whose institutional structure and historical evolution is worth exploring further

The Economic Effect in Corruption Models

We have attempted, in this dissertation, to show the effect of corruption on economic variables that have an effect on the populace. We focused on variables that have an effect both in terms of economic growth and development of human potential. We also developed the models by incorporating urban density to control for the effect of that variable, so that we can see the effect of corruption much more clearly.

We studied four new variables: tax revenue, foreign direct investment (FDI), total government spending, and government spending on education. All the variables were measured in % of GDP.

The tax revenue decreased by 1.58% points for every 1 unit increase in CPI (Corruption Perceptions Index) and is a major effect statistically. Remittances had a large effect. A 1% increases in remittances increased tax revenue by 0.25%. Inequality also had a large effect. A 10 point change in the Gini coefficient decreased the tax revenue collected by 1.59% points. Long term colonial effects also exist. Spanish colonies were observed to collected less tax revenue (7.76% points less than British colonies and very statistically significant).

Foreign direct investment (FDI) decreased by 0.549% points for every 1 unit increase in CPI which is a decent effect. Natural resource rents and remittances have a large effect. FDI inflow increases by 0.613% points for every 10% point increase in natural resources rents. And FDI inflow increases by 1.296% points for every 10% point increase in remittances. Colonial effects are major. Internal colonies on average have an 8% FDI inflow compared to an average 6% FDI inflow for external colonies. Also, among the major European powers, French colonies have the largest FDI inflow.

Corruption has a large effect on total government spending. A 1 unit increase in CPI has a 1.452% point decrease in government spending. Since in this data set the mean government spending is 15.5%, a unit change in CPI has a nearly 10% decrease in government spending which is a substantial effect. Urban density has a statistical significant effect at the 1% level. For every increase of 1000 people per square kilometer, government spending goes down by 0.216 percentage points. This suggests more efficiency in spending due to density effects. GDP per capital has a small but substantial effect (at the 1% level). For every $1000 increase in GDP per capita, government spending decreases by 0.088% points.

Inequality has the other large effect. A 10 point change in the Gini coefficient increases government spending by 1.2% points. Which makes sense, since a country with a high level of inequality will require more state services to provide for the more plentiful poor. The only major colonial effect noticed is that former Spanish colonies have much lower government spending (approximately 12% of GDP) compared to the higher average of the other European colonies (greater than 16.5% of GDP).

Finally, corruption decreases government spending on education. A one unit increase in CPI has a 0.291% point decrease in education spending. Given that the average state spending on education among all countries in the dataset is 4.59%, this is quite a decent effect. A one unit increase in CPI has a 6.3% decreases in state educational spending.

The only other variable that has a major effect here is remittances. A 10% point increase in remittances increases state educations spending by 0.51% points. The major colonial effect observed is that former Spanish colonies spend less government money (4.46% of GDP) on education that other former European colonies (greater than 5.7% of GDP).

Corruption and Spatial Effect Models

In this dissertation we explored three kinds of spatial effect models. Spatial effect models are models where the population are located in space and can observe their near (and sometimes far) neighbors. The key aspect in these models is that individuals might change their behavior (or threshold parameters) by observing their neighbours. And as a population gets more dense, it is easier to observe the change in behavior (and their outcomes) of ones neighbours. But the very act of the very small individual behavior change can change the large scale observed characteristics of the population.

In the Granovetter model every individual had a threshold of corruption and the entire nations population had a threshold distribution with a given mean and standard deviation. We demonstrated that it is very possible for a sudden transition to happen such that the entire population could shift to fully corrupt if the standard deviation of the population threshold distribution went over a critical sigma (σ^*). A population could have a stable mean but a shifting standard deviation over the years and thus can suddenly become corrupt without the nation realizing such a quick shift could be possible. This also implies that governments can be blindsided into how much corruption has pervaded their society if they fixate on the wrong measure.

The Schelling dynamic segregation model shows that while preferences (for living in an honest/corrupt society) can change smoothly, the emergent equilibrium outcome of the population distribution is not smooth but sudden. If the median honest individual desires to live in a 40% or more honest society, the honest and corrupt population are stably separated from each other and do not interact. However if the median honest individual is satisfied living in a less than 40% honest society; then suddenly the emergent, equilibrium, stable population is 50% corrupt and 50% honest. There is no intermediate stable equilibrium. It is this ability to explain the suddenness of the population composition change that is the strength of the Schelling model, and which also makes it a good complement to the Granovetter model.

The Renormalization group percolation model shows how the average propensity of an individual to corruption can affect its neighbours across a two dimensional lattice. The spatial aspect is key, since we can move around an honest (uncorrupted) individual and still affect the individuals beyond. The model then shows that if the propensity of individuals in a location is below 38% then corruption throughout the lattice remains local and isolated. But if the propensity (or probability) is above 38% then corruption cascades throughout the system (or nation). This explanation of suddenness of propagation in a useful property of the model and is a good complement to the Granovetter model. In the Renormalization group model we consider the possible shifting change in propensity to corruption while in contrast to the Granovetter model where we consider the shifting standard deviation.

Corruption and Network Theory Models

In this dissertation we explored three kinds of network connection models. These models tell us how the structure of interconnection between individuals have an effect on how far, on average, everyone is maximally connected to each other and how easy it might be to find a desired chain of links within the network (like between the briber and the bribed). The key aspect in these models is how very simple rules (how links are established) in developing the network leads to very far reaching consequences in the ability of arbitrary nodes to connect across the network.

The Erdős-Rényi random network model shows that than even in a totally randomly connected world, the maximum traversed distance between a potential briber and bribed falls between 4 and 6 links. This upper bound requires no structure at all in the network and tells us that the traversed value can only decrease if structure can be introduced.

The Watts-Strogatz model shows that unlike Erdős-Rényi, we do not need to a lot of individuals to have far off links. In fact the usual structures that exist in normal human populations are actually useful. Most people belong to groups (or cliques) and it is the property of cliques that most of the individuals know each other (the friends of my friend are also usually my friends). Given this clique structure across a network population, only a small addition of randomness in linking to non-clique members leads to a massive change in the traversal structure of the network.

Watts & Strogatz shows with numerical simulation that a 1% rewiring of links is sufficient to shirk the traversal length of the network by a factor of ~ 6 . A Watts-Strogatz small world network has approximate the same diameter (total number of hops to traverse the network) as a random network but it is so structured in intra-clique connections that finding an appropriate link to connect the briber across the network to a bureaucrat becomes very easy.

In contrast the Barabasi-Albert preferential attachment network shows is that even if we assume no structure like that of the Watts-Strogatz model and we revert to random linking model, a very conservative assumption that hubs (people who have more friends or connections than average) exist leads to network with a very small diameter. And in fact the diameter is smaller than Erdős-Rényi models once we are above a population of $\sim$ 10k. The usefulness of hubs is that by definition more than an average number of people know them and they are usually the first ones newcomers prefer to connect to. This leads to hubs being the easiest was to find the appropriate link between briber and bribed.

The real world is small and it does not need much structure to make it so. This enables corruption to propagate much easily than one would naively expect. These three modes gives us a useful bound in how hard it would be to find a link between a briber and the bribed while at the same time telling us how simple the population network structure needs to be to make the world small and discoverable.

Further Work

We wish to continue our work in corruption and especially try to adapt it to urban extent models that are not dependent on night-time lights. Furthermore there is newer data on corruption perceptions (both in new data in time and new data from different kind of surveys) that we believe we can incorporate into our models to make them better.

One of the reasons we attempted this dissertation is to bring novel mathematical ideas into heterodox political economy. Neoclassical economics has sort of dominated by adopting every mathematical idea it can take and heterodox political economy has been left behind in a way. If we can take a lead on ideas that are still not been thoroughly coopted, maybe we can usefully apply a heterodox perspective upon them.

We believe that the ideas explored here from spatial effect models and network theory models have the potential to be adapted usefully in many different parts of heterodox political economy. This is why in this dissertation we take the first steps with it by exploring it with corruption data.

We believe there are many opportunities for further work using these models (and their more advanced counterparts) and adapting them to ideas in political growth theory. Also another avenue is in exploring how novel economic ideas are adopted and especially novel ideas regarding monetary models of the economy.

APPENDIX A

LIST OF COUNTRIES IN THE DATA SET

The data set consists of 148 countries that are sorted below by continent and then by their Gleditsch-Ward country code.

North America: United States, Canada, Haiti, Dominican Republic, Jamaica, Barbados, Dominica, Grenada, Mexico, Belize, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, Panama.

South America: Colombia, Venezuela, Guyana, Suriname, Ecuador, Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay.

Europe: United Kingdom, Ireland, Netherlands, Belgium, France, Switzerland, Spain, Portugal, Germany, Poland, Austria, Hungary, Czech Republic, Slovak Republic, Italy, Malta, Albania, Macedonia, Croatia, Bosnia and Herzegovina, Slovenia, Greece, Cyprus, Bulgaria, Moldova, Russia, Estonia, Latvia, Lithuania, Ukraine, Belarus, Finland, Sweden, Norway, Denmark, Iceland.

Africa: Cape Verde, Equatorial Guinea, Gambia, Mali, Senegal, Benin, Mauritania, Niger, Ivory Coast, Guinea, Burkina Faso, Ghana, Togo, Cameroon, Nigeria, Gabon, Central African Republic, Chad, Republic of Congo, Uganda, Kenya, Tanzania, Burundi, Rwanda, Djibouti, Ethiopia, Eritrea, Angola, Mozambique, Zambia, Malawi, South Africa, Namibia, Lesotho, Botswana, Swaziland, Mauritius, Morocco, Algeria, Tunisia, Libya, Sudan, Egypt.

Asia: Armenia, Georgia, Azerbaijan, Iran, Turkey, Syria, Lebanon, Jordan, Israel, Saudi Arabia, Yemen, Kuwait, Bahrain, Qatar, United Arab Emirates, Oman, Turkmenistan, Tajikistan, Kyrgyz Republic, Uzbekistan, Kazakhstan, China, Mongolia, South Korea, Japan, India, Bhutan, Pakistan, Bangladesh, Myanmar, Sri Lanka, Nepal, Thailand, Laos, Vietnam, Malaysia, Philippines, Indonesia.

Australasia: Australia, Papua New Guinea, New Zealand.

APPENDIX B

LIST OF COUNTRIES ACCORDING TO THEIR COLONIAL STATUS

Countries never colonized (22):

United Kingdom, Netherlands, France, Switzerland, Spain, Portugal, Germany, Austria, Italy, Russia, Sweden, Denmark,

Ethiopia, Iran, Turkey, Saudi Arabia, China, Mongolia, Japan, Bhutan, Nepal, Thailand.

Countries colonized by Britain (44):

United States, Canada, Jamaica, Barbados, Dominica, Grenada, Belize, Guyana,

Ireland, Malta, Cyprus,

Gambia, Ghana, Nigeria, Uganda, Kenya, Tanzania, Zambia, Malawi, South Africa, Namibia, Lesotho, Botswana, Swaziland, Mauritius, Sudan, Egypt

Jordan, Israel, Yemen, Kuwait, Bahrain, Qatar, United Arab Emirates, Oman, India, Pakistan, Bangladesh, Myanmar, Sri Lanka, Malaysia

Australia, New Zealand, Papua New Guinea.

Countries colonized by France (24):

Haiti, Mali, Senegal, Benin, Mauritania, Niger, Ivory Coast, Guinea, Burkina Faso, Togo, Cameroon, Gabon, Central African Republic, Chad, Republic of Congo, Djibouti, Morocco, Algeria, Tunisia,

Syria, Lebanon, Cambodia, Laos, Vietnam.

Countries colonized by Spain (18):

Dominican Republic, Mexico, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Ecuador, Peru, Bolivia, Paraguay, Chile, Argentina, Uruguay, Equatorial Guinea.

Countries colonized by Portugal (4):

Brazil, Cape Verde, Angola, Mozambique

Countries colonized by Netherlands (3):

Surinam, Belgium, Indonesia

Countries colonized by Others (9):

Iceland (Denmark), Norway (Sweden), Finland (Russia pre Soviet (pre 1917)), Burundi (Belgium), Rwanda (Belgium), Eritrea (Ethiopia), Libya (Italy), South Korea (Japan), Philippines (United States).

Internal colony of the Soviet Union (14):

Moldova, Estonia, Latvia, Lithuania, Ukraine, Belarus, Armenia, Georgia, Azerbaijan, Turkmenistan, Tajikistan, Kyrgyz Republic, Uzbekistan, Kazakhstan.

Internal colony of the Austro-Hungarian Empire (6):

Hungary, Czech Republic, Slovak Republic, Croatia, Bosnia and Herzegovina, Slovenia.

Internal colony of the Ottoman Empire(4):

Albania, Macedonia, Greece, Bulgaria.

APPENDIX C

GRAPHS

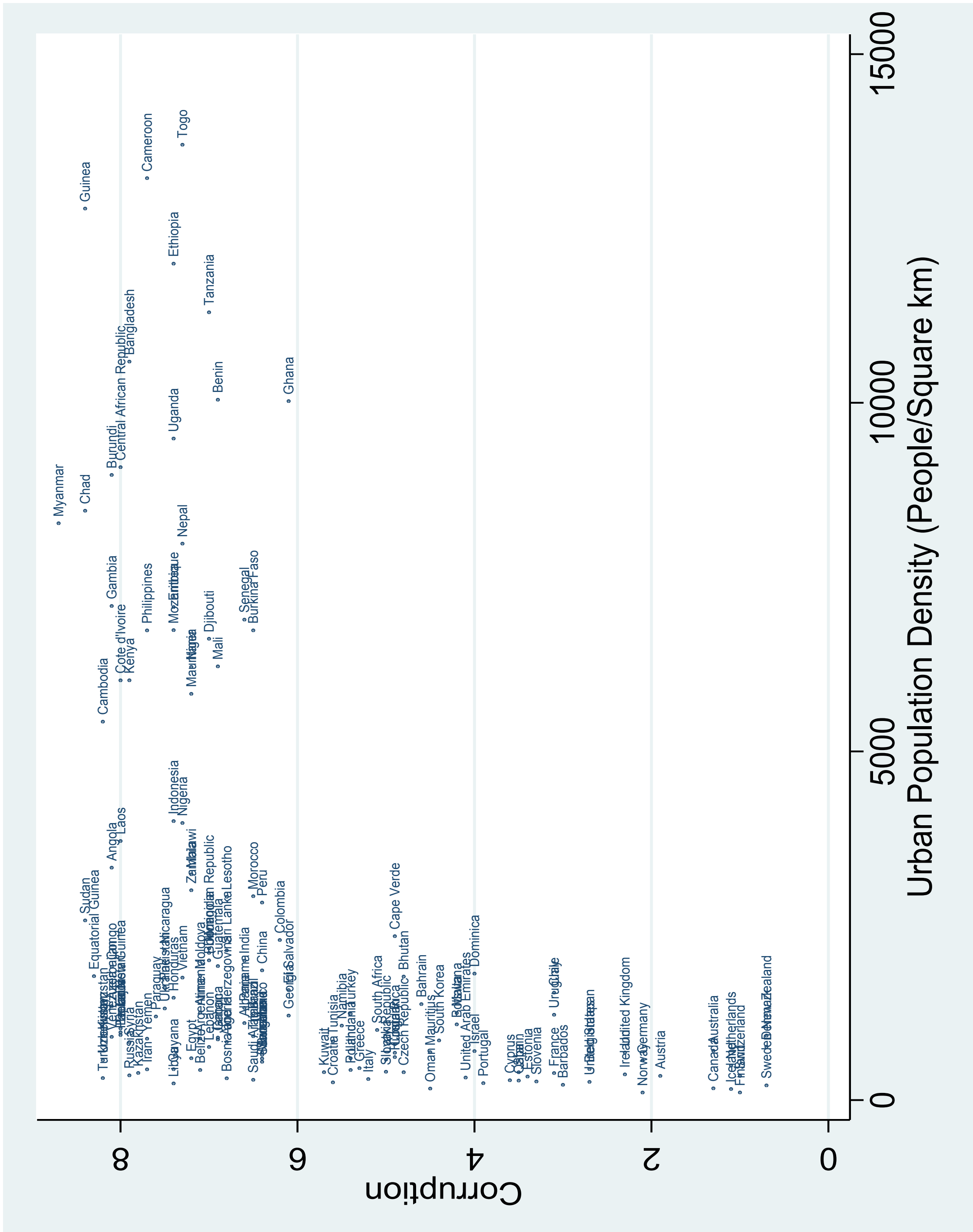

Figure C.1: Plot of Corruption against Urban Population Density (2008)

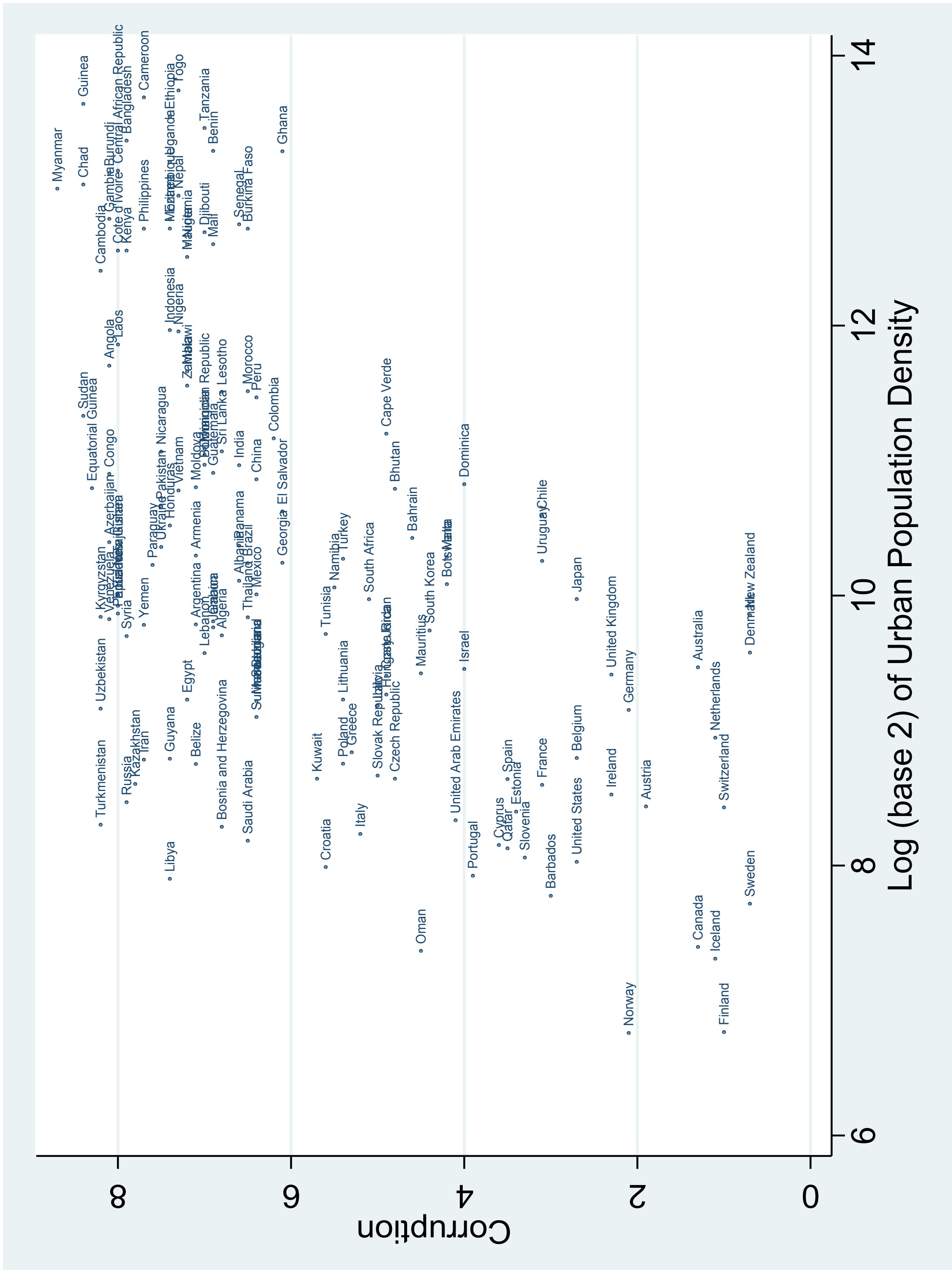

Figure C.2: Plot of Corruption against Log (base 2) of Urban Population Density (2008)

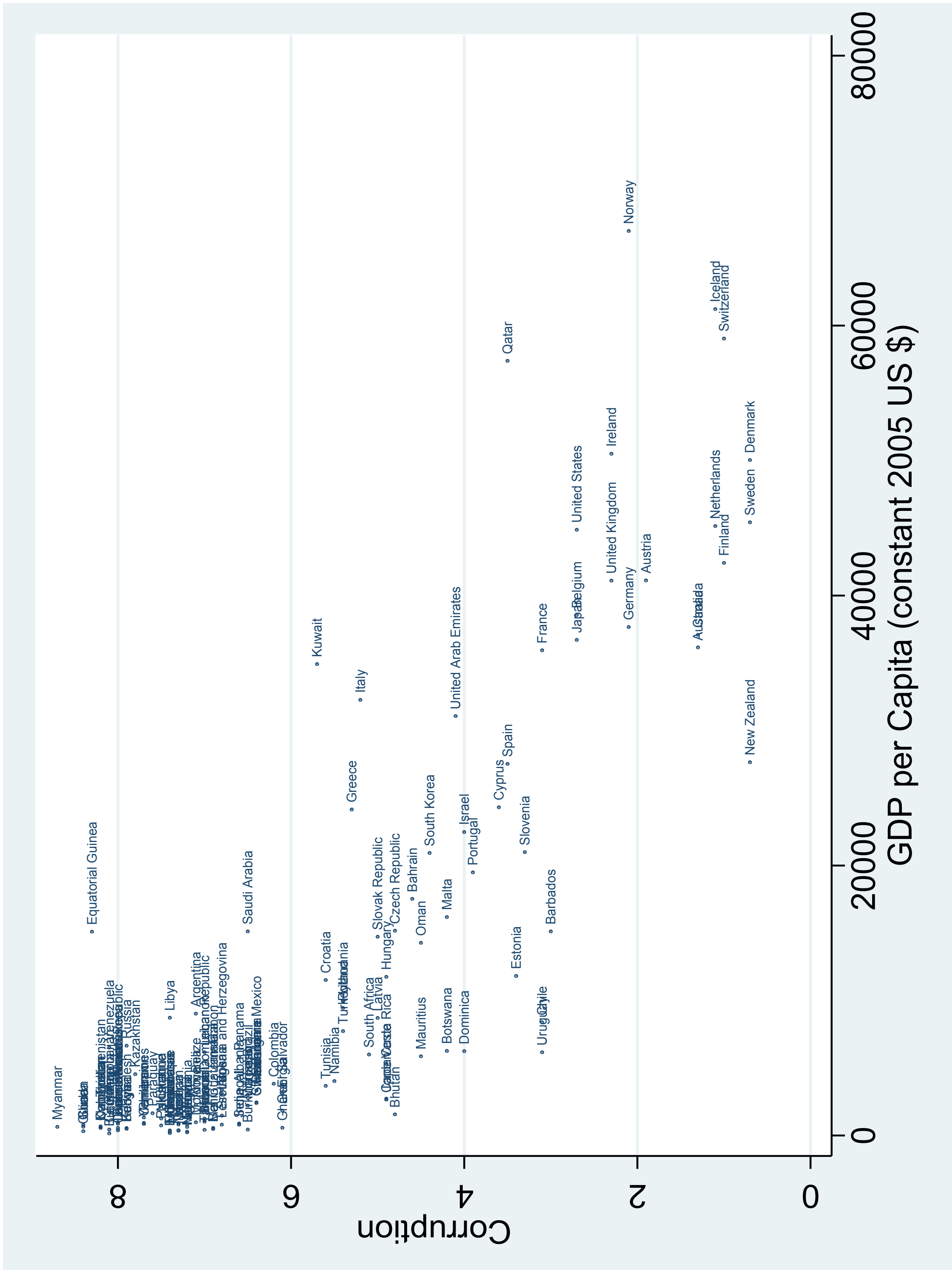

Figure C.3: Plot of Corruption against GDP per Capita(2008)

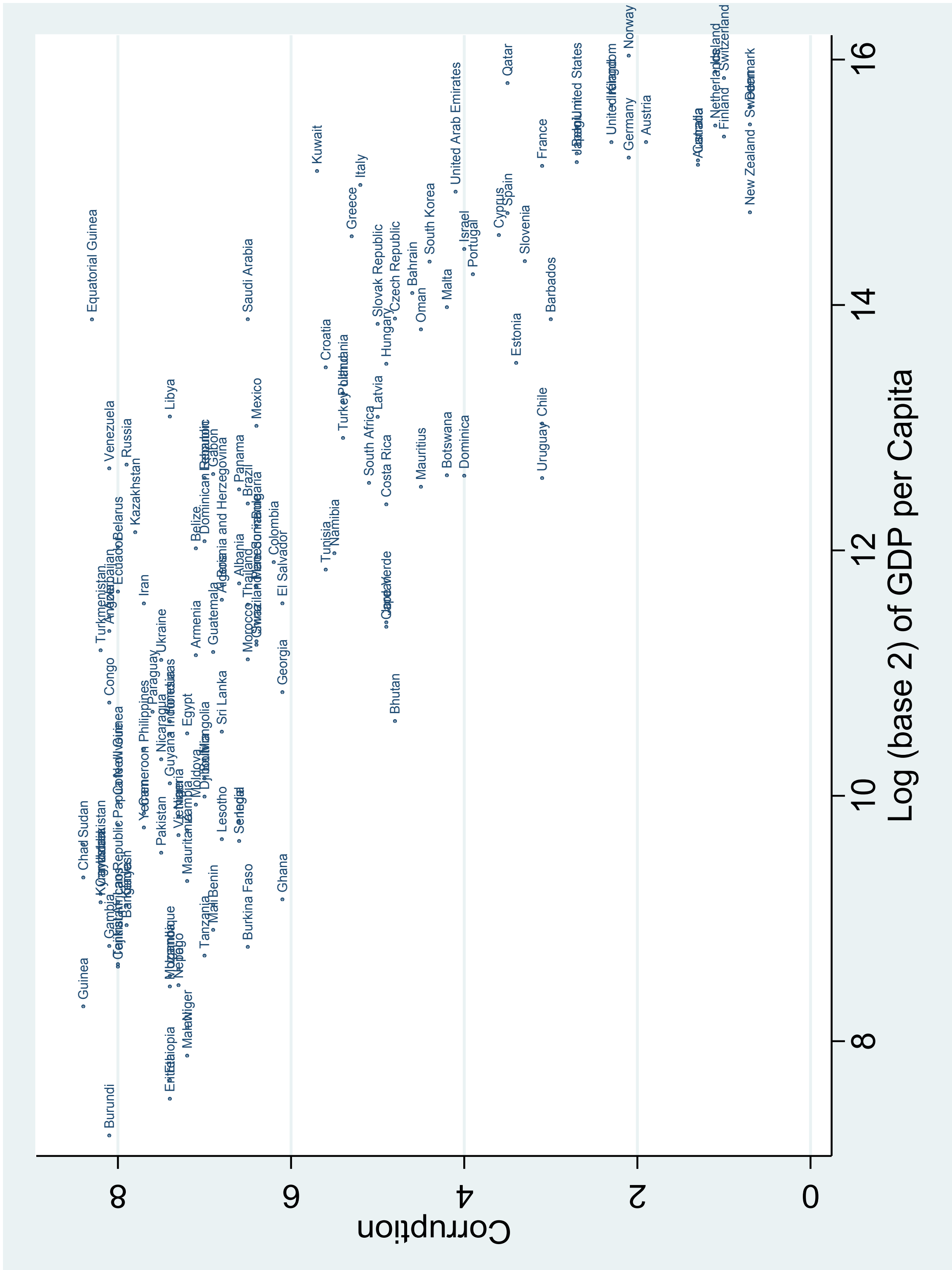

Figure C.4: Plot of Corruption against Log (base 2) of GDP per Capita (2008)

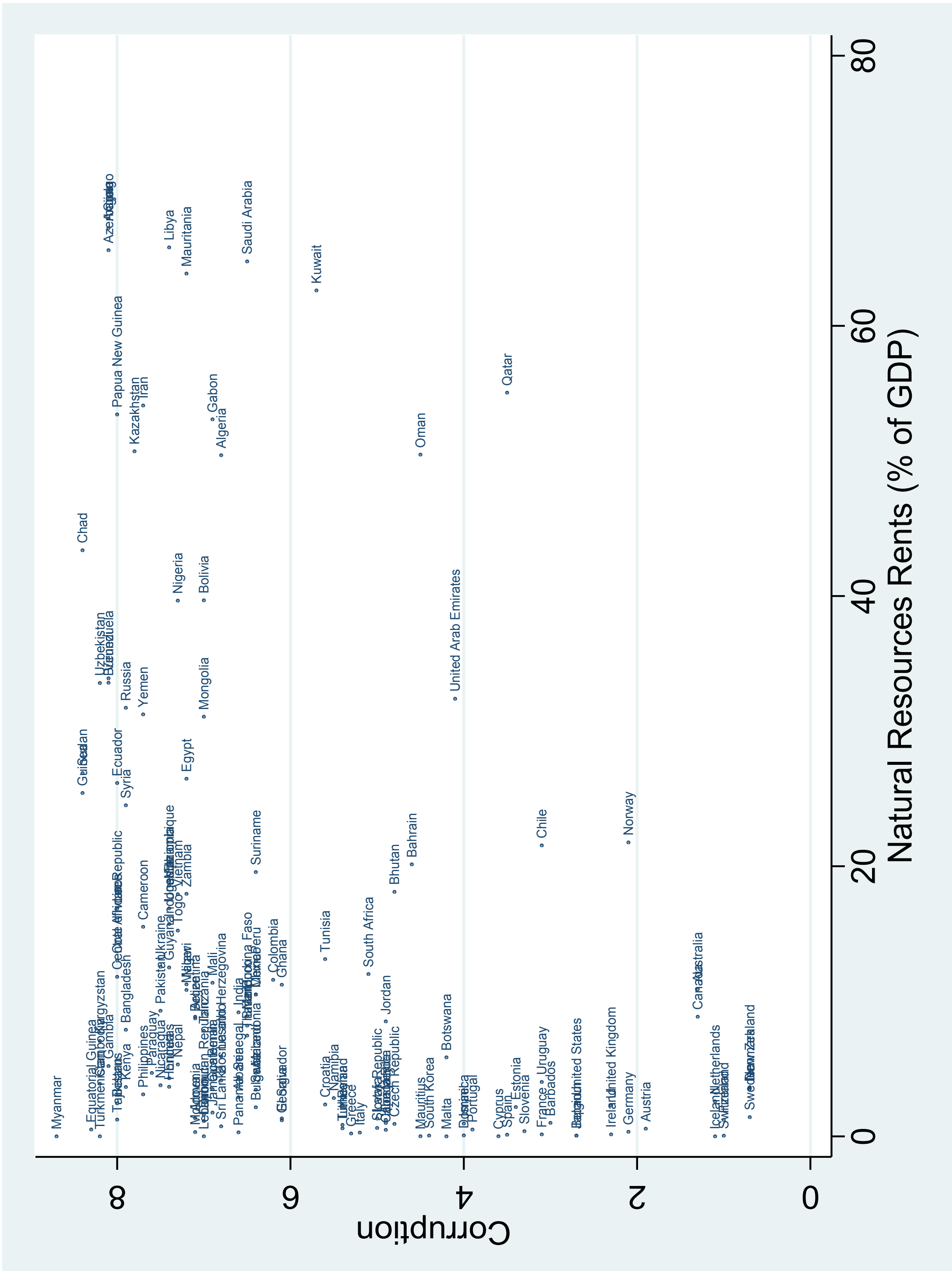

Figure C.5: Plot of Corruption against Natural Resources Rents (2008)

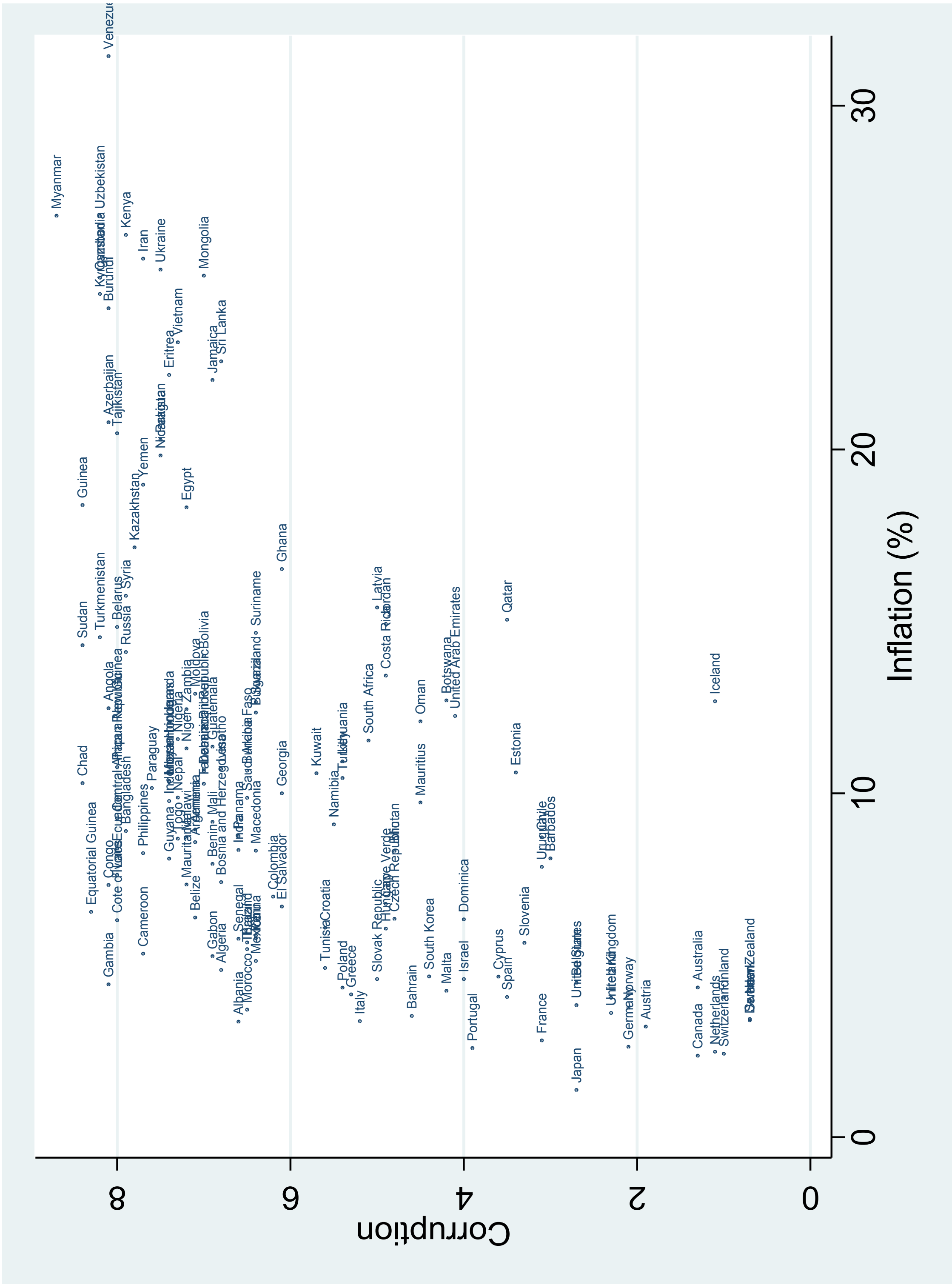

Figure C.6: Plot of Corruption against Inflation (2008)

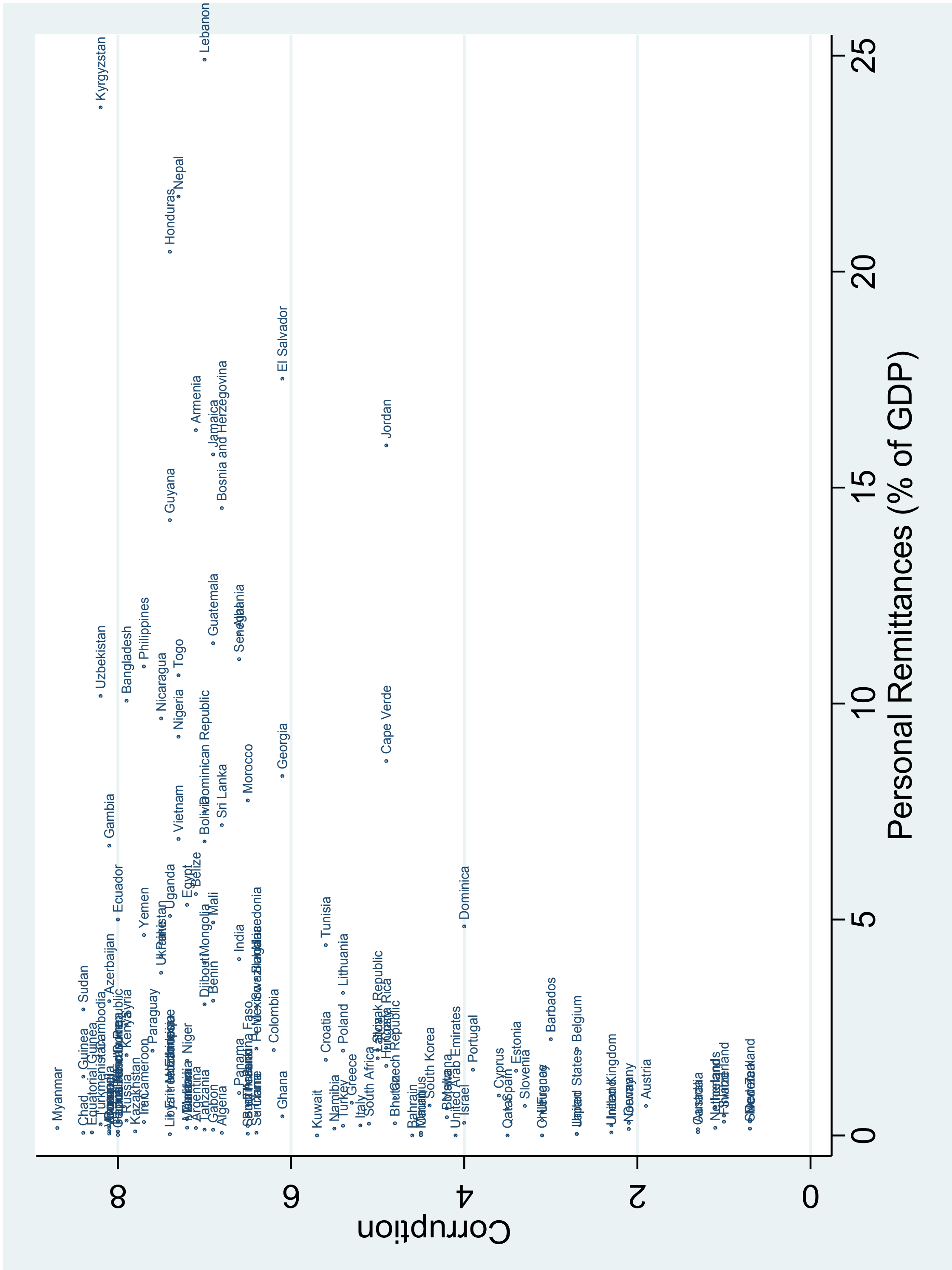

Figure C.7: Plot of Corruption against Personal Remittances (2008)

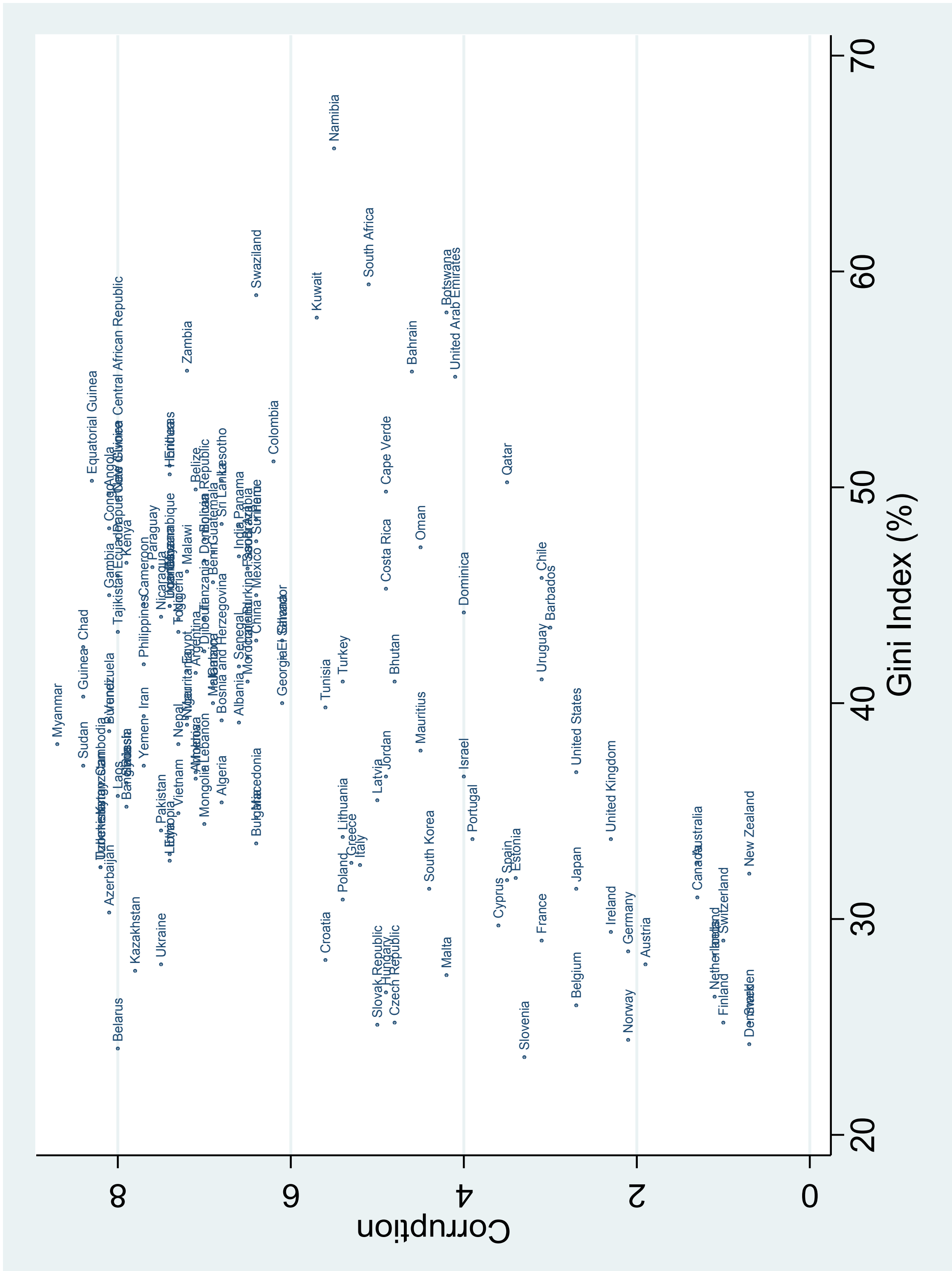

Figure C.8: Plot of Corruption against Gini Index (2008)

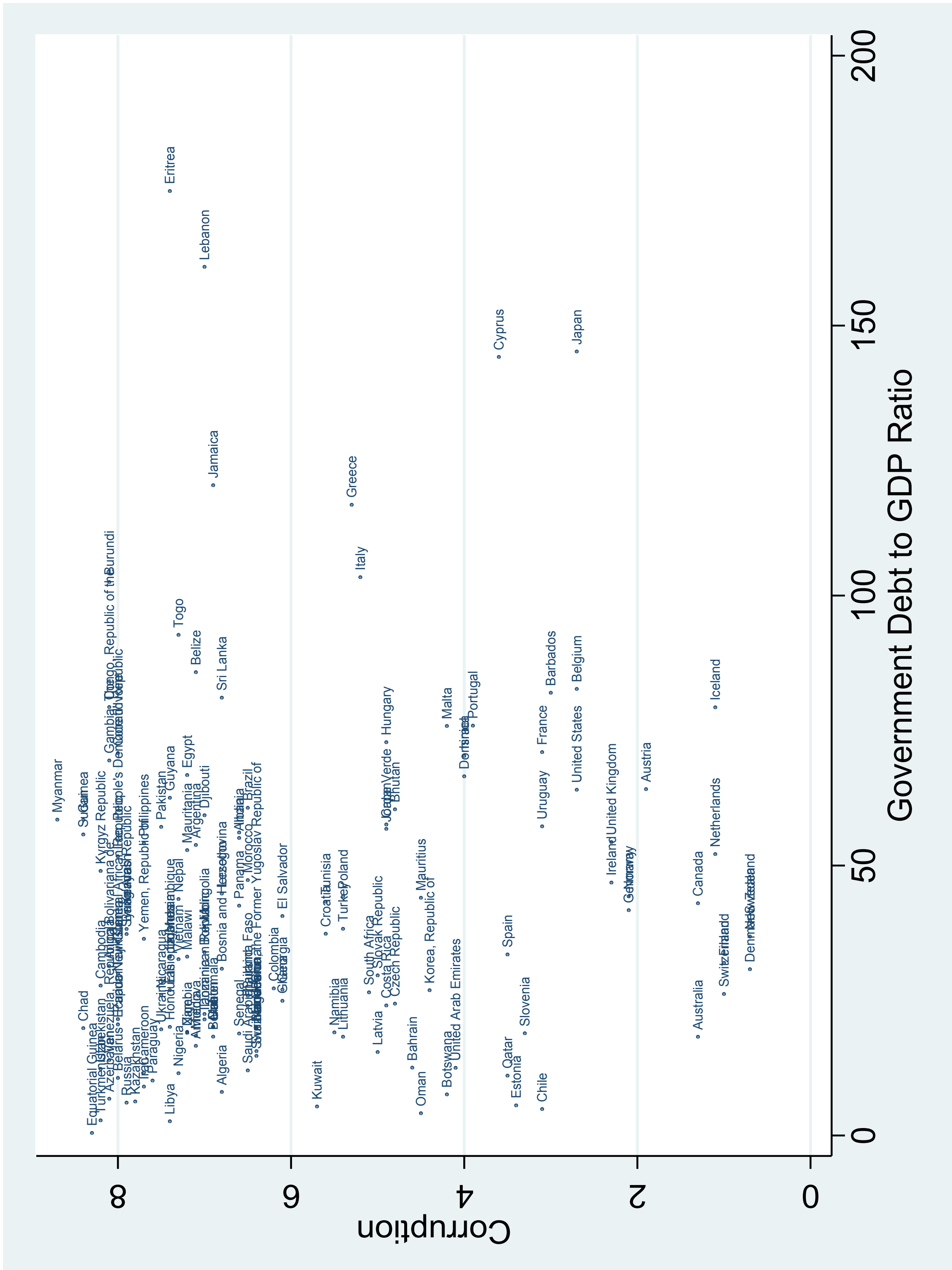

Figure C.9: Plot of Corruption against Government Debt to GDP Ratio (2008)

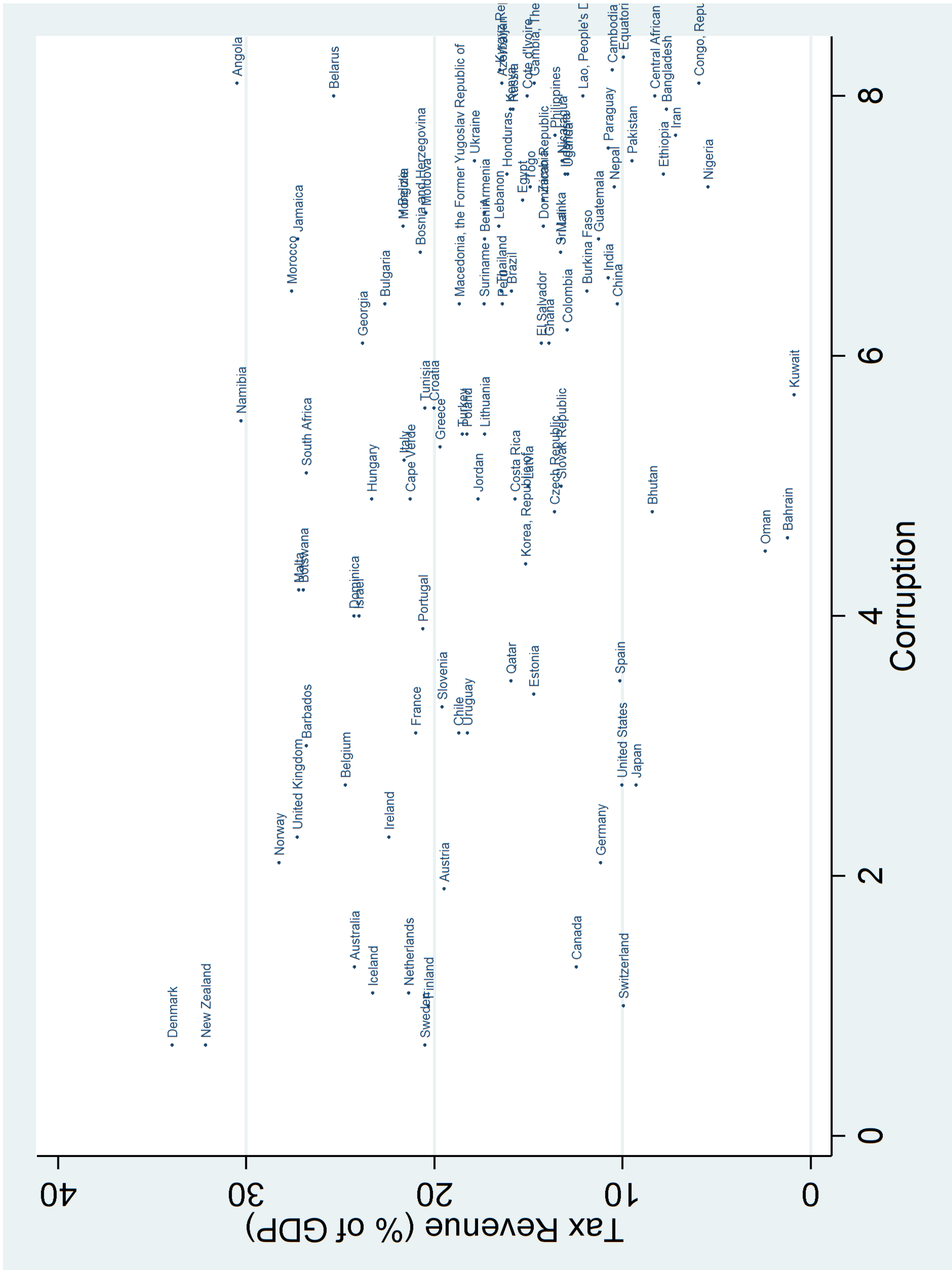

Figure C.10: Plot of Tax Revenue (as % of GDP) collected against Corruption (2008)

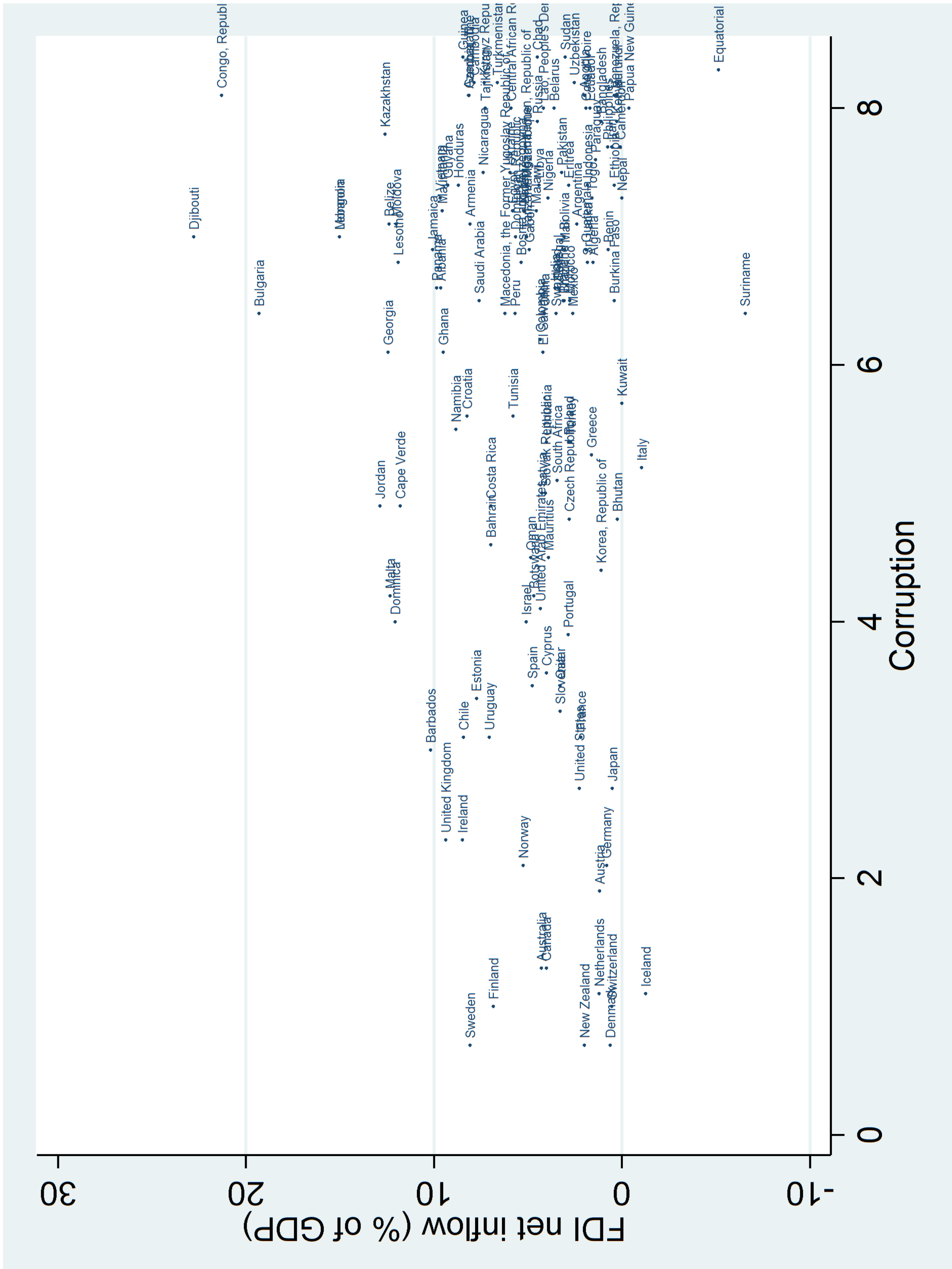

Figure C.11: Plot of Foreign Direct Investment (FDI) net inflow (as % of GDP) against Corruption (2008)

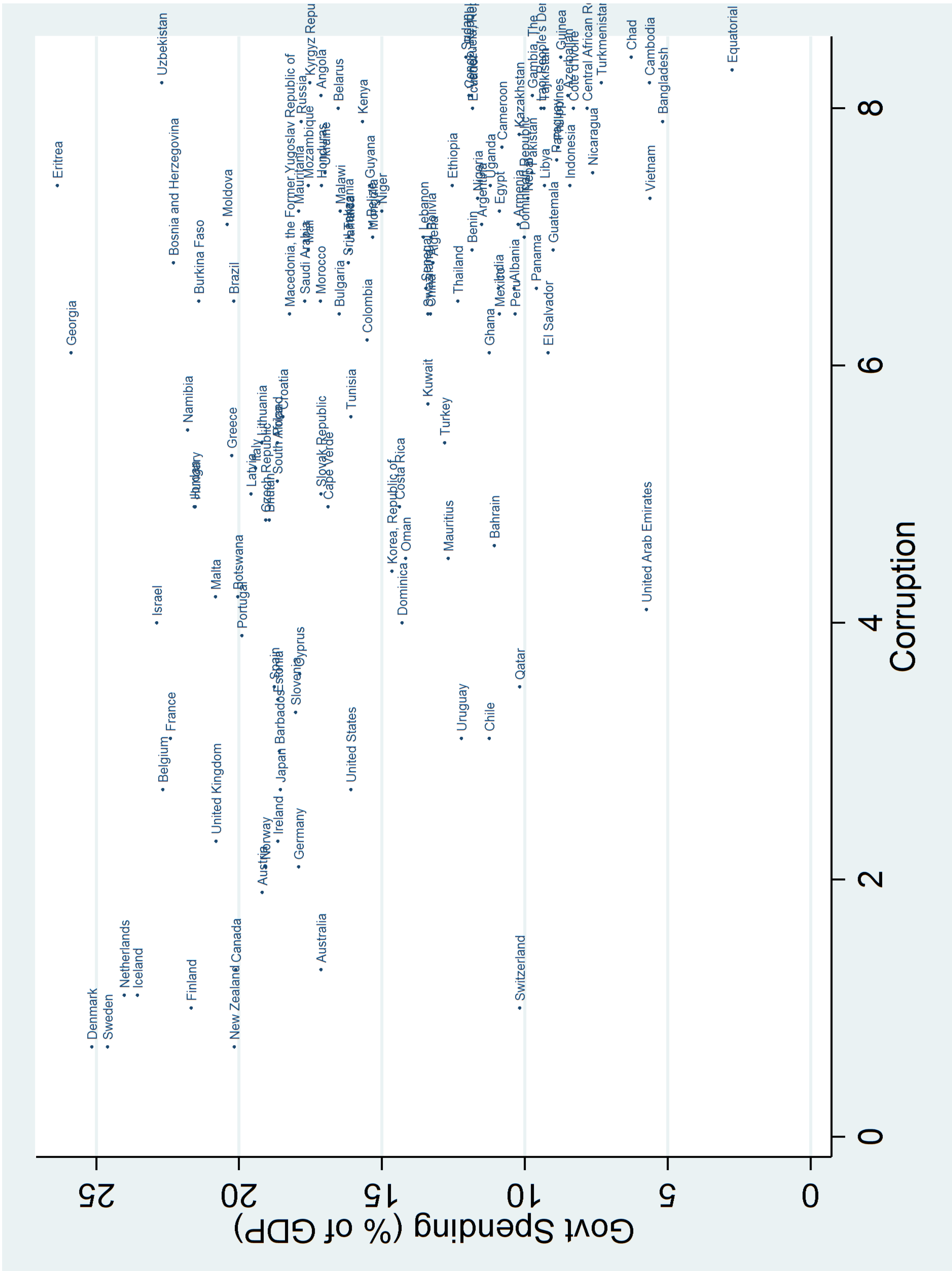

Figure C.12: Plot of Government Spending (as % of GDP) against Corruption (2008)

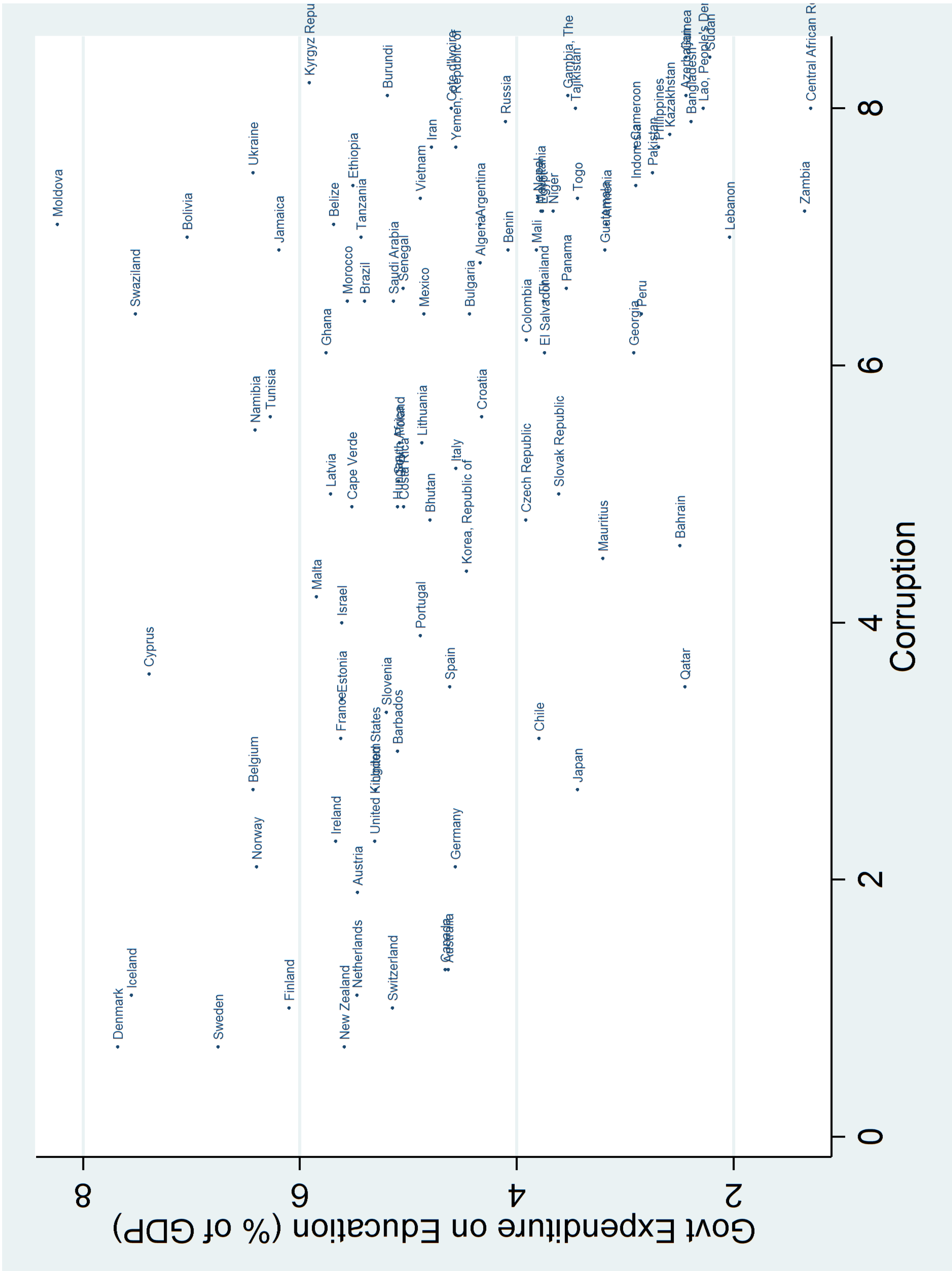

Figure C.13: Plot of Government Spending on Education (as % of GDP) against Corruption (2008)

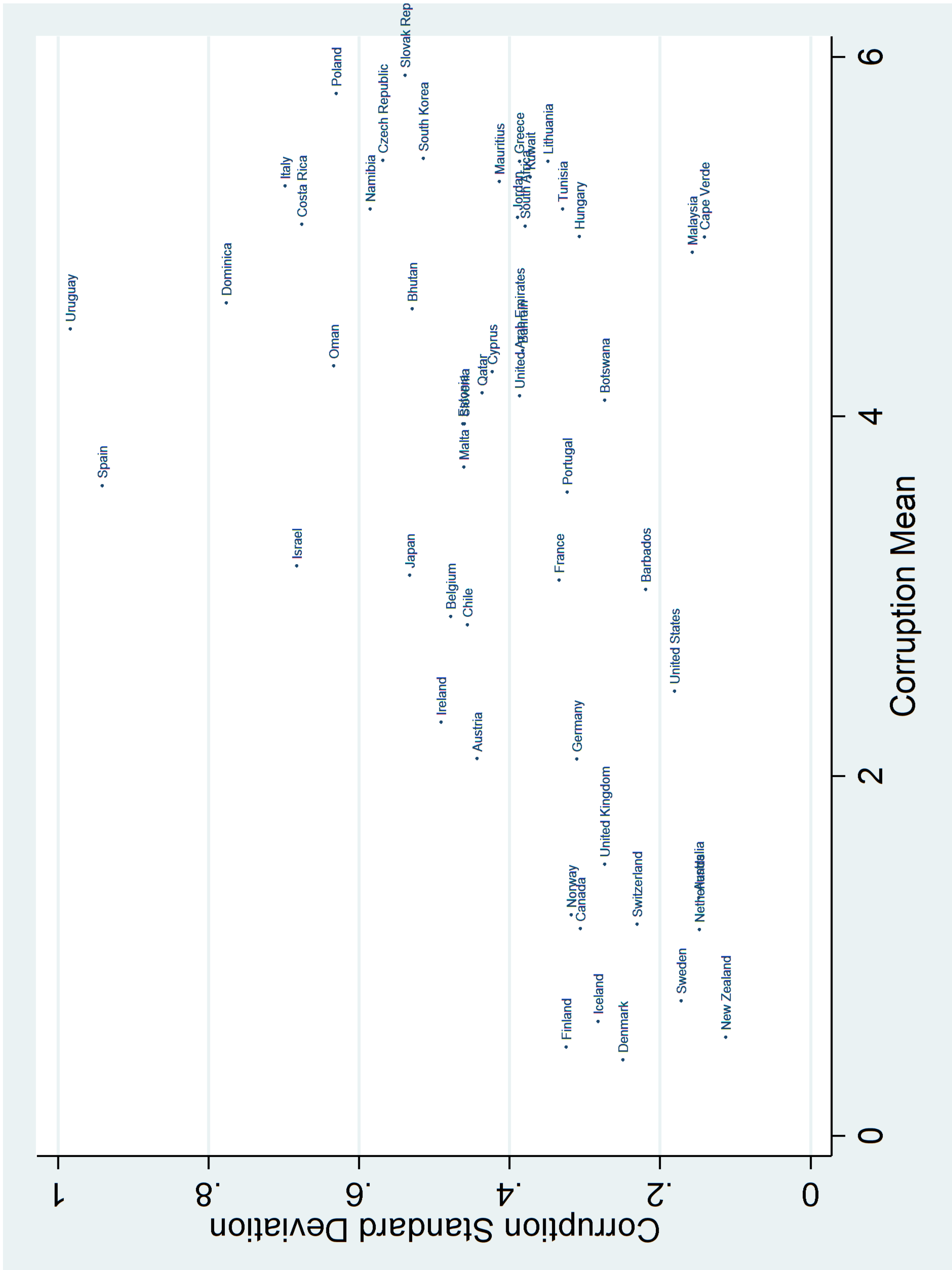

Figure C.14: Plot of Corruption Standard Deviation against Corruption Mean for Low Corruption Countries (CPI from 0 to 6)

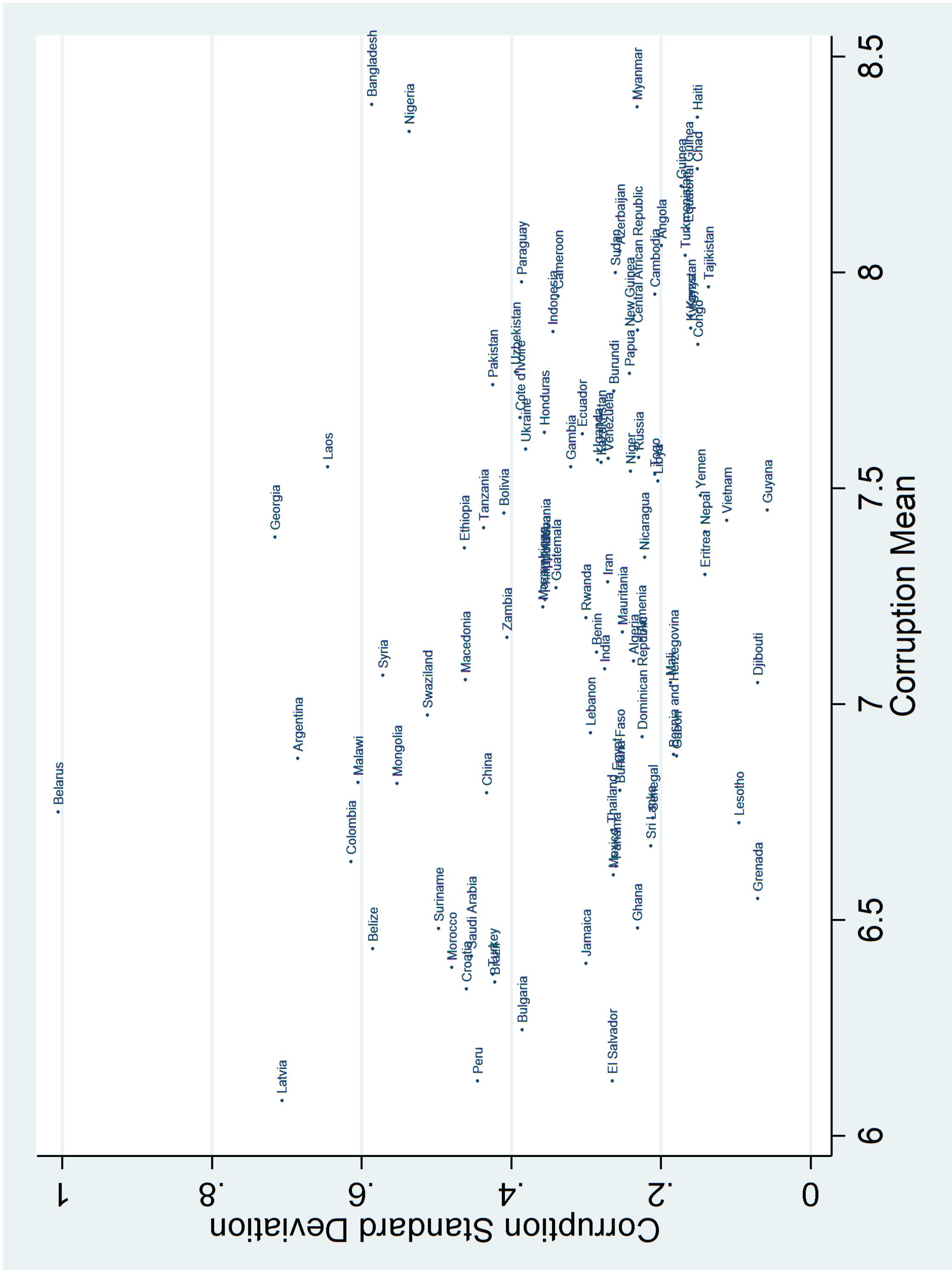

Figure C.15: Plot of Corruption Standard Deviation against Corruption Mean for High Corruption Countries (CPI from 6 to 10)

APPENDIX D

REGRESSION OUTPUT AND P-VALUES

Table D.1: Regression Output of the Corruption Models along with the P-Values

	(Model 1) Corruption	(Model 1) P–Values	(Model 2) Corruption	(Model 2) P–Values
Urban Density	0.000105***		0.000138***	
	(0.000008)	12.78	(0.000009)	15.13
GDP per capita	−0.0000895***		−0.0000804***	
	(0.0000036)	–24.68	(0.0000038)	–21.20
Natural resource rents	0.0298***		0.0317***	
	(0.0018)	16.69	(0.0018)	17.15
Remittances	0.0258***		0.0210***	
	(0.0044)	5.89	(0.0039)	5.44
Inflation	0.0047**		0.0045**	
	(0.0021)	2.17	(0.0021)	2.11
Government Debt	0.0022***		0.0047***	
	(0.0008)	2.64	(0.0008)	5.57
Gini	0.0029		0.0224***	
	(0.0039)	0.74	(0.0049)	4.56
Democracy Duration	−0.0107***		−0.0082***	
	(0.0015)	–7.28	(0.0015)	–5.57
% Christian	0.0059***		0.0012	
	(0.0015)	3.83	(0.0017)	0.69
% Muslim	0.0075***		0.0059***	
	(0.0015)	5.10	(0.0016)	3.78
% Buddhist	0.0105***		0.0095***	
	(0.0022)	4.74	(0.0024)	4.00
Constant	5.669***			
	(0.189)	29.94		
Observations	1340		1340	
Adjusted R^2	0.7958		0.8089	

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D.2: Colony Dummy Variables Coefficients along with the P-Values – Regression Output of the Corruption Models

	(Model 1) Corruption	(Model 2) Corruption	(Model 2) P–Values
Never Colonized		4.669***	
		(0.244)	19.10
Colony of Britain		4.453***	
		(0.275)	16.16
Colony of France		4.578***	
		(0.274)	16.69
Colony of Spain		5.061***	
		(0.298)	16.96
Colony of Portugal		4.396***	
		(0.308)	14.29
Colony of Netherlands		5.396***	
		(0.265)	20.36
Colony of Others		4.749***	
		(0.292)	16.26
Internal Colony of		5.370***	
Austro-Hungarian Empire		(0.205)	26.16
Internal Colony of		5.579***	
Ottoman Empire		(0.242)	23.03
Internal Colony of		5.435***	
Soviet Union		(0.236)	23.01
Observations	1340	1340	
Adjusted R^2	0.7958	0.8089	

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX E

THE VARIANCE COVARIANCE MATRIX OF THE COLONIAL DUMMY VARIABLES

Table E.1: Variance Covariance Matrix of the Colony Dummy Variables with Each Other - Part 1 of Wide Table

	Never	Britain	France	Spain	Portugal
Never	.05975424				
Britain	.06270452	.07589928			
France	.06074166	.07159587	.0752256		
Spain	.0640836	.07569379	.07353093	.08909117	
Portugal	.0642864	.07609141	.07407236	.07998852	.0945809
Netherlands	.05880866	.06756448	.06632142	.0691595	.06943423
Soviet Union	.05017508	.05799121	.05674091	.06266127	.06080647
Ottoman	.05164612	.06023168	.05845447	.06429304	.06277576
Austro-Hungarian	.04253135	.0486955	.0467534	.05263842	.05122637
Other	.06043352	.06708728	.06545293	.06826087	.06977873

Table E.2: Variance Covariance Matrix of the Colony Dummy Variables with Each Other - Part 2 of Wide Table

	Netherlands	Soviet Union	Ottoman	Austro-Hungarian	Other
Never					
Britain					
France					
Spain					
Portugal					
Netherlands	.0702674				
Soviet Union	05354687	.05576712			
Ottoman	.05570749	.05019903	.05866731		
Austro-Hungarian	.04510381	.04140949	.04279378	.04216054	
Other	.06303068	.05383949	.05574303	.04585846	.08536494

APPENDIX F

VARIANCE INFLATION FACTORS

Table F.1: Variance Inflation Factors

	(1) Tax Revenue	(2) FDI Inflow	(1) Govt. Spending	(2) Govt. Spending on Edu.
Corruption	5.74	5.30	5.27	5.71
Urban Density	1.72	1.67	1.68	1.77
GDP per capita	6.23	5.48	5.49	7.44
Natural resource rents	1.67	1.64	1.63	1.53
Remittances	1.38	1.28	1.29	1.35
Inflation	1.18	1.13	1.13	1.11
Government Debt	1.48	1.32	1.34	1.31
Gini	3.30	2.77	2.81	3.24
Democracy Duration	2.83	2.56	2.59	3.10
% Christian	4.42	4.73	4.71	4.87
% Muslim	4.11	3.88	3.87	4.28
% Buddhist	1.78	1.86	1.86	1.99

APPENDIX G

FINDING THE SOLUTION TO THE GRANOVETTER EQUATIONS

The equations that we need to solve are

$$0.01 = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{G.1}$$

$$0.01x = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]. \tag{G.2}$$

Here erf is the error function. This is a nonlinear simultaneous equation model. The easiest way to actually solve this is using the Solve package in Excel for every value of μ. For every value of μ (integer value from 1 to 99) we create a table with the values of the two equations (the PDF and the CDF). We then use the software package to vary the value of sigma between 0 and 50 that solves the simultaneous equation for the different values of μ. The following table gives us the result of the calculations. These values are then used for the graphical analysis.

Mu	Critical Sigma	Mu	Critical Sigma	Mu	Critical Sigma
1	0.08597	34	18.7521	67	17.947
2	0.42986	35	19.5747	68	17.1667
3	0.85972	36	20.4258	69	16.4087
4	1.28957	37	21.307	70	15.6628
5	1.71943	38	22.2157	71	14.9459
6	2.14929	39	23.1563	72	14.2342
7	2.57915	40	24.1338	73	13.5521
8	3.00901	41	25.1529	74	12.8745
9	3.43887	42	26.219	75	12.2204
10	3.89532	43	27.3381	76	11.5772
11	4.38223	44	28.5211	77	10.9432
12	4.86914	45	29.7853	78	10.3353
13	5.35606	46	31.1481	79	9.72731
14	5.8486	47	32.6419	80	9.13927
15	6.38029	48	34.3342	81	8.56807
16	6.91198	49	36.3708	82	7.99687
17	7.44367	50	39.865	83	7.44367
18	7.99687	51	36.3708	84	6.91198
19	8.56807	52	34.3342	85	6.38029
20	9.13927	53	32.6419	86	5.8486
21	9.72731	54	31.1481	87	5.35606
22	10.3353	55	29.7853	88	4.86914
23	10.9432	56	28.5211	89	4.38223
24	11.5772	57	27.3381	90	3.89532
25	12.2204	58	26.219	91	3.43887
26	12.8745	59	25.1529	92	3.00901
27	13.5521	60	24.1338	93	2.57915
28	14.2342	61	23.1563	94	2.14929
29	14.9459	62	22.2157	95	1.71943
30	15.6628	63	21.307	96	1.28957
31	16.4087	64	20.4258	97	0.85972
32	17.1667	65	19.5747	98	0.42986
33	17.947	66	18.7521	99	0.08597

APPENDIX H

THE DEGREE DISTRIBUTION OF A RANDOM NETWORK IS POISSON

In a random network the probability that a node i has exactly k links has three factors:

i) The probability that the node has k links, which is p^k.

ii) The probability that the node does not have the remaining $(n-1-k)$ links, which is $(1-p)^{n-1-k}$.

iii) The number of ways we can choose k links from the $(n-1)$ potential links a node could have, which is $\binom{n-1}{k}$.

Therefore the degree distribution of a random network is the binomial distribution

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}. \tag{H.1}$$

However most random networks (like real networks) are sparse. This means k is much smaller than n, or $k \ll n$. In such case

$$\binom{n-1}{k} = \frac{(n-1)!}{(n-1-k)!k!} \approx \frac{(n-1)^k}{k!}. \tag{H.2}$$

We also know that the average degree of a random network $\langle k \rangle$ is the product of the probability p that two nodes are connected and $(n-1)$ which is the maximum number of links any node can have in a n sized network. Or

$$\langle k \rangle = p(n-1). \tag{H.3}$$

Or

$$p = \frac{\langle k \rangle}{(n-1)}. \tag{H.4}$$

Also from taylor expansion of natural logarithms and the fact that $k \ll n$we get

$$\begin{aligned} ln[(1-p)^{n-1-k}] &= (n-1-k)ln(1-p) \\ &= (n-1-k)ln\left(1 - \frac{\langle k \rangle}{(n-1)}\right) \\ &\approx -(n-1-k)\left(\frac{\langle k \rangle}{(n-1)}\right) \\ &\approx -\langle k \rangle. \end{aligned} \tag{H.5}$$

Taking exponentials on both sides leads to

$$(1-p)^{n-1-k} = e^{-\langle k \rangle}. \tag{H.6}$$

Therefore the binomial distribution of the random network from equation H.1 in the limit of large n becomes

$$\begin{aligned} p_k &= \frac{(n-1)^k}{k!} p^k e^{-\langle k \rangle} \\ &= \frac{(n-1)^k}{k!} \left(\frac{\langle k \rangle}{(n-1)}\right)^k e^{-\langle k \rangle} \\ &= e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \end{aligned} \tag{H.7}$$

Equation H.7 is the Poisson Distribution. Thus for large n the random network $G(n, p)$ has a Poisson degree distribution and is also called a Poisson random network.

REFERENCE LIST

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5):1369–1401.

Ades, A., & Di Tella, R. (1999). Rents, competition, and corruption. *American Economic Review*, *89*(4):982–993.

Aidt, T. S. (2003). Economic analysis of corruption: A survey. *The Economic Journal*, *113*(491).

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1):47.

Alvaredo, F., Atkinson, A., Piketty, T., & Saez, E. (2020). *World inequality database.* https://WID.world.

Barabási, A.-L., & Pósfai, M. (2016). *Network science*. Cambridge University Press.

Bardhan, P. (1997). Corruption and development: A review of issues. *Journal of Economic Literature*, *35*(3):1320–1346.

Biggs, N., Lloyd, E. K., & Wilson, R. J. (1986). *Graph theory, 1736-1936*. Oxford University Press.

Boix, C., Miller, M., & Rosato, S. (2018). Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2015.

Boix, C., & Stokes, S. C. (2003). Endogenous democratization. *World Politics*, *55*(4):517–549.

Bollobás, B., & Riordan, O. (2004). The diameter of a scale-free random graph. *Combinatorica*, *24*(1):5–34.

Braun, M., & Di Tella, R. (2004). Inflation, inflation variability, and corruption. *Economics & Politics*, *16*(1):77–100.

Brown, D., & James, P. (2015). Religious characteristics of states dataset, phase 1: Demographics (RCS). *Association of Religion Data Archives*.

Calabrese, F., Dahlem, D., Gerber, A., Paul, D., Chen, X., Rowland, J., Rath, C., & Ratti, C. (2011). The connected states of America: Quantifying social radii of influence. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 223–230. IEEE.

Cohen, R., & Havlin, S. (2003). Scale-free networks are ultrasmall. *Physical Review Letters*, *90*(5):058701.

de Sola Pool, I., & Kochen, M. (1978). Contacts and influence. *Social Networks*, *1*(1):5–51.

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics*, *112*(4):1203–1250.

Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, *5*(1):17–60.

Gilbert, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, *30*(4):1141–1144.

Glaeser, E. L., La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth*, *9*(3):271–303.

Graham, B. A. T., & Tucker, J. R. (2016). *The International Political Economy Data Resource.* V1.5, Harvard Dataverse. https://doi.org/10.7910/DVN/X093TV.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6):1360–1380.

Granovetter, M. S. (1978). Threshold models of collective behavior. *American Journal of Sociology*, *83*(6):1420–1443.

Grimmett, G. (1999). *Percolation*. Springer.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, *102*(2):994–1028.

Huntington, S. P. (1968). *Political order in changing societies*. Yale University Press.

Kadanoff, L. P. (1966). Scaling laws for Ising models near T_c. *Physics Physique Fizika*, *2*(6):263.

Kaufmann, D. (2005). Myths and realities of governance and corruption. *World Economic Forum Global Competitiveness Report 2005-2006*.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2009). Governance matters VIII: Aggregate and individual governance indicators, 1996-2008. *World Bank Policy Research Working Paper*, (4978).

Kranton, R. E. (1996). Reciprocal exchange: A self-sustaining system. *The American Economic Review*, *86*(4):830–851.

Kunicova, J., & Rose-Ackerman, S. (2005). Electoral rules and constitutional structures as constraints on corruption. *British Journal of Political Science*, *35*(4):573–606.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., & Vishny, R. (1999). The quality of government. *The Journal of Law, Economics, and Organization*, *15*(1):222–279.

Lange, M., & Dawson, A. (2009). Dividing and ruling the world? A statistical test of the effects of colonialism on postcolonial civil violence. *Social Forces*, *88*(2):785–817.

Ledeneva, A. V. (1998). *Russia's economy of favours: Blat, networking and informal Exchange*, volume 102. Cambridge University Press.

Leff, N. H. (1964). Economic development through bureaucratic corruption. *American Behavioral Scientist*, *8*(3):8–14.

Maddison, A. (2003). *The world economy: Historical statistics*. OECD.

Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, *110*(3):681–712.

Mbaye, S., Badia, M. M. M., & Chae, K. (2018). *Global debt database: Methodology and sources*. International Monetary Fund.

Méon, P.-G., & Weill, L. (2010). Is corruption an efficient grease? *World Development*, *38*(3):244–259.

Milgram, S. (1967). The small world problem. *Psychology Today*, *2*(1):60–67.

Montgomery, J. (2009). *Mathematical Models of Social Systems*. University of Wisconsin. https://www.ssc.wisc.edu/%7Ejmontgom/376textbook.htm.

Montinola, G. R., & Jackman, R. W. (2002). Sources of corruption: A cross-country study. *British Journal of Political Science*, *32*(1):147–170.

Newman, M. (2018). *Networks*. Oxford University Press.

Oldenburg, P. (1987). Middlemen in third-world corruption: Implications of an Indian case. *World Politics*, *39*(4):508–535.

Pan, W., Ghoshal, G., Krumme, C., Cebrian, M., & Pentland, A. (2013). Urban characteristics attributable to density-driven tie formation. *Nature Communications*, *4*:1961.

Robertson, E. (2014). Kickstarting memes and movements: A distribution-based threshold model. *Santa Fe Institute Research Experience for Undergraduates Program*.

Sayama, H. (2015). *Introduction to the modeling and analysis of complex systems*. Open SUNY Textbooks.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, *1*(2):143–186.

Schläpfer, M., Bettencourt, L. M., Grauwin, S., Raschke, M., Claxton, R., Smoreda, Z., West, G. B., & Ratti, C. (2014). The scaling of human interactions with city size. *Journal of the Royal Society Interface*, *11*(98):20130789.

Shleifer, A. (1997). Government in transition. *European Economic Review*, *41*(3):385–410.

Small, C., Elvidge, C. D., Balk, D., & Montgomery, M. (2011). Spatial scaling of stable night lights. *Remote Sensing of Environment*, *115*(2):269–280.

Stauffer, D., & Aharony, A. (1991). *Introduction to percolation theory*. Taylor & Francis.

Transparency International (2012). *Corruption perceptions index 2012*. Berlin, Germany. Transparency International.

Travers, J., & Milgram, S. (1969). An experimental study of the small world problem. *Sociometry*, *32*(4):425–443.

Treisman, D. (2000). The causes of corruption: A cross-national study. *Journal of Public Economics*, *76*(3):399–457.

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.

Ugo, P. (2001). Electoral rules, political systems, and institutional quality. *Economics & Politics*, *13*(3):311–342.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, *393*(6684):440–442.

Wei, S.-J., & Zeckhauser, R. (1999). Dark deals and dampened destinies: Corruption and economic performance. *Japan and the World Economy*, *11*(3):443–454.

Wiedermann, M., Smith, E. K., Heitzig, J., & Donges, J. F. (2020). A network-based microfoundation of Granovetter's threshold model for social tipping. *Scientific reports*, *10*(1):1–10.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

World Bank (2015). *World development indicators 2015*. Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/21634.

Ziliak, S., & McCloskey, D. N. (2008). *The cult of statistical significance: How the standard error costs us jobs, justice, and lives*. University of Michigan Press.

Zhou, N., Hubacek, K., & Roberts, M. (2015). Analysis of spatial patterns of urban growth across South Asia using DMSP-OLS nighttime lights data. *Applied Geography*, *63*:292–303.

VITA

Rafed Amin Al-Huq was born in August 30, 1977 in Dhaka, Bangladesh. He attended the California Institute of Technology, where he earned a Bachelor of Science degree in Engineering & Applied Science in 2000.

In 2010 Rafed moved to Kansas City to attend the University of Missouri - Kansas City where he earned a Master of Arts degree in Economics in 2012. A year later he joined the Ph.D. program in Economics.

Since 2017, Rafed has been a visiting Instructor of Economics at various liberal arts colleges in the East Coast. From 2019 to 2022, he taught at Franklin & Marshall College. In 2022, he moved to New Orleans and is currently teaching at Tulane University as a visiting lecturer.