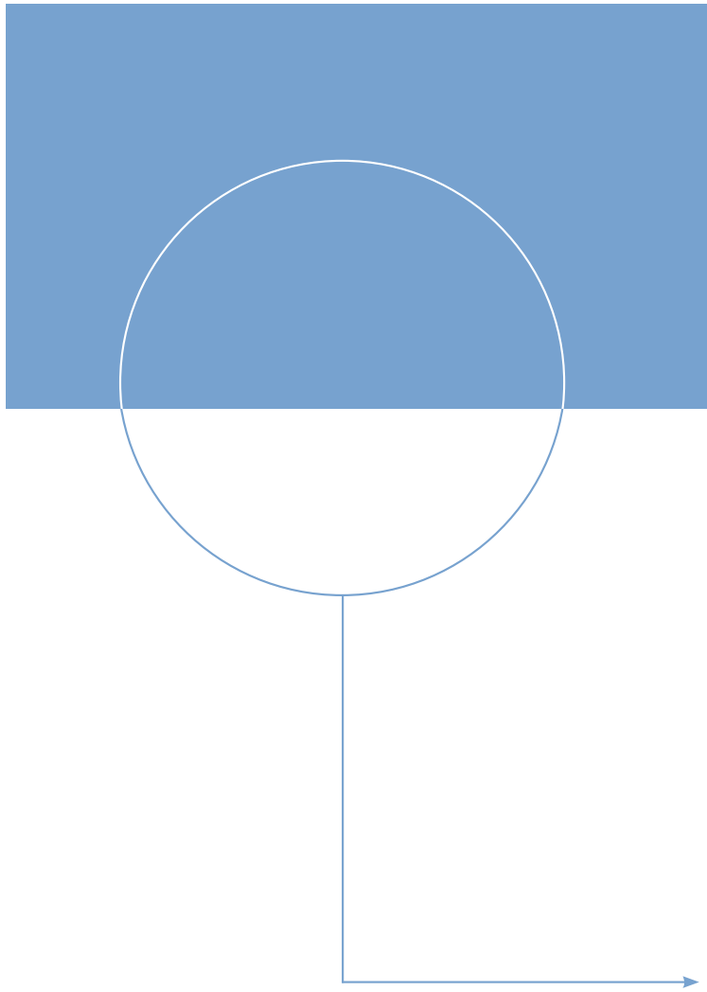


Master's thesis

2019

Andreas Knutson

NTNU
Norwegian University of
Science and Technology
Faculty of Social and Educational Sciences
Department of Sociology and Political Science

Andreas Knutson

Cutting Out the Middlemen

An empirical analysis of e-government as an anti-corruption tool from 2003-2018

June 2019

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Cutting Out the Middlemen

An empirical analysis of e-government as an anti-corruption tool from 2003-2018

Andreas Knutson

Master's thesis in Political Science
Submission date: June 2019
Supervisor: Indra de Soysa

Norwegian University of Science and Technology
Department of Sociology and Political Science

Abstract

This paper study the effect of e-government on political corruption levels. E-government relates to extended use of information and communication technology (ICT) in the public sector. Both international organisations and research view e-government as a way of boosting efficiency, transparency and accountability in public institutions. Researchers on corruption have long identified these traits as important for determining corruption levels which has led to the synthesis of the two fields looking at the effect of e-government on corruption. However, the research has been characterised by anecdotal evidence and a technological determinism painting e-government as a universal tool for combating political corruption. Since e-government is evolving rapidly, I set out to investigate the question *does e-government lead to lower levels of political corruption?* Increasing data availability allows for constructing longer datasets than what has been available to previous researchers. Another novelty is the use of the Varieties of democracy political corruption index. In my empirical analysis I employ a time-series analysis from 2003-2018. Moreover, a Granger causality test and an Instrumental variable regression (IV regression) are used to determine the direction of causality and minimise endogeneity bias. In contrast to previous studies, the empirical results are unclear. An interaction term between e-government and democracy, however, indicates that e-government is dependent on democratic institutions to work as a corruption deterrent. Weak democracies see an increase in corruption with higher levels of e-government while advanced democracies experience the opposite effect. These findings suggest that the technological capabilities needed for e-government services can be used to increase censorship and strangle the democratic accountability needed to prevent corrupt practices. Regarding the causal direction, the Granger causality test yield ambiguous results, while the IV regression indicates that the arrow points from e-government to lowered corruption levels. E-government does therefore seem to not have a uniform impact on corruption globally. However, somewhat unclear results and rapidly changing technology present a need for further research.

Sammendrag

Denne studien undersøker effekten av digital forvaltning på nivået av politisk korrupsjon. Digital forvaltning omhandler den utstrakte broken av informasjons- og kommunikasjonsteknologi I offentlig sektor. Både internasjonale organisasjoner og forskere ser digital forvaltning som en måte å øke effektivitet, transparens og ansvarlighet. Korrupsjonsforskningen har lenge sett disse trekkene som viktige for å forutse korrupsjonsnivåer, og en syntese av disse feltene har ført til fremveksten av en ny litteratur. Likevel har forskingen vært preget av anekdotiske bevis og en teknologisk determinisme som tegner digital forvalting som en universalløsning for korrupsjonsbekjempelse. Siden digital forvaltning er i rask utvikling, trengs alltid ny forskning og jeg har derfor valgt å undersøke spørsmålet *fører digital forvaltning til lavere nivåer av politisk korrupsjon?* Økende tilgjengelighet til data tillater meg å undersøke større datasett enn det som finnes i tidligere forskning. En annen nyvinning er bruken av Varieties of democracys indeks for politisk korrupsjon. I min empiriske analyse benytter jeg en tidsserie analyse for årene 2003-2018. I tillegg bruker jeg en Granger kausalitetstest og en instrumentell variabel-regresjon (IV-regresjon) for å undersøke de kausale forholdene og for å minimere endogenitetsbias. I motsetning til tidligere studier, er de empiriske resultatene uklare. Et interaksjonsledd mellom digital forvaltning og demokrati indikerer at digital forvaltning er avhengig av demokratiske institusjoner for å virke preventivt mot korrupsjon. Svake demokratier ser en økning i politisk korrupsjon med digitalisering, mens i avanserte demokratier får man de motsatte resultatene. Disse funnene gir indikasjoner på at den teknologiske kapasiteten som ligger til bunn for den digitale forvaltningen kan brukes til å øke sensur og kvele den demokratiske ansvarligheten som trengs for å redusere korrupsjon. Angående retningen på kausaliteten, så viser Granger-testen uklare resultater, mens IV-regresjonen tyder på at årsaksretningen går fra digital forvaltning til en reduksjon i korrupsjonsnivåene. Uansett, så tyder resultatene på at digital forvaltning ikke har en uniform effekt på korrupsjon globalt. Uklare resultater og teknologi i rask endring presenterer likevel et behov for videre forskning.

Forord

Denne masteroppgaven setter et punktum for et snart seks år langt studieløp på NTNU og studietiden. Jeg kom til byen i 2013 med en plan om å være i ett år. Tilfeldighetene ville det annerledes. Eller, det er kanskje rettere å si at studentbyen Trondheim ville det annerledes, og det er mange som fortjener en takk for at jeg har stortrivdes her.

De fire årene jeg har fått ved institutt for sosiologi- og statsvitenskap har gitt meg mye og fortjener en spesiell takk. Takk til Indra de Soysa som har veiledet meg gjennom både bachelor- og masteroppgaven, ved å alltid ha døra på gløtt. På instituttet har jeg også fått lov til å prøve meg som både læringsassistent og som vitenskapelig assistent for Stefan Geiss, noe jeg er svært takknemlig for. Ikke minst fortjener gjengen en stor takk for å ha gitt meg plass i et fantastisk klassemiljø.

Utenfor Dragvoll er det mange som fortjener en takk som vennene mine fra Bodø. Likevel vil jeg vise spesiell takknemmelighet til Anja som har betydd så mye for meg siden vi møttes i ISFiT høsten 2016. Du er en stor støtte i hverdagen med din tålmodighet - spesielt gjennom de siste månedene med masterskriving. Takk for at du inspirerer meg til å alltid gjøre mitt beste.

Takk til mamma, pappa, Espen, Øyvind og Tom. Dere vil alltid ha en enorm betydning for meg.

Andreas Knutson

Table of contents

List of figures

List of tables

Abbreviations

BEEPS	Business Environment and Enterprise Performance Survey
BLUE	Best Linear Unbiased Estimation
CPI	Corruption Perceptions Index
DK	Driscoll-Kraay Standard Errors
EGDI	E-Government Development Index
EoC	Economics of Crime
FE	Fixed Effect
FGLS	Feasible General Least Squares
G2B	Government to Business
G2C	Government to Citizen
G2G	Government to Government
GDP	Gross Domestic Product
GPT	General Purpose Technology
HCI	Human Capital Index
ICRG	International Country Risk Guide
ICT	Information and Communication Technology
IRT	Item Response Theory
ITU	International Telecommunication Union
IV	Instrumental Variable Regression
LDV	Lagged Dependent Variable
LM	Lagrange Multiplier
NGO	Non-Governmental Organisation
NW	Newey-West Standard Errors
OLS	Ordinary Least Squares
OPEN	Online Procedures ENhancement for civil application
OSI	Online Service Index
OTD	Optical Transient Detector
RE	Random Effect
RESET	Ramsey Regression Equation Specification Error Test
SD	Standard Deviation
SDG	Sustainable Development Goals
SMG	Seoul Metropolitan Government
SORM	System for Operative Investigative Activities
TII	Telecommunication Infrastructure Index
TSCS	Time Series Cross Section
UN	United Nations
UNDP	United Nations Developmental Programme
UNESCO-UIS	United Nations Educational, Scientific and Cultural Organisation - Institute for Statistics
V-DEM	Varieties of Democracy
VIF	Variance Inflation Factor
WB	World Bank
WDI	World Development Indicators
WENAO	Western Europe and Oceania
WGI	World Governance Indicators
WWW	World Wide Web

1 Introduction

Does e-government reduce political corruption? International organisations such as the United Nations and the World Bank have identified various forms of corruption as one of the single greatest obstacles to development (Elbahnasawy, 2014). The United Nations (2018) see corruption in the public sector and politics as an obstacle to reaching the Sustainable Development Goals. Meanwhile, the ubiquity of information and communication technology (ICT) has inspired policymakers and researchers to propose e-government, the extended use of ICT in in the public sector, as a method of controlling political corruption. E-government is supposed to reduce corruption and build inclusive institutions at all levels through increasing transparency, accountability and efficiency (United Nations, 2018; Wirtz & Daiser, 2018; Yildiz, 2012). Researchers generally agree that corruption in various forms distorts both political and economic institutions intended to increase and maintain inclusive development (Acemoglu & Robinson, 2012; Elbahnasawy, 2014; United Nations, 2018). Combating corruption is therefore essential when maintaining and building inclusive institutions that allow for an economic and political development that comprises the whole of society.

Acemoglu and Robinson (2012) define inclusive institutions as characterised by democratic values such as power-sharing, rule of law, and the provision of public services. Extractive institutions are, on the other hand, characterised by the lack of these values which can lead to an eventual decline in growth, with the risk of bringing development to a complete halt. By providing political stability and a reliable justice system, inclusive institutions incentivise long-term investment, innovation and lead to inclusive growth. Political corruption levels are arguably an indication of how inclusive, or rather, how extractive institutions are. We can therefore count corruption as a possible root to a vicious circle of underdevelopment comparable to the forces driving extractive institutions which eventually only serve the ones in power (Acemoglu & Robinson, 2012; Bueno de Mesquita & Smith, 2012).

E-government as an anti-corruption tool makes intuitive sense when interpreting corruption as a principal-agent problem. The two issues of both decentralised information and information asymmetry keep the principals (citizens), uninformed and prey to corrupt institutions (the agents) (Della Porta & Vannucci, 1999; Laffont & Martimort, 2002; Rose-Ackerman, 1978, 2006). E-government is perceived to have an immense potential for improving government and problems with corruption by creating a networked structure of interconnectivity, aiding in

service delivery, increasing efficiency and effectiveness, interactivity, decentralisation, transparency, and accountability. In other words, e-government arguably minimise the discretionary powers of civil servants, while empowering citizens. Moreover, knowing how to build and sustain inclusive development is increasingly important in a politically and economically volatile world where democratic and inclusive institutions are experiencing a backslide and a "third wave" of autocratisation (Lührmann et al., 2019). I therefore set out to ask the question *does e-government cause lower levels of political corruption?*

Previous research has indeed found positive effects of e-government on corruption control and levels. However, much of the research is still relying on anecdotal evidence and the demand for rigorous quantitative studies is not being met. Indeed, the size of the quantitative literature is growing, but the availability of data has been scant, and the literature has been characterised by a large focus on model specification and less on rigorous analysis of the how and whys of e-government and corruption. I believe that the literature on the importance of inclusive and stable institutions can help situate the e-government research more firmly in place in an already rich literature. I set out to contribute to closing that gap in the e-government literature by adding an auxiliary objective to further studying the role of democratic institutions in the nexus between e-government and corruption. Studying how e-government behaves in different institutional contexts is increasingly important seeing as the technology is changing rapidly and research is catching up to realities by leaving technological determinism behind. The backslide of democracies and emergence of autocratic institutions globally is a cause for worry as the technology behind e-government can be abused in a myriad of ways by governments.

By employing a new dataset based on the Varieties of democracy project data on corruption and democracy, I add further novelty to my theses. Former quantitative research on e-government has relied on even shorter panels and, I argue, less valid operationalisations of political corruption and democracy (McMann, Pemstein, Seim, Teorell, & Lindberg, 2016). Former studies are dominated by statistically significant findings which gives the impression that e-government is a quick fix for corruption issues. From the empirical analysis in my theses, I show how e-government is associated with corruption is dependent on both model specifications and the sample employed. That is, I did not find a universally positive or significant improvement in corruption levels associated with e-government. Moreover, I find that the effect of e-government as an anti-corruption tool is dependent on democratic institutions. In fact, weak democratic control leads e-government to increase corruption. E-government seems to strengthen already existing institutions whether these are inclusive or

extractive. This leads me to conclude that previous studies have had a much too optimistic view of e-government and its capabilities for improving government.

The rest of the paper is structured as follows: in chapter 2 I present and discuss a collection of theoretical and empirical literature on corruption, e-government and development theories which leads to the presentation of the research hypotheses. Chapter 3 is a rigorous discussion on the methods and data employed in the empirical analysis. Both time-series methods and the operationalisation of the variables are taken into consideration, and I end the methods and data chapter by discussing the treatment of missing values. In chapter 4 I present the empirical models and describe the results. The discussion in chapter 5 is an analysis of the empirical models guided by the previous research, theoretical foundations and hypotheses. Finally, in chapter 6 in the conclusion, I summarise my findings and answer my research question in addition to presenting possibilities for future research.

2 Theory

In this chapter, I present the theoretical scaffolding for my thesis that e-government leads to lower corruption. Firstly, I present the consequences of political corruption on development to illustrate the real-world value of fighting corruption. Thereafter, I define political corruption relying on a broad swath of literature. Furthermore, I illustrate the logic of corrupt behaviour in a principal-agent framework, demonstrating how individuals are motivated to act corruptly. Having presented the definition of political corruption and the principal-agent framework, I discuss institutional and other macro-level determinants of corruption to illustrate what factors might determine the institutional controls on the individual's ability to be corrupt. In the following part of the theory chapter, I aim at illustrating what we understand as e-government, and to what extent the technology needed for e-government is distributed globally. This last point is often referred to as the digital divide. In the two last subchapters, I present a summary of the theory and previous research

2.1 Institutions, Inclusiveness and Corruption.

Globalisation involves the internationalisation of markets, increased competition between nations, new ICT and increased interconnectedness, but also the rising relevance and volatility of markets. These factors have led to increases in financial openness, trade, foreign direct investment, ICT capital investment & use in addition to migration and mobility of workers (Mills, 2008, p. 3). The global economy and political institutions are with increased globalisation experiencing rapid changes driven by a closer interconnectedness. In the decade

after the financial crisis in 2008, global political institutions and the globalised economy has become an even more controversial topic in many countries with increased scepticism towards international cooperation, migration and trade. The scepticism has manifested itself in the rise of populist parties, the comeback of the strong man and a feeling of being left behind for many people in both developed and developing nations. Proponents of globalisation have seen these factors as a powerful force for lifting millions out of poverty and contribute to development (ibid.). Critics have argued that globalisation does not lift all boats, but instead lift all yachts, allowing the rich to become richer while the rest is left struggling with taking part in the globalised economy and political life leading to a loss of perceived efficacy in may people's life (Buckman, 2008).

Industrialised countries and developing nations experience different effects of globalisation. Developed countries will to some extent experience deindustrialisation, weaker labour bargaining power and a shift toward higher pay for high-skilled labour and lower pay for low-skilled workers. Developing economies are hypothesised to experience industrialisation, capital inflows and new jobs, and increased wage levels for low-skilled workers and a wage reduction for high-skilled labour (Mills, 2008). While these effects in sum may have lifted millions out of poverty, the rapid changes and hypothesised increases in inequality within and between nations might cause political grievances, and a backlash against international cooperation. However, it has been argued that national (inclusive) institutions such as the educational systems, employment and industrial relations systems, welfare regimes can filter away some of the unwanted effects of globalisation, effectively slowing down the speed of changes, allowing societies to adapt (Mills, 2008, p. 6). With a rapidly changing and interdependent global economy and political system and autocratic values on the offensive, protest movements such as the Yellow Wests in France and Canada as well as the rise of populism, understanding the mechanisms that drive inclusive economic and political development is increasingly important.

As touched upon in the introduction, both policymakers and researchers seem to agree that institutions matter for development. While it is unclear what comes first, institutions or economic growth, I will for simplicity assume that institutions come first (Treisman, 2007). This allows us to study the role of institutions in development using the theory proposed by Acemoglu and Robinson (2012). Central to their theory is the idea of growth being a side-effect of inclusive institutions, while extractive institutions will hamper development. Inclusive institutions manage to secure private property, equality before the law and public services like education that provides a level playing field. Extractive institutions are therefore the absence of

inclusive institutions. How political institutions are put together defines the rules of the game and who gets to play it (North, 1990). Power can be distributed broadly and narrowly. Institutional power-sharing results in pluralistic institutions, while a narrow distribution of power is absolutist. Even if power is widely distributed, it is still necessary for the state to have a monopoly of legitimate violence, or else stability will not be achieved, and the political institutions can be characterised as extractive. Inclusive and extractive institutions are self-enforcing and create virtuous and vicious circles, respectively. Bueno de Mesquita and Smith (2012) demonstrate how these self-enforcing mechanisms are maintained. For both autocratic and democratic leaders maintaining power is done through forming a winning coalition that is as small as possible. No ruler rules alone and her essential supporters must be bought to stay loyal. In a country, most likely a democracy, with pluralistic power-sharing, the incumbent is dependent on a larger group of supporters and is incentivised to invest in public services to regain their trust come election time. In an absolutist political environment, the public is much less important, and the incumbent has no incentive in investing others than her cronies. The luckiest autocrats even have access to large supplies of natural resources or foreign aid, keeping themselves and their few supporter's fat and happy, while leaving the rest of the nation underdeveloped, enforcing resource curse mechanisms and other extractive institutions. An example of a typical extractive institution imposed by autocratic leaders is rampant corruption. Staying in power is expensive, and cronies are to be paid. When an autocratic leader comes to power, excessive taxing and other forms of plundering the public is essential for paying off essential supporters, leading to a vicious circle.

Corruption levels are, therefore, a good indication on the inclusiveness or extractiveness of a country's political and economic institutions. While globalisation is debated to have both positive and negative effects, it all seems to depend on the institutions within each country, and possibly neighbouring countries, therefore, serving as a good example of the importance of institutions. Corruption can be seen as the antithesis to core-values of inclusion and democracy such as: "public trust, responsive representation, impartial adjudication, mutual persuasion, generalised trust and reciprocity and voluntary exchange" (Warren, 2015, p. 43). Corruption is not also detrimental to development in autocracies, it can also lead to lowered trust in democracy, itself, its institutions and its political class (Della Porta & Vannucci, 1999, p. 10). Warren (2015) argues that this loss of trust will negatively impact people's impression of their own political efficacy, making them doubt the values of deliberation and public speech and their own personal impact through these institutions. This can, in turn, result in people

preferring to only care about their own "narrow domains of self-interest", substantially narrowing down the scope and impact of democracy (Warren, 2015, p. 45). The impact of voting is also weakened through "breaking the principle-agent relations essential to democratic representation" (Warren, 2015, p. 45). Della Porta and Vannucci (1999, p. 10) further argues that the corruption of civic norms can result in a snowball-effect where the non-corrupt must follow the corrupt rules of the game to obtain political power. This dynamic can consequently lead the political system to enter a vicious circle (ibid.). In sum, corruption will eventually erode inclusive values, changing them for extractive ones, making the state an apparatus for the special interests of the well-off and well-connected. Corruption will in due course create extractive institutions to the benefit of the few. In practice, officials will start conducting secret violations of contracts that work against the citizen's interest but in favour of a third party, such as a higher-level official to receive a reward that is bigger than what is reaped from honest work (Della Porta & Vannucci, 1999).

However, some early economic research on corruption suggested that it could promote growth and welfare, something Rose-Ackerman (2006) attributes to the economist's belief in prize mechanisms. Some economists would see bribery as a way of speeding up interactions with the government (Huntington, 1968)."Grease-the-wheels-of-bureaucracy" type of corruption is seen as a method for overcoming governmental red tape and increasing growth (Jain, 2001, p. 92). In contrast to this argument, Kaufmann and Wei (1999) find that companies who engage in bribery spend more time negotiating with bureaucrats resulting in higher costs. Greasing the wheels is not as much a reality as corruption being a way of putting "sand in the machine" (Jain, 2001, p. 92). In the end, the corrupt agents would have an incentive for increasing the price of bribes. Moreover, even if one interaction with government officials would become more efficient, the same officials could recognise opportunities for further profit by adding regulatory demands making even more frequent interaction with corrupt agents necessary. In total, we are left with a loss of efficiency (ibid.).

Researchers seem to mostly agree that corruption distorts institutional incentive systems meant to maintain growth, sustainable development, and impedes the success of the United Nations Sustainable Development Goals (Elbahnasawy, 2014, p. 114; United Nations, 2018). Global institutions concerned with development such as the World Bank and the UN has identified corruption as the single greatest obstacle to economic development (Elbahnasawy, 2014). Jain (2001, p. 92) sums up the argument that "corruption aids development in certain instances" as unlikely, stating that it "… does not merit much attention from serious scholars…". Introducing

unofficial price-mechanisms in the public sector also harm the poor proportionately harder than the rich. Some scholars suggest that corruption may be viewed as a regressive tax on the absolute poor, who would pay more significant fractions of their salary compared to the well-off (Todaro & Smith, 2015; Uslaner, 2015). Research finds that both the quality and quantity of public health care and educational services is worse in countries struggling with corruption (Jain, 2001, p. 92). Corruption, therefore, begets corruption and is highly persistent, something that is recognised by the international development community and policymakers in both democracies and autocracies alike. Even autocracies like China, which started a process of opening up their economy in 1978, recognised the need to keep their economic institutions inclusive and hold corrupt officials at bay to ensure some version of the democratic core values continue to drive development (Fukuyama, 2012[1992]; He, 2000). Institutions matter and combating corruption is necessary for sustainable inclusive development.

2.2 Defining Political Corruption

Moving towards a robust definition of our central concept of interest is undoubtedly essential in any scientific endeavour. Philip (2015, p. 17) argues that a definition of a word has "two dimensions: it may articulate the meaning and use of a word, and it can provide a tool in the construction of an explanation.". Both dimensions are important. However, I will emphasise the latter. Corruption is a many-faceted concept which needs to be defined rather stringently to be useful as a scientific tool. In the following paragraphs, I will discuss the concept to reach a set of criteria that will make us able to separate the concept of corruption from other phenomena (Philip, 2015, p. 17).

Historically corruption has been a topic of interest for many of the greatest political thinkers, and the way corruption has been defined has varied through time. Dobel (1999 [1978], p. 16) argues that in a classical conception, corruption represented a degeneration of the body politic in general: "for Machiavelli it was the destruction of citizens' virtues; for Montesquieu the perversion of a good political order into an evil one; for Rousseau, the inevitable consequence of the very struggle for power". Corruption was, therefore, something that happened to the whole of the citizenry. In modern times, morals are, at least in some parts of the world, secularised and the population, in general, ever more pluralised and it is argued that corruption revolves more around the individual citizen maintaining fairness of competition than the pursuit of moral goals for the whole of the polity (Della Porta & Vannucci, 1999, p. 16).

The concept of corruption is, therefore, changing over time and is dependent on societal factors, rather than solely on some ideal universal standards. Criterions for defining an act as corrupt therefore varies across societies and time periods. A relativist definition of corruption is therefore hard to define and operationalise: Leys (1965, p. 216) puts it like this: "an act is therefore presumably only corrupt if society condemns it as such, and if the doer is afflicted with a sense of guilt when he does it". This way of seeing corruption resonates with the first dimension of a definition, where the everyday use of a word provides its true meaning. However, such a relativist approach may render a concept useless in a cross-cultural and cross-temporal context (McMann et al., 2016).

Philip (2015) argues that a relativist definition will leave us with the unsatisfying situation where local definitions are not transferable to other contexts. On the other hand, we run the risk of oversimplifying the concept, leaving us only with a slightly shallow and technical understanding of corruption. A strict behaviourist approach leaves room only for a handful of obvious acts of corruption, such as bribery. Others opt for the generally accepted definition of political corruption as "the use of public office for private gain", which is not wrong, but can be too broad and can lead to a murky analysis. Broad definitions welcome subjective conceptions of corruption with little external validity, because it doesn't really define what this definition includes nor what it excludes. This leads me to follow McMann et al. (2016) and their discussion on the Varieties of Democracy Institute's corruption index. They base their definition of corruption on the common definition, "the use of public office for private gain". However, they are not "interested in capturing all instances where non-specialists, media, and pundits use the word corruption as a label for a behaviour of which they do not approve" and to achieve that goal, specifications are in order (McMann et al., 2016, p. 9). The Varieties of Democracy Institute's corruption index (V-Dem), includes mainly adhere to the general definition of corruption as "the use of public office for private gain", but it includes several types of corruption involving multiple branches of government such as the executive, legislative, judicial and the bureaucracy. In addition, it includes grand and petty corruption. The index also covers specifying bribes, undocumented extra payments, kickbacks, contracts for personal gain, future employment, theft, embezzlement and misappropriation. The catch-all term "material inducements" is also included (McMann et al., 2016, p. 9). So as opposed to settling for a general description of political corruption, McMann et al. (2016) explicitly strive to state which types of corruption are included in their definition.

V-Dem's approach includes a broad array of phenomena we can consider as political corruption. However, we need our definition to not only include behaviour we consider corrupt but also exclude other phenomena from our analysis, enabling us to concretise our criteria. McMann et al. (2016, p. 10) specify the agents and type of corruption they include in their definition. By strictly considering government officeholders, other public servants are not included. Therefore, a payment to a university admissions officer in return for admission is not included in this definition. Another example is the case where both public and private media outlets can receive bribes in favour of positive coverage. Forms of corruption where the private gain is not immediate and undeniable are also excluded. This can be exemplified by acts such as vote-buying or distribution of government jobs to a specific district to increase the chances of re-election for a political incumbent. Other behaviours like adultery are also excluded from this definition of corruption (McMann et al., 2016). In sum, this conception of corruption relates intimately to corruption in government understood most often by governments themselves and the academic community (McMann et al., 2016, p. 8). That is, it adheres to the broad definition of corruption as "the use of public office for private gain", but at the same time defining a variety of specific types of corruption included in this definition, something that is contributing to the definition's validity.

Defining corruption, a many-faceted concept, is an exercise in balancing a plethora of concerns. Political corruption is something "you recognize when you see it" but it is hard to define clearly. To begin with, I quoted Philip (2015) stating that defining a term for scientific purposes is a balance between treating the word as it is commonly used and defining it in stricter scientifically viable terms. In the following sections, I discussed briefly how the concept of corruption changes over time and space, leaving us with several caveats when using corruption as a tool for scientific analysis. By opting for the V-Dem definition of corruption, I would argue that it allows enough wiggle room for cultural and temporal differences, while at the same time being stringent enough to make a scientifically valid operationalisation possible.

2.3 Political Corruption as a Principal-Agent Problem

Understanding the general logic behind corruption demands a theoretical model. There is a multitude of ways to approach the theoretical discussion on corruption. Rothstein and Teorell (2015) mention some general theories of corruption, including explanations based on "bad norms" or "collective action theories". However, placing this text in the realm of political economy inspired by Della Porta and Vannucci (1999); Rose-Ackerman (1978, 2006), the agent-principal model is a natural approach. Firstly, we have to understand and analyse the

relationship existing between the agent, the person entrusted to a decision-making role, and the principal, on whose behalf the agent acts (Della Porta & Vannucci, 1999, p. 16).

The agent-principal relationship contains two central ingredients; conflicting objectives and decentralised information (Laffont & Martimort, 2002). Agents might have conflicting interests with the principal, making them deviate from the agreed upon objectives. The principal will most likely not have enough information needed to control the agent's behaviour. If this was the case, and the principal had complete information about the agent, the principal would be able to tailor a contract, assuming agents follow contracts, according to the agent's skill set and motivations. It is evident that the principal needs to find a way to align the agent's private incentives with institutional goals. This would minimise the costs inflicted upon the principal for not having complete information, nor any obvious way of obtaining it efficiently (Laffont & Martimort, 2002).

Della Porta and Vannucci (1999) schematically identify corruption when three elements occur simultaneously: 1) a secret violation of a contract that implicitly or explicitly defines the responsibility and the exercise of some discretionary power; 2) when an agent acts against the principal's interest in favour of a third party, from which he receives a reward; 3) lastly, in a political context, we treat the citizenry, or the state, as the principal. In other words: political power is entrusted to agents of the state to serve the citizenry which is then misused for personal gain or to serve a rent-seeking institution. Political corruption then thrives in the shade of information asymmetry, absence of accountability and when private wealth and public power overlap, resulting in gains for the few at the cost of the many (Rose-Ackerman, 1978, p. 6; 2006, p. xvii).

Preventing corruption is, therefore, about aligning personal and institutional goals, assuming these goals adhere to non-corrupt standard in an inclusive institution. Theories seldom lack critics, this one being no exception. Rothstein and Teorell (2015) problematise the agent-principal theory arguing that if agents are solely motivated by incentives, changing incentive-structures would be the definitive answer to eroding corrupt practices. Following this logic we would see that "When a society is constructed so that fear is larger than greed, things would go well", and corruption would vanish (Rothstein & Teorell, 2015, p. 89). Human interaction with formal institutional rules is, as we might expect, "often partially backed, supplemented, or contradicted by implicit rules such as taboos, customs, traditions, codes of conduct, routines, conventions, and so forth" (Voigt, 2013, p. 5). However, Laffont and Martimort (2002) point out that, of course, norms and other factors play a role in human interaction, but we should not

underestimate the importance of self-interest. Institutions need ways of minimising information asymmetry to prevent and make sure real accountability punish corrupt behaviour at an individual level, effectively minimising the importance of contextual factors such as norms (Kolstad & Wiig, 2009). As Hobbes (2000 [1651], p. 368) eloquently put it: "… covenants, without the sword, are but words…".

2.4 Contextual Determinants of Political Corruption

The agent-principal theory says something about how to analyse human behaviour in institutional settings. Implied in the theory is the possibility of designing institutions capable of curbing corruption. But what type of institutions are important, which are bad for lowering corruption levels? Acemoglu and Robinson (2012) argue for inclusive institutions. Rothstein and Teorell (2015) goes more into specifics at the organisational level and identify three "institutional causalities" of corruption: 1) type of democracy, 2) type of public administration and 3) size of government. Treisman (2007) looks at some of these institutional factors but adds to the debate a focus on the economic determinants of corruption, such as openness to trade and economic growth – both indicators of inclusiveness. I will start with having a look at the institutional factors, followed by a discussion of economic and other factors determining the climate for corruption.

2.4.1 Democratic institutions and accountability

Democracies theoretically present higher accountability costs for corrupt agents (McMann, Seim, Teorell, & Lindberg, 2017). McMann et al. (2017, p. 6) illustrate four types of accountability: 1) free and fair elections lets citizens reward or punish the incumbents for their performance in office, known as vertical accountability; 2) Legislative and judicial constrictions on the executive such as high courts able to monitor and sanction the executive results in vertical accountability; 3) An active and free civil society and media allowed through freedom of association and freedom of expression, allows society to perform informed choices in elections and provide checks and balances between institutions resulting in societal accountability. In sum, democratic accountability should contribute to the continuation and emergence of inclusive institutions.

Rothstein and Teorell (2015) as well as McMann et al. (2017), however, acknowledges that the connection between democracy and corruption is not linear. An increase in the measured level of democracy does not automatically decrease the levels of perceived corruption. The relationship is curvilinear - either U-shaped or J-shaped since the mere introduction of free and

fair elections presents corruption opportunities (McMann et al., 2017; Rothstein & Teorell, 2015). Although Treisman (2007) also finds that democracies are perceived to be less corrupt but emphasise that the causal direction might point both ways. Newer democracies seem to be more corrupt than autocracies, but some old democracies like Italy and Greece do have fairly high levels of perceived corruption suggesting a sort of self-enforcing mechanism in play. There are also regional differences; for example, northern Italy is much less corrupt than southern regions (Rothstein & Teorell, 2015). Moreover, different institutionalisations of democracy also have consequences for the level of corruption. Rothstein and Teorell (2015, p. 85) mention some variants of democracy stating that they can be "federal or unitary, presidential or parliamentary, multi-party or two-party, have strong or weak local government". Research shows conflicting results and concludes that democracy shouldn't be counted as a primary causal factor but combined with other variables in explaining corruption. They question whether the real importance lies not in the formal institutionalisation, but how it is *de facto* implemented (Rothstein & Teorell, 2015). McMann et al. (2017) further argue that it is not *democracy* per se, but several components of democracy that need to work together, something that arguably is evident even in one-party states like China where bureaucrats and political leadership have strived to curb corruption to preserve regime legitimacy (He, 2000).

2.4.2 Public administration types

Rothstein and Teorell (2015) state that the type of public administration in a country is of often attributed characteristics associated with either increases or decreases in levels of corruption. Especially the lack of "Weberianism" is pointed out as a source of corruption. However, Rothstein and Teorell (2015) also argue that the causal direction in this finding is not clear since it is almost "self-evident" that following the rules would decrease corruption. Other researchers argue that variables such as the relative pay of civil servants can have an impact on the levels of corruption. However, Dahlström, Lapuente, and Teorell (2012) find that a strictly meritocratic recruitment process is the essential feature for preventing corruption (Rothstein & Teorell, 2015). Charron, Dahlström, and Lapuente (2015) show that meritocracy correlates highly with more traditional measures of indicating the quality of governance. Two examples are the "Rule of Law" and the "Government Effectiveness" indicators from the World Governance Indicators (WGI). These respectively have a correlation coefficient of 0.77 and 0.72 (Charron et al., 2015, p. 15). A strict recruitment process adhering to the norm of impartiality could also be thought to have a strong signalling effect. Rothstein and Teorell (2015) mention that the same could be said of promoting gender equality. In a comprehensive

literature review, Treisman (2007) does, however, not find any evidence for the gender-equality-low corruption nexus. He argues that this effect is a result of elements in liberal democracy frequently observed in low-corruption countries. A different explanation based on modernisation theory might be that developed nations are less disposed towards valuing kinship, but might be more impersonal (Nam, 2018).

2.4.3 Government size and macro-economic factors

Some economists argue that the size of government is itself a source for corruption. The idea is that the size of government increases opportunities for corruption and rent-seeking (Rothstein & Teorell, 2015, p. 86). Empirical findings suggest that the opposite might be right. Fiscal sociology implies that this paradoxical result can be credited to the degree of ownership citizens feel towards government finances. One study based on interviews of people in Uganda show that citizens are keenly aware of the high corruption levels among the political elite and do take a moral stand against this, but that they do not usually act. The reason for this inaction towards corruptions is attributed to the low level of income- and direct taxes paid to the government (Rothstein & Teorell, 2015, p. 86). Rothstein and Teorell (2015) ask whether this lack of taxation also can drive the corruption behind regimes with other sources of income, such as natural resources and foreign aid. Research also assume international trade to have a mitigating effect on rent-seeking activities from bureaucrats. On the other side will international trade, according to Treisman (2007) increase market competition. This will eventually undermine monopolies and gradually "competition erodes profit margins and officials are driven out of the market" (Treisman, 2007, p. 236). Despite criticisms liberalism-driven economic globalisation is shown by some researchers to make countries follow the book and reduce domestic corruption partly through increasing trade, as described above, but also through external pressures to follow rules that allow countries to participate in trade blocks or aid programs (Nam, 2018). Economic freedom enhanced through economic liberalism is also hypothesised to increase demands for predictable trade (Nam, 2018). Through trade can thus function as a regional and even global driver for reducing corruption as policy learning can happen across borders. Nam (2018, p. 280) argue therefore that having good neighbours might even assist countries suffering from the *resource curse*. That is, countries suffering from extractive institutions due to resource rents allowing incumbents to take control of the state, keeping the outsiders "poor, ignorant and unorganized" through ,among other things, lack of inclusive resource management (Acemoglu & Robinson, 2012; Bueno de Mesquita & Smith, 2012, p. 91; Kolstad & Wiig, 2009; Nam, 2018).

Not only economic freedom and personal investment in the state through moderate taxation that finances public goods contribute to lowering corruption. Economic growth and wealth are essential for lowering incentives to act corruptly. Treisman (2007, p. 223) states that "By far the strongest and most consistent finding of the new empirical work is that lower perceived corruption correlates closely with higher economic development". Current levels of perceived corruption are highly correlated with economic levels of development almost 200 years ago. These results are robust to the inclusion of control variables such as democracy, trade, and inequality. Moreover, the variability of interest rates is hypothesised to increase corruption (Treisman, 2007). Interest rate volatility increases difficulties in monitoring government finances, leading to corruption (Elbahnasawy, 2014; Treisman, 2007).

Despite rigorous research on corruption and economic factors, the direction of the causal chain is still unclear. Treisman (2007, p. 225) sums up this debate by mentioning three schools: the ones who believe that good institutions drive economic development, those who believe that economic development drove the emergence of good institutions and the ones who believe that human capital caused both. Treisman (2007, p. 212) concludes that the measures of perceived corruption are lower in "economically developed, long-established liberal democracies, with a free and widely read press, a high share of women in government, and a history of openness to trade". Countries that are regarded as more corrupt are more dependent on natural resource exports, have "more intrusive business regulations and unpredictable inflation" (Treisman, 2007, p. 212).

2.5 Defining E-government and ICTs

In this chapter I am going to debate previous research on information and communication technology and the closely-knit concept of e-government. Understanding the meaning of the concepts involved is not a straightforward task, making a thorough review of the literature necessary. I will begin by defining information and communication technologies, followed by a general discussion on e-government as a concept and research field to create a general understanding of what researchers have emphasised in the past and how the field is currently understanding e-government.

2.5.1 Information and Communication Technologies

Clarifying the concept of ICTs is central to the discussion e-government. The concept of ICTs is used in a plethora of contexts. However, Zuppo (2012, p. 13) states that "the primary definition of information and communication technologies revolves around the devices and

infrastructures that facilitate the transfer of information through digital means". From a technical perspective, ICT is a complex infrastructure of computers, databases, protocols, and media. Software solutions such as groupware designed to make several individuals collaborate as efficiently as possible are also included in this definition. We are in sum talking about a collection of devices, software and a move towards greater integration between data- and communications technologies. In a lack of a better word, we call this combination of technologies ICTs (Baldersheim, Øgård, & Haug, 2008, p. 60). Doong and Ho (2012) provide a useful way of thinking about ICTs through the concept of GPTs - *general-purpose technologies*. GPTs are technologies that spread throughout the economy and have the potential for improving productivity in a plethora of sectors. This text is adhering to a broad definition of ICTs following the GPT concept, that is broad categories of technology we might not necessarily associate with the daily workings of government, but the easy access to these technologies is important for making e-government ever so slightly more sensible and practical.

2.5.2 E-government

Now I turn to the process of defining the concept of e-government, a concept closely related to ICTs. In similarity to defining corruption, the literature on e-government is many-faceted and finding clearly defined concepts is challenging. Bannister and Connolly (2015) argue that the study of e-government is a field of study where the absence of rigorous theories has caused a sense of angst among researchers. They are even discussing whether e-government studies can have the right to recognise itself as its own field (Bannister & Connolly, 2015, p. 10). However, Wirtz and Daiser (2018) points out that the field is relatively young, and has moved from being practitioner-oriented to a field attracting the interest from academics in related fields such as public management, political science, innovation studies and technology studies (Bannister & Connolly, 2015; Yildiz, 2012, p. 345). Baldersheim et al. (2008, p. 13) note that even though political scientists have been slow to develop technology centred theory, there exist exceptions. Already in the 1960s researchers such as David Easton and Karl W. Deutsch adopted concepts from the field of cybernetics as analytical tools. This strand of literature did, however, not evoke interest from the general research community (Baldersheim et al., 2008, p. 13).

Today the topic of e-government is gaining traction in the academic community. It is no longer possible to discuss themes such as "[…]government reform, globalization, governance, public policy, citizen participation, accountability, citizen orientedness, government service-quality, and many others without including technology/e-government into the equation" (Yildiz, 2012, p. 345). The potential of e-government is perceived to be immense (Wirtz & Daiser, 2018;

Yildiz, 2012). Dispersion of ICTs in government and society is believed to help create a networked structure of interconnectivity, aid in service delivery, efficiency and effectiveness, interactivity, decentralisation, transparency, and accountability. E-government has evolved to become a catchphrase for all of these functions (Yildiz, 2007, p. 650).

Even with increasing theorising of the field, researchers disagree on how to define e-government or how to spell it. Lee (2010, p. 229) sums up some of the troubles of explaining e-government clearly: "The concept of e-Government involves an abundant pool of organisational, managerial, and technological issues, not only because it is a new area but also because it is a complex phenomenon involving various stakeholders and technologies". Bannister and Connolly (2015) identified two definitions scholars generally adhere to. The first and narrowest considers e-government to be "the availability of government services online 24/7", others consider e-government to cover all use of ICT in government or at least in the formulation of policy and the delivery of services" (Bannister & Connolly, 2015, p. 9). Homburg (2008) argues that e-government is a strategic use of ICTs in the public sector to achieve a more democratic and efficient bureaucracy and service production.

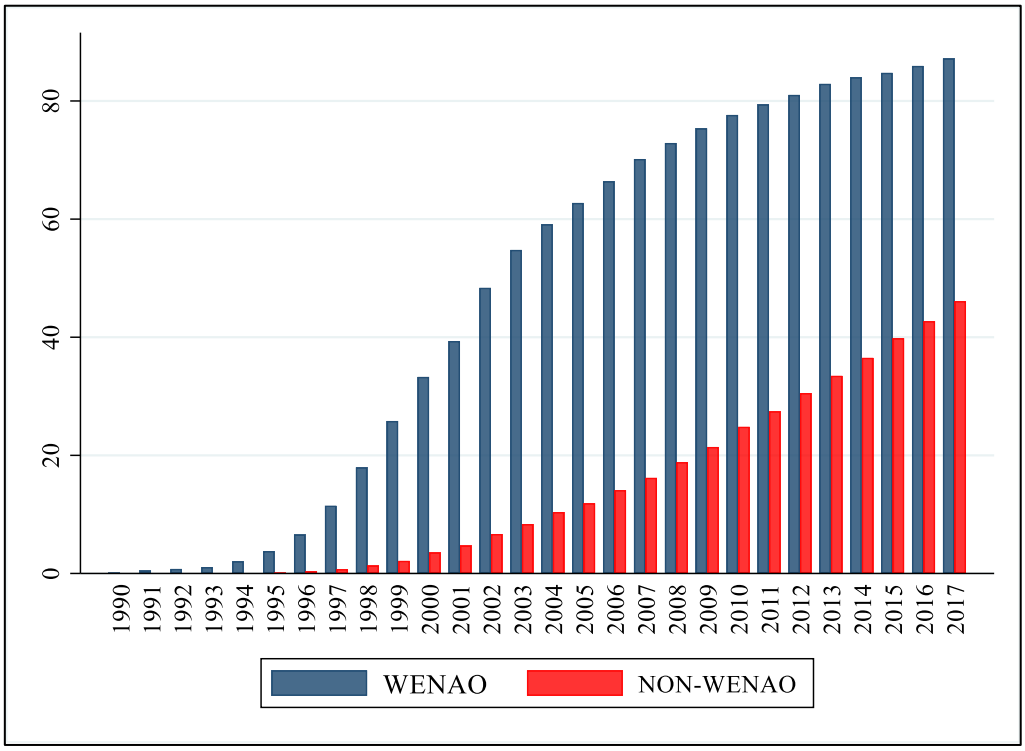
Amongst the unclear definitions and alleged theoretical absence of e-government, it is still possible to identify some concepts and theories. In their meta-analysis of e-government research, Wirtz and Daiser (2018) define stage-models/maturity models as an essential strand of e-government literature. These models seek to identify the development progress of ICTs in government and will usually consist of between three and six stages and vary between a descriptive and more academic nature (Bannister & Connolly, 2015, p. 7). Other typologies of e-government such as G2G (government-to-government), G2C (government-to-citizen) and G2B (government-to-business) are also helpful for identifying functions of e-government (Andersen et al., 2010; United Nations, 2018).

The complexity and reach of e-government might vary over time and space, further illustrating the fogginess of the concept, making it clear that e-government isn't easily definable. Yildiz (2007) argues that this is one of the inherent weaknesses of the idea. It is also a concept concerning ICTs, which are technologies that are changing rapidly. The United Nations (2018) has adapted to this reality in their biannual global e-government report, by slightly changing their definition to fit new developments, but the most basic definition is that e-government is "a tool for information and service provision" (United Nations, 2018, p. 220). Keeping these caveats in mind, I wish to approach the concept with a broad understanding of its implications for government, corruption, and democracy. I understand the idea of e-government as *a tool for*

developing policy and delivering services and information through the extensive use of ICT, (Bannister & Connolly, 2015, p. 9; United Nations, 2018). This broad definition is necessary as I aim at studying a cross-section of countries at the macro-level over a range of years, making it too costly to dive deeply into specifics.

2.5.3 Global ICT Dispersion and the Digital Divide

Figure 2.1 Annual global average of internet users in WENAO and non-WENAO countries[1]



As stylistically demonstrated in figure 2.3, the number of internet users globally has increased rapidly from the inception of the World Wide Web in the early 1990s. Approximately half of the global population is now online, but internet use is much more widespread in Western Europe, North America and Oceania (WENAO), as shown in the figure above. While some researchers attribute this spread of internet and related ICT to accelerated growth in the global economy (Doong & Ho, 2012), others like Milner (2006, p. 176) argues that political factors within countries also "exert a powerful influence". Chinn and Fairlie (2007) find that governance indicators are more important than trade openness. Misuraca, Rossel, and Glassey

[1] WENAO = Western Europe, North America, Oceania

(2010) point out that there is a considerable variation between democracies. Different ways of organising the state will potentially have a high impact on the success of e-government implementation. Misuraca et al. (2010) exemplify this by concluding in the Swiss context that, maybe contrary to intuition, a federalist state can be a strength when integrating e-government measures. The reason being that federalism is "…capable of simultaneously ensuring the homogeneity and diversity of how various facets of E-Government can be integrated" (Misuraca et al., 2010, p. 214). Moreover, Dalgaard, Bentzen, Selaya, and Andersen (2011) argue that new technologies might create both political and economic losers, leading incumbents and entrepreneurs to block the adoption. The adoption rate of new technologies is, therefore, likely endogenous and determined by both unevenly distributed economic and political capabilities – an argument which squares with the idea of inclusive institutions as innovative and adaptive.

Arguably more exogenous factors such as population size and density might also impact ICT dispersion. Andersen (2009, p. 203) argues according to Urban density theory that population size and density should impact the level of ICT dispersion since implementing the necessary infrastructure is a fixed cost meaning that the per capita cost is decreasing by the number of users, while the density reduces the required size of infrastructure projects. Furthermore, urban density should be associated with higher availability of skilled labour. Andersen (2009) summarises the argument with three points: 1) urban areas proved higher density of already existing ICT, 2) there is possibly a higher supply of ICT know-how and 3) learning effects and spill over. Elbahnasawy (2014) partially agrees with the urbanisation argument, arguing that a rural population will have less knowledge of the government bureaucracy and will, therefore, be more tolerant of an official's questionable behaviour. On the other side, in a rural society where "everybody knows everybody" it will cost more to be corrupt since the corrupted official will gain a reputation as a person who can "be bought".

Although ICT dispersion is gaining speed, it is not uniform around the world, which has led the term "digital divide" to catch on amongst researchers. The concept concerns the gap between the "haves" and the "have nots". As the access to ICT becomes more ubiquitous than before, the debate has moved towards discussing multiple divides (United Nations, 2018). Even if computers are cheaper and more available, the know-how and recourses needed to develop and maintain effective e-government solutions are not evenly distributed between nations and within nations. Dada (2006) warns that developing nations might not have the necessary infrastructure to implement e-government solutions from developed countries. By studying an

e-government project in rural India, Kumar and Best (2006) found that the project achieved several intended goals in addition to reducing corruption. However, factors such as changing commitment to the project from authorities and lack of qualified personnel ended in the failure of the project. One suggested solution is to build infrastructure for mobile-only solutions, as mobile communication technologies are getting cheaper, and the need for a cable-based network meant for technology, such as the land-line based telephone. However, the United Nations (2018, p. 35) recognises that when it comes to infrastructure "leapfrogging into mobile-only solutions" isn't as viable as initially thought, and many developing nations are still investing in necessary fixed-line technologies. Maybe ironically, investing in landlines is increasingly necessary with the introduction of more powerful 5G mobile networks, which requires new fibre cables (United Nations, 2018, p. 35).

Unfortunately, even investing in large infrastructure projects might present corruption opportunities (Charoensukmongkol & Moqbel, 2014). Officials in charge of public projects might collude with contractors to inflate the cost estimates and budgets through an inefficient and obscured bidding process. This collusion between public officials and contractors results in promises of kickbacks, future bribery, and gains for the ones involved in the conspiracy. The problems don't stop at after necessary infrastructure is constructed and e-government services are up and running. Dada (2006) fears that these e-government systems will continue to exclude more impoverished citizens with less access to both the technology and the necessary knowledge to benefit from it, implying that building the necessary infrastructure is far from enough to eliminate the many digital divides. There are concerns about levels of digital literacy, gender equality and accessibility for persons with different disabilities (United Nations, 2018). In sum, we are talking about a global digital divide between countries and a domestic digital divide within countries that impacts the outcome of e-government initiatives (Doong & Ho, 2012). Studying e-government implementation at an organisational level, Heeks (2003) estimates that only a minimal number of e-government projects in developing countries are successes (15%), while the majority are total failures (35%) or partial failures (50%). "eGovernment failures come in more varieties than Heinz", but Heeks (2003, p.5) have identified three archetypes: Hard-soft gaps, private-public gaps and country-context gaps. Hard-soft gaps widen when the strict rationality of e-government systems collide with soft factors: "people, politics, emotion and culture" (ibid.). Public-private gaps occur when IT firms and consultancies fail to identify the differences between the public and private sector realities and end up shoehorning square pegs in round holes. Country context gaps occur when ready-

to-use e-government systems are imposed on developing and transitional countries. The variety of contextual factors determining the outcome of an e-government project implying that e-government might not be a universal solution to corruption problems in public institutions.

Al-Ali (1991) illustrate this issue with a case on technology transfer from the UK to Kuwait. Bridging the digital divide is not only about selling ready-to-use solutions to developing and middle-income nations. Directly transferring technology to other contexts might make developing countries dependent on the suppliers through the lack of a sizeable and qualified local workforce. The absence of infrastructure and human recourses can lead to a situation where the potential of the new technology is not harnessed to its maximal capacity. Relying on a case study of Jordan, Dada (2006) fears that developing and middle-income nations are not ready to implement e-government systems, since only privileged segments of the population will be able to access services more efficiently, and the rest are left to bribing new intermediaries resulting in no improvements of democracy or as implied, corruption levels.

2.6 Previous Research on e-government and Corruption

Above, I mainly focused on defining the concepts of corruption, ICT, e-government and the digital divide. I have also dealt briefly with previous empirical research. The literature on e-government research demonstrate that researchers understand e-government as having an almost unprecedented potential for improving government, although the discussion on the digital divide provide the insight that countries have vastly different capabilities for succeeding with their e-government projects (Andersen et al., 2010; Bannister & Connolly, 2015; Ojha & Palvia, 2012; United Nations, 2018; Wirtz & Daiser, 2018; Yildiz, 2012). Below I will move on from discussing corruption and e-government separately to focus on how e-government impacts corruption. This chapter will start by examining case studies to demonstrate how e-government initiatives can work in practice. Subsequently, I review macro-level quantitative studies. Furthermore, I will look at how e-government might increase corruption and disinformation in low-democracy states before I review the literature on the direction of causality between e-government and corruption.

With regards to theorising, e-government research has been practitioner-oriented and mainly approached through stage-models and typologies such as G2B (Government-2-business) (Andersen et al., 2010; United Nations, 2018; Wirtz & Daiser, 2018). However, when studying e-government in a corruption context, researchers are often opting for clearly stated analytical frameworks, such as the principle-agent model (Ojha & Palvia, 2012). Others contribute to the

literature by examining the macro-level determinants of e-government succeeding as an anti-corruption tool (Nam, 2018). Methodically the research is diverse. Researchers such as Elbahnasawy (2014) and Wirtz and Daiser (2018) state a need for more quantitative analysis for e-government research in general as well as the e-government-corruption connection. I have also added a discussion on how e-government might work in low-democracy frameworks, an issue I argue is under-communicated in the e-government literature.

2.6.1 Micro-level research

Research on e-government and corruption can be split into two categories determined by their analytical levels: macro-level and micro-level research. What I identify as micro-level studies are generally case studies examining the implementation of e-government initiatives and provide some insight into the practical implementation process that macro-level studies can't provide.

Ojha and Palvia (2012) studied five e-government initiatives in India to demonstrate e-government in practice. Theoretically, the authors lean on three common theoretical frameworks in political economy: agency-theory, the economics of crime and transaction cost economics (TCE). Ojha and Palvia (2012) draw on five concepts derived from these theories: 1) probability of conviction (EoC), 2) information asymmetry (agency-theory), 3) disintermediation (agency-theory), 4) asset specificity (TCE) and 5) uncertainty (TCE) to construct an analytical framework.

According to agency-theory, decreasing the information symmetry between the agent and the principal is a central goal when combating corruption. Through lowering the costs of monitoring and removing the agent's discretion might be a way of mitigating the risk of corruption occurring (Kim, Kim, & Lee, 2009; Ojha & Palvia, 2012). By eliminating the agent's role through computerisation of the public sector, might be a solution to lowering monitoring cost and achieving what Ojha and Palvia (2012) call 'disintermediation'. Removing the agent's discretionary powers by making service delivery follow the rules embedded into software, is an example of partial disintermediation. Citizens being able to serve themselves is an example of total disintermediation. However, it can be thought that digital systems are not going to prevent corruption completely. Corrupt public servants might learn how to circumvent the systems leading to corruption growing (Bhatnagar, 2003; Elbahnasawy, 2014). Ojha and Palvia (2012) hypothesise that lowering (human) asset specificity might reduce corruption by making citizens and other principals dependent on an agent able to put itself in a veto player position. Reducing

uncertainty is hypothesised to reduce corruption by informing the citizen of his or her rights, procedures, etc. which in turn minimise bribe takers' opportunism. Related to reducing information asymmetry and uncertainty is increasing the probability of conviction through digitising bureaucratic procedures, making thorough audits by third-parties possible.

Ojha and Palvia (2012) conclude that in 3/5 reviewed projects, digitising services decrease corruption, while 2/5 partially succeeded. One example of success is the *Bhoomi* ('land') project, which involved the digitisation of 20 million records of land ownership (Bhatnagar, 2003). This G2B initiative drastically reduced bureaucratic corruption related to the transaction of property. Bhoomi thus increased the probability of conviction, reduced information asymmetry, reached partial disintermediation, reduced human asset specificity and reduced uncertainty. On the other hand, the computerisation of passport offices in India did not prove to be a success. While the application forms for a new passport was disintermediated and done online by the applicant, this paper had to be printed out and delivered to the resident's passport office. Only digitising a small part of the process, left the rest of the back-end process unchanged, resulting in minimal gains corruption-wise (Ojha & Palvia, 2012).

In a study on the anti-corruption measure OPEN (Online Procedures ENhancement for civil application) in the Seoul Metropolitan Government (SMG), Kim et al. (2009, p. 47) identified four anti-corruption strategies: "prevention, enforcement, access to information and empowerment, and capacity building". Prevention was enforced through standardising tasks, and minimising discretion, while enforcement was done through making real-time tracking of the internal processes possible. Furthermore, enforcement was enhanced by adding feedback mechanisms, and citizens can report unwanted conduct to the audit department or the mayor directly. Kim et al. (2009) conclude that the OPEN system reduced corruption, where it was implemented in the SMG. I would argue that there are striking similarities between the anti-corruption described by Kim et al. (2009) and Ojha and Palvia (2012). The OPEN system's efficiency was recognised even by international organisations and suggested implemented in developing countries (Kim et al., 2009, p. 44). However, as discussed above, the digital divide describes a situation where the prerequisites needed for e-government initiatives to succeed dependent on a variety of institutional and geographical contexts.

2.6.2 Macro-level research

Researchers have stated a need for more quantitative research on the relationship between e-government and corruption (Elbahnasawy, 2014; Wirtz & Daiser, 2018). Andersen (2009)

further argue that macro-level investigation could reveal general equilibrium effects; he fears that micro-level impact studies and "anecdotal evidence" can't account for the possibility that e-government initiatives in one sector could lead corruption to "migrate elsewhere in the economic system" (Andersen, 2009, p. 202). But as previously mentioned; the quantitative research reviewed here tend to lend less attention to "the mechanisms linking internet to corruption" (Dalgaard et al., 2011, p. 415). Several of the quantitative studies find that ICT-diffusion such as an increased number of internet users and e-government is associated with lower levels and changes in corruption (Andersen, 2009; Dalgaard et al., 2011; Elbahnasawy, 2014; Lio, Liu, & Ou, 2011; Nam, 2018). Studies vary both in sample size, units of interest and variables included. One type of study focuses mainly on the relationship between ICT or e-government and reduction in corruption (Andersen, 2009; Dalgaard et al., 2011; Lio et al., 2011). While others emphasise contextual factors that might have an impact on the efficiency of ICT and e-government on corruption prevention (Kock & Gaskins, 2014; Nam, 2018).

2.6.3 Contextual factors

Other studies focus on intermediating factors and contextual settings, in contrast, to solely studying the internet/e-government-corruption relationship. Kock and Gaskins (2014) rely on a panel data set of 24 Latin American and 23 Sub-Saharan African countries from 2006 to 2010 and path analysis to investigate the relationship between internet diffusion and corruption mediated by a measure of voice and accountability which resonates with the argument that transparency isn't sufficient, and real accountability is needed to prevent corruption (Kolstad & Wiig, 2009). They find that in Latin America and Sub-Saharan Africa, internet diffusion leads to less corruption if the effect is mediated through the effect of voice and accountability. When including the interaction term, the direct effect of internet diffusion on corruption. Kock and Gaskins (2014) also control for other variables such as GDP per capita, geographical region, and cultural dimensions. Elbahnasawy (2014) uses a large panel dataset consisting of 190 countries from 1995-2009. I find this study interesting because of the dataset's size and the variety of variables used. As a measure of e-government, Elbahnasawy (2014) uses the United Nations e-government development index (EGDI). He further estimates how the three subcomponents of the index, scope and quality of e-government services in addition to the telecommunication infrastructure index and the human capital index have an impact on reducing corruption, while the human capital index does not seem to have impact under any model specification. This finding contrasts the hypothesis by Kolstad and Wiig (2009) that education lets citizens utilise the available government information effectively. It should, of

course, be mentioned that Kolstad and Wiig (2009) focused on natural resource-rich countries. However, this finding also contradicts the worry about digital illiteracy presented in the debate on the Digital Divide (United Nations, 2018). A possible explanation might be that different types of ICT might demand different levels of competency (Chinn & Fairlie, 2007, p. 18). In addition, he finds that internet adoption and e-government implementation should be considered as compliments. That is, internet adoption drives the efficiency of e-government. Other factors such as the bureaucracy's willingness to adopt new technologies can also create barriers. The same goes for citizens who might prefer analogue ways of communicating with government services (Bertot, Jaeger, & Grimes, 2010). Another obstacle for e-government's anti-corruption capabilities, is that officials can potentially learn to adapt their corrupt behaviours to new e-government systems. Essentially reversing their intended use as anti-corruption tools (Bertot et al., 2010).

In contrast to Treisman (2007), Elbahnasawy (2014) doesn't find any evidence that inflation variation and trade restrictions worsen corruption. Nam (2018) also controls for contextual factors like economic and political development and cultural factors. Nam (2018) finds that cultural factors such as individualism positively impact the efficiency of e-government as a corruption deterrent.

While it seems like e-government and ICT leads to less corruption, other studies find that overinvesting in ICT infrastructure can lead to new corruption opportunities, as previously mentioned in the discussion on the digital divide (Charoensukmongkol & Moqbel, 2014). Charoensukmongkol and Moqbel (2014) find a u-shaped relationship between ICT investment and corruption. They suggest that this is a reason Kolstad and Wiig (2009) didn't find a significant relationship between transparency and reduction in corruption in resource-rich countries. However, Kolstad and Wiig (2009) point out that transparency is not sufficient, especially without the ability to process it (e.g. lack of education), equal access to it, and if there is an absence of incentives to act on information and absence of democratic accountability. Kolstad and Wiig (2009, p. 526) list other ways information might be manipulated to create an impression of real transparency, for example through unequal access, incomplete information and information overload, among others.

2.6.4 E-government in weak democracies

The problem of disinformation and information overload is one arguably under communicated areas of e-government-corruption research. While much of the early literature has been

characterised by optimism, an increasing swath of the literature contest this view instead regarding e-government and the internet "as enhancing isolation, threatening privacy, raising the dialogue to cacophonous levels, extending the possibility of the "big brother state" (Reece, 2006, p. 73). Mejias and Vokuev (2017) found that internet censorship and the spread of disinformation can explain some a potential lack of ICT and e-government's anti-corruption capabilities, citing several instances of governments using ICT to increase information asymmetry between the citizenry and government. For example, in February 2014 the Russian telecommunication agency, *Roskomnadzor*, was ordered by Russia's general prosecutor and his deputies to block web sites containing calling for demonstrations, a decision taken without any court verdict. In March the same year Russian authorities blocked web sites blocked three opposition web sites in addition to democracy and anti-corruption activist Alexei Navalny. While Russia has used SORM ("System for Operative Investigative Activities") devises to "collect, analyse and store all data that is transmitted or received on Russian networks, including calls, email, website visits and credit card transactions" since the late 1990s, they are not alone in the censorship and surveillance club (Mejias & Vokuev, 2017, p. 1032). In 1999 Malaysia and Saudi Arabia began censoring internet access as an official government policy. Turkey passed a law banning the airing of "pessimism" online in 2002 (Bertot et al., 2010, p. 267). In 2001 USA started restricting access and increased surveillance in federal libraries (Bertot et al., 2010). China began arresting people for what they wrote online the same year and further restricted internet access in 2002 by closing 92% of the internet cafes in the country. Since then, Chinese authorities have implemented measures to dictate information flows by requiring that firewall software is pre-sale installed in all computers sold in the country. The way e-government capabilities are employed is therefore widely different across countries. While China is inspired by western international organisations and advanced democracies in their e-government efforts, cultural differences are apparent. Recently Chinese authorities announced a project to make a social credit system by 2020, creating headlines like "China Wants to Tap Big Data To Build A Bigger Brother" (Stepan & Mokry, 2015).

Nevertheless, Stepan and Mokry (2015, p. 18) argue that the primary motivation of the Chinese government is mainly to prevent online fraud and "creditworthiness", and that "security bodies play a minor role, if any". Even if the G2C social credit system is intended as a benign implementation to reduce risk in online transactions, it is not hard to imagine how such a system might be misused. It is apparent that e-government as a corruption deterrent is most likely highly dependent on already existing institutions and to what degree these allow power-sharing.

Highly developed autocracies might exploit existing software and technology to further strangling civil society and cementing regime power, essentially turning the initially discussed principal-agent problem on its head. That is, in low-democracy countries, the public has lower power leverage on the state, and the government can treat the citizenry as the agents, and e-government solutions are made to protect the executives and other higher-ups in the political system.

2.6.5 The direction of causality in previous research

The direction of causality between e-government and corruption is unclear. As discussed in the subchapter *Global ICT dispersion and the Digital Divide*, there are good reasons to believe that causal direction is determined by endogenous factors. Andersen (2009) suggest that e-government led to reductions in corruption in non-OECD countries examining in the years between 1996 and 2006. He also controls for press freedom, population size, urban population, GDP per capita, phone lines per 1000 inhabitant and country fixed effects by way of a first-differences estimator. Dalgaard et al. (2011) find the same results examining the relationship between internet diffusion and corruption in U.S. states and a global country panel. By employing an instrumental variable estimation, they find that the causal relationship runs from internet diffusion to reduced corruption. In contrast to this finding, Lio et al. (2011) uses Granger causality and determine a bi-directional relationship between internet adoption and corruption in their sample of 70 countries between 1998 and 2005, Elbahnasawy (2014) likewise find a bi-directional relationship between internet adoption and corruption, in addition to finding that e-government unidirectionally reduces corruption, strengthening the idea that e-government and corruption are likely endogenous.

2.7 Hypotheses

2.7.1 E-government and corruption

This paper starts out with asking whether e-government leads to less political corruption. Through the theoretical discussion, we have studied the problem of corruption as a sign of extractive institutions to situate the problem in a development context. Furthermore, a principal-agent framework of corruption has provided an understanding of how e-government is supposed to prevent principals from indulging in corrupt practices on an individual level. In the review of previous studies, it is shown that researchers mostly find that e-government reducing corruption is plausible and to varying degrees confirm it. The qualitative literature seems to confirm that e-government reduces corruption by mitigating the principal-agent problem

between government and citizens, by increasing transparency and reducing the discretionary powers of civil servants. Several quantitative studies confirm this relationship, however, still no certain unidirectionality in the causal chain is proven. Based on these considerations, I formulate my hypothesis as:

H_1: E-government causes lower levels of corruption

2.7.2 Democracy and Institutions

In the discussion on the importance of institutions, I demonstrated how corruption is an important ingredient in keeping extractive institutions strong by obstructing development through corruption. Inclusive institutions are, on the contrary, characterised by a similar value set as democracies. The relationship between e-government and reduction of corruption does, therefore, not exist in a vacuum and is affected by contextual factors and institutions. According to agency-theory, minimising information asymmetry and real accountability are essential for keeping agents in check on the individual level (Kolstad & Wiig, 2009; Laffont & Martimort, 2002; Ojha & Palvia, 2012). On an institutional level, pluralistic power-sharing is needed to ensure inclusivity. Democracies provide incumbents with incentives to create good policies and institutions. I, therefore, expect that stronger democratic institutions will be associated with lower levels of political corruption. Furthermore, I assume democracy to be exponentially more efficient at preventing and stopping corruption, and I expect to find the same curvilinear relationship found by McMann et al. (2017). Incentives for keeping corruption levels high are also supposedly lowered when more people are granted real influence on who gets political power.

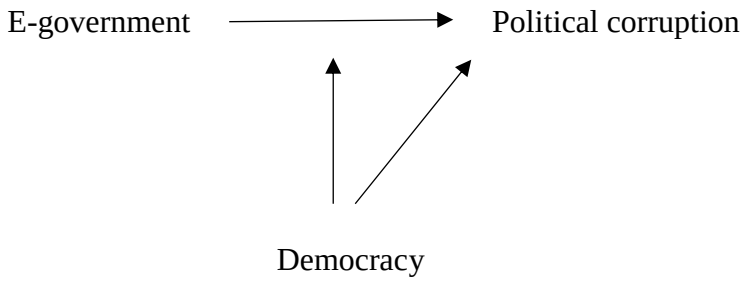


Figure 2.2 Illustration of the mediating effect of democracy on e-government.

According to Kock and Gaskins (2014), an interaction term between e-government, accountability (here represented by measures of democracy) and corruption control should be controlled for. That is, the effect of e-government is dependent on the democratic accountability. A mediating effect of democracy on corruption additionally squares with the discussion on inclusive institutions, suggesting that the anti-corruption effect of e-government is dependent on governments willingness to put checks on their power, and not use e-government to weaken citizen empowerment.

H_{2a}: Democracy causes lower levels of corruption

H_{2b}: Democracy has a mediating effect on e-government

H_{2c}: Democracy has a negative and curvilinear effect on corruption

2.7.3 Contextual factors

In trying to isolate the effects of e-government, accounting for well-established control variables should prove useful. To reiterate, the previous discussion on corruption determinants, Treisman (2007, p. 212) describe low-corruption countries as:"… economically developed, long-established liberal democracies, with a free and widely read press, a high share of women in government, and a history of openness to trade", in other words, countries with inclusive institutions. Moreover, countries perceived as corrupt to be more dependent on natural resource exports and have "more intrusive business regulations and unpredictable inflation", nations characterised by extractive institutions (Treisman, 2007, p. 212). Treisman (2007) especially highlights factors such as GDP per capita as a robust variable. The debate on the digital divide also reminds us that adopting e-government solutions is not free. A reasonable belief is that wealthier nations can afford to allocate more resources to implementing e-government infrastructure. Richer countries can also allocate more resources to other measures combating corruption. Trade can impact corruption levels through, for example, increasing competition and probably affecting business culture (Nam, 2018; Treisman, 2007). However, Elbahnasawy (2014) and Chinn and Fairlie (2007) do not find trade to be important.

According to the debate on the digital divide, we should also worry about factors such as digital illiteracy other domestic divides affecting preconditions for utilising ICT and e-government successfully (Kolstad & Wiig, 2009; United Nations, 2018). Elbahnasawy (2014) does this by

testing for the effect of Human Capital. Andersen (2009) argues that since constructing telecommunication infrastructure is a fixed cost; urban population density should be positively correlated with the availability of technology such as the internet as well as the supply of ICT know-how and spill over learning-effects. Elbahnasawy (2014) argues that a rural population will have less knowledge of the government bureaucracy and will therefore be more tolerant of an official's questionable behaviour. On the other side, in a rural society "everybody knows everybody", and being corrupt has a higher social cost. Depending on which factor counts the most, the control for urban population could result in either a positive or negative regression coefficient. However, according to theory, self-interest is a driving force in human behaviour, and real accountability presented through e-government solutions, not social pressure, should drive the results and I expect urban population density to reduce corruption.

Following the discussion above, I expect the contextual variables to cause low corruption and dependency on natural resource rents to cause high corruption levels. The effect of urban population is not clear, but I expect it to be negatively associated with corruption. While I have no hypotheses for the contextual variables, I expect increased trade, GDP per capita and urbanisation to cause lower corruption and natural resource rents to cause higher corruption.

3 Methods and Data

In this chapter, I will present the data structure required for time-series analysis and describe the special requirements for time-series analysis. Furthermore, a variety of estimation methods are described, including the fixed effect estimator and the Driscoll-Kraay standard errors. Following a general discussion on relevant estimators, I present the operationalisation of my variables and a discussion on the validity and reliability of the measures included, with particular emphasis on the dependent and primary independent variable of interest.

3.1 Data structure in time-series analysis

In this study I employ a time series cross sectional dataset (TSCS) containing approximately 162 countries with observations from 1990-2018, but most of the data in use cover the 15 years from 2003-2018. This depend on the variables in use. Since not all countries have observations for all years on all variables, the data is unbalanced – typical problem for panel-type data. How this is dealt with is discussed in further detail below. One reason why TSCS data is popular among researchers, is that we are often interested in investigating dynamic processes, something the repeated observations on the same units enables us to do. Over time concepts change, and we want to understand the development of these changes. "We are interested in the

fact that an actor ends up at point B from point A, but also the process by which they got there" (Pevehouse & Brozek, 2008, p. 456). TSCS methods, therefore, provides more potential for inferring causality (Wooldridge, 2013). Time-series data also allows for employing methods that makes it possible to control for unobserved variables thus account for individual heterogeneity (Wooldridge, 2013). A potential drawback of using time-series methods is that we get cluster structures in the data because the observations are not independent of each other, even if data is selected at random, in addition to other forms of serial correlation (Baltagi, 2005; Mehmetoglu & Jakobsen, 2017). As touched upon above, the issue of missing values can be problematic for estimation techniques, especially if the missing rate is high and not random (Mehmetoglu & Jakobsen, 2017). However, treated correctly and dependent on the question at hand, time-series data can arguably provide more statistically significant insights than other methods relying on single observations only providing snapshots of our empirical reality. Opting for TSCS thus answers the call for more generalizable macro-studies of e-government and corruption put forward by, for example, Elbahnasawy (2014) and Andersen (2009). I, therefore, have theoretical and methodical reasons for employing TSCS based methods.

3.1.1 Assumptions for linear regression

Inferential statistics rely on a set of assumptions, or caveats to having in mind when we try to draw conclusions from our data. This is necessary to follow so that we obtain unbiased estimations we can trust, based on the law of large numbers and the central limit theorem (Wooldridge, 2013). Following Wooldridge (2013), Mehmetoglu and Jakobsen (2017) in addition to Skog (2004) I will now turn to this issue.

Most authors rely on some version of the Gauss-Markov assumptions when discussing assumptions of the classical ordinary least squares (OLS) estimator. If you can argue that your model fulfils these assumptions, your model is BLUE. That is, it is the Best Linear Unbiased Estimator (Wooldridge, 2013). When this is true, you need not look further for a better estimator than OLS (Wooldridge, 2013, p. 98). However, in many cases, the model we build will not follow these assumptions, and other estimators are therefore closer to being BLUE. I will elaborate on that issue below. Skog (2004, p. 236) denotes the linear cross-sectional regression model as:

$$Y_i = b_0 + b_1 \cdot X_1 + \varepsilon_i$$

moreover, sums up the OLS assumptions[2] in three points: 1) linearity, 2) the error terms, ε_i are a) homoscedastic, b) normal and c) uncorrelated (absence of autocorrelation), and finally, 3) the independent variables and the error term are uncorrelated. Mehmetoglu and Jakobsen (2017) and Wooldridge (2013) adds to the list that there should be no perfect collinearity, while Wooldridge (2013) and Skog (2004) also highlights the problem of errors-in-variables. Another issue related to the errors is the case of outliers. The first assumption means that the relationship between the predictor variables and the dependent variable is linear. That is, one unit increase in X, equals one unit increase in Y. However, this assumption is flexible since "y and the independent variables can be arbitrary functions of the underlying variables of interest, such as natural logarithms and squares" (Wooldridge, 2013, p. 79). The error term should be homoscedastic, meaning that the error variation around the predicted value should be equal for all values of x. Achieving normality of the error terms is vital for achieving correct t-values and f-values, necessary for accurate hypothesis testing (Skog, 2004, p. 250; Wooldridge, 2013, p. 343). The absence of autocorrelation is in most time-series hard to avoid but can lead to spurious correlations. More extended time series will in addition to the autocorrelation problem, often encounter persistent trends causing problems with non-stationarity. That is, the mean and variance of a variable is not constant over time due to persistent trends in the underlying data-generating process. Trends, therefore, undermine the assumption that our variable is the result of some stochastic phenomenon (Wooldridge, 2013). In addition to temporal dependence, panel data can contain spatial dependence, or in this case, cross-country dependence. That is, neighbouring countries might assert influence on each other (Mehmetoglu & Jakobsen, 2017). Fulfilling these assumptions means that we can obtain the same variance formulas as in cross-sectional regression under random sampling, thus adding reliability to our statistical inference (Wooldridge, 2013, p. 362). The third assumption summarised by Skog (2004) is the absence of missing variables causing omitted variable bias. That is, there exist one or several variables Z explaining both X and Y. Z's effect will thus be included in the error term resulting in biased estimates (Skog, 2004). The problem of perfect multicollinearity means that the correlation between regressors can lead to inflated standard errors (Wooldridge, 2013, p. 508). Errors-in-variables or measurement errors can have different meanings for the model if the error is found in the dependent and or the independent variable. While measurement errors in the dependent

[2] These are not purely the original Gauss-Markov assumptions. However, these can be found in (Mehmetoglu & Jakobsen, 2017, p. 134) or Wooldridge (2013, p. 101)

variable do not result in biased estimates, errors in the independent variables will (Skog, 2004, p. 257).

3.1.2 Special data requirements for time-series

In contrast to a cross-section dataset, the total number of observations is determined by multiplying the number of units in the cross-section N by their observed time periods T. This means that TSCS methods are not as dependent on a large N as other panel data methods, because the number of observations will become sufficient as T increases (Beck, 2001). However, the number of observations in itself is not necessarily the greatest concern when using time-series data. Park (2011) distinguishes between *long* and *short* panel data. Long panel datasets have many time periods (large T) but few entities (small N), while short panels have many objects of observation (large N) and few time periods (small T). Being aware of this matter because a too small N can lead to Type I errors, while too large N can lead to type II errors by failing to reject a false null hypothesis (Park, 2011, p. 3). Thus, defining strict minimum limits of T and N is hard, but Beck (2001) argues that 10 time periods should be enough, although Pevehouse and Brozek (2008) warn that many time-series estimators are most stable when T is larger than 50, dependent on the estimator in use. My data include a maximum of 15 years, which can be considered a short panel, something that might be a cause for concern when choosing estimation procedures.

3.2 Estimation methods

Estimation methods have different strengths and weaknesses, dependent on factors such as data structure and which assumption the data breaches. Panel data presents special considerations, which are usually solved using estimation techniques such as ordinary least squares (OLS) or feasible generalised least squares (FGLS). As mentioned above, many time-series estimators are most stable when T is greater than 50, while others argue that a T of 10 is enough, but this should be taken with a grain of salt. As a way of checking my model's robustness, I opt for a range of estimators and standard errors, in an attempt to make up for shortcomings in data and breaches of assumptions. King and Roberts (2015), however, argue that "robust standard errors" only conceals specification issues. After all, it has been said that "all models are wrong; the practical question is how wrong do they have to be not to be useful" (Bell & Jones, 2015, p. 143; Box & Draper, 1987, p. 87). For my analysis, I will, therefore, run regressions with a range of different estimation techniques based on Pooled OLS and FGLS. OLS based methods include regressions using the Newey-West (NW) and Driscoll-Kraay (DK) standard errors as well as

Instrumental Variable regression and Fixed effects. I will also run feasible general least squares (FGLS) based estimations. I will now turn to a discussion on the different estimators included in my analysis where I present the rationales for including these.

The pooled OLS estimator is a widely used workhorse of regression analysis. OLS is the estimator of choice when your data complies with the modified Gauss-Markov assumption listed above (Wooldridge, 2013). However, the repeated observations on the same units in the panel data leads to the dependency between observations. Often, this leads to autocorrelation, which leads to biased standard errors and thus inflated t- and f-statistics (Mehmetoglu & Jakobsen, 2017). Inflated test statistics can lead to an increased risk of false positive findings, type I errors (Mehmetoglu & Jakobsen, 2017). Autocorrelation can also lead to a correlation between the errors and the X-variables which can lead to, among other things, heteroscedasticity, which results in unprecise estimates (Mehmetoglu & Jakobsen, 2017). A frequently used answer to the autocorrelation and heteroscedasticity problems faced by the OLS estimator is the Newey-West standard errors (Petersen, 2008). By employing a lag up to some length (for example one year), the Newey-West estimator accounts for autocorrelation in the error term.

Not only the error-term is lagged in many TSCS models. It is normal to lag the dependent and independent variables to account for historical factors impacting the estimated values of Y. Intuitively, the effect of x should also come before Y in time. An additional reason is that adding a lagged dependent variable (LDV) is that it can account for trending series containing a unit root (Mehmetoglu & Jakobsen, 2017). However, adding an LDV can make the model unstable when the data has a small T (<20) and suppress the explanatory power of other independent variables (Achen, 2000; Beck & Katz, 2011)

While the Newey-West standard errors do take into account autocorrelation, the estimator does not control for spatial dependency. That is, proximate countries could influence each other through common cultures, policies, and so forth. Even in microeconomic panels with units selected at random, herd behaviour and norms can cause these types of dependency between units (Hoechle, 2007). With Driscoll and Kraay standard errors, it is possible to control for such serial dependency (Hoechle, 2007). Hoechle (2007) argues that standard errors in Newey-West estimations will be downward biased when the spatial dependency is not accounted for. Another strength of the DK estimator is robustness when used on short panels. That is, in panels where N is larger than T, the Driscoll and Kraay approach “eliminates the deficiencies of other large-

T -consistent covariance matrix estimators …" (Hoechle, 2007, p. 284). Theoretically, the Driscoll-Kraay method, therefore, fits my short panel well.

When we suspect that our data include unobserved individual heterogeneity that is constant over time, we can apply the fixed effect estimator. To arrive at this estimator, we need time-demeaned data, this process is also known as the within transformation. We are now left with time-varying variables only, allowing us to investigate the time-varying X variables on the time-varying Y variable while excluding unit-specific characteristics which can cause bias in the estimates of a standard pooled OLS (Wooldridge, 2013). We can also include time fixed effects by including year-dummies in the regression controlling for unobserved effects that vary over time, such as the financial climate (Mehmetoglu & Jakobsen, 2017, p. 250). Because of the fixed effects ability to control for unobserved variables, it is generally thought of as a convincing tool for estimating *ceteris paribus* effects and making causal claims (Wooldridge, 2013, p. 477). However, due to the relative inefficiency of the FE estimator, it can have some trouble picking up significant effects from slow-changing variables, such as democracy, resulting in an increased risk of committing a type II error (Beck, 2001; Wilson & Butler, 2007). When T is relatively small compared to N, the FE estimator can produce "seriously biased coefficients" (Nickell, 1981, p. 1418). One possible solution would be to test subsets of N to decrease the size of N relative to T. However; my T is not very large either with a maximum of 15 years. If the covariance between our error and the independent variable is assumed small or non-existing, the random effects model (RE) is more efficient (Mehmetoglu & Jakobsen, 2017). RE also allows us to model in time-invariant factors such as culture, gender or whether a country is located in the global south (Bell & Jones, 2015). The RE model is harder to interpret substantially since the coefficients include both the within and between effects (Mehmetoglu & Jakobsen, 2017; Park, 2011). When estimating random effects models in Stata, the standard estimation technique is not OLS but FGLS. In their widely cited article, Beck and Katz (1995) argue that FGLS should be avoided when the number of time periods is smaller than the number of units (T<N).

A consequence of using FGLS on short panels is hugely underestimated standard errors (Beck & Katz, 1995). Because of my datasets limited time periods, I will only use the FGLS estimation techniques as robustness checks. Should we then choose the FE or RE estimator? Although the case where the covariance between the unobserved effect and the regressors is 0 is to be regarded as "the exception rather than the rule" (Wooldridge, 2013, p. 478). We can get a suggestion of which model to use by using a Hausman test. However, the question if there is a

dependency between the errors and the independent variables, is a theoretical one (Wooldridge, 2013, p. 478).

The last estimation technique I wish to include is an Instrumental Variable estimation (IV estimation). When facing the problem of omitted variable bias (unobserved heterogeneity), errors-in-variables and endogeneity, we have four choices: 1) we can ignore the problem, 2) we can find a proxy for the unobserved variable, 3) we can assume that the omitted variable does not change over time, and use the fixed effect estimator to rule this out or 4) we can recognise the problem by introducing a fitting instrument (Wooldridge, 2013, p. 491). Invoking a strong instrument allow us to remove some bias from our estimator and infer some statement of causality. It should, however, be mentioned that the IV estimation can contain fixed effect estimations. Wooldridge (2013) presents two requirements for an instrument to be valid: instrumental exogeneity and instrument relevance. Exogeneity involves the instrument being uncorrelated with the error term. Relevance is achieved when the instrument correlates with x and can be said to be "relevant for explaining variation in x" (Wooldridge, 2013, p. 492). IV regression should, however, be used with caution. Evoking weak instruments, that is they do not comply with at least one of the two requirements, might result in estimates that contain more bias than we would get using pooled OLS.

Replicating Dalgaard et al. (2011), I apply lightning strikes as an instrument for internet users. I also use the proposed instrument on e-government, since they theoretically should pick up some of the same effects. The IV regression will consist of a simple replication of one of the models done by Dalgaard et al. (2011) using their data material. I deviate from their approach by using annual data and not differenced data. Following the simple models checking the validity of the instrument, I run two full models with all controls and time-fixed effects included.

3.2.1 Granger causality

In testing the unclear causal direction between e-government and corruption, I employ a Granger causality test. The null hypothesis is that e-government does not *Granger cause* corruption or vice versa. I follow Elbahnasawy (2014) in checking for several lag lengths, although Wooldridge (2013) state that when using annual data, a lag length of around two years is often enough. In practice, I do this test manually by running OLS regressions with the Driscoll-Kraay standard errors experimenting with lag lengths and alternate between using the EGDI and Corruption as the dependent variable. In postestimation, I employ the f-test to test

the direction of causality or non-causality. This test, therefore, is not testing whether x is both sufficient and necessary for y, but it provides an idea of the sequential relationship between the variables of interest. This is, therefore, not a test of true causality and should be interpreted with caution. A classic example is the rooster *Granger causing* the sunrise.

3.3 Operationalisation of variables, validity and reliability

In this section, I am going to discuss the operationalisation of factors mentioned in the theoretical discussion and subsequently in the review of previous research but most importantly, in my listing of the hypotheses. I will explain the strengths and weaknesses of the variables in question; this includes a discussion on validity and reliability concerns. Although I mention a plethora of possible predictors of corruption, I follow Achen (2005) in choosing to keep the number of variable lows. Achen (2005) argue that regression estimates are not getting more precise by adding additional variables. Good models are usually built relying on sound theory. Mehmetoglu and Jakobsen (2017) and Skog (2004) support this by arguing that it is not the lack of variables that causes problems, rather the omittance of the right variables, although the "right" variables can sometimes be many.

The rest of this section is organised as follows: I will start by discussing the dependent variable and its reliability and validity by following McMann et al. (2016). A thorough discussion of the dependent and independent variable is warranted when measuring arguably soft concepts such as corruption and e-government. Through this discussion, I will present points to be aware of when assessing a dataset and its strengths and weaknesses. Following my discussion on the dependent variable, I will present my main independent variable, followed by a listing of relevant controls. By the end of this subchapter, the reader should be informed about the quality of my data and the rationale behind the variables chosen to test my hypotheses.

3.3.1 Dependent variable – The political corruption index

My dependent variable is the Political corruption index constructed by the V-Dem institute found in the V-Dem dataset v9 from 2019 (Michael Coppedge et al., 2019). The index measures corruption in an interval between 0, low corruption, and 1, high corruption. Unlike many other measures of corruption, the V-Dem measure includes different areas of the "polity realm" by including six different types of corruption over three branches of government (Michael Coppedge et al., 2019). That is, the executive, legislative and judicial. Finally, the measure differentiates between the highest levels of politics on the one hand and the public sector in general at the other. By separating these two levels of politics, the index enables the analyst to

measure both "petty" and "grand" corruption, which includes both bribery and theft, in addition to corruption influencing policy making and implementation.

McMann et al. (2016) have thoroughly evaluated the validity and reliability of their corruption index, and they "… define validity as the absence of systematic measurement error, and reliability as the absence of unsystematic (or random) measurement error." (McMann et al., 2016, p. 4). The authors propose three tools assessing the dataset's: 1) content validity, 2) data generation and 3) convergent validity. The first step involves evaluating the measure's ability to capture relevant components of the theoretical construct and contains several subcomponents: a) resonance/face validity, b) domain, c) differentiation, d) fecundity, e) consistency of the measure, and f) causal utility. I will now turn to these concepts.

Resonance or face validity of the concept in question should function as a good translation of the underlying idea. In this case, McMann et al. (2016) aim at measuring the common definition of corruption which is "the use of public office for private gain" by including several types of corruption, exemplified in the paragraph above. When evaluating the domain, the researcher wants to know whether the definition of the concept captures the meaning for the relevant audiences. As discussed in the previous chapter, this definition of corruption is aimed at an academic audience and policymakers, not the public's understanding of corruption, keeping subjective moral evaluations to a minimum.

Differentiation strengthens the definition by including relevant measures, but also excluding irrelevant ones. McMann et al. (2016, p. 9) for example argue for the strength of their measure by comparing the V-Dem indexes against other corruption measures, and among other factors, how they define corruption. In short, their main point is that others, only measure "a very narrow slice of the universe of all corruption forms falling under the definition, 'the use of public office for private gain'". McMann et al. (2016) illustrate this by mentioning, among others, Transparency International's Corruption Perception Index and the World Bank's Business Environment and Enterprise Performance Survey (BEEPS) which do include measures of public sector corruption but omit measures of executive, legislative and judicial corruption. Another example is the World Government Indicators which include electoral corruption something that is arguably not directly associated with personal gain.

The coherency of the corruption index is tested by a factor analysis of the six subcomponents of the index. Results show that each of the concepts probably measures the same underlying concept of "public office for private gain" by only loading on one unique factor, thus increasing

fecundity by reducing the complexity of the concept to a coherent and simple dimension. Furthermore, McMann et al. (2016, p. 11) argue that the V-Dem indicators are consistent by including "numerous sufficient attributes depicting corruption", making the index more applicable across time and space. In sum, these considerations improve the causal utility of the measure.

The second tool for evaluating a dataset's validity is studying its data generation. Who funds the dataset? Is it a non-governmental organisation (NGO), a multilateral bank or a university? Who is surveyed? Is it businesses, the general public or an expert survey? V-Dem is an academic venture led by scholars from leading universities and is funded mostly by northern European countries. V-Dem data is also collected by their own experts, and this avoids problems of aggregating potential biases following from opting for a poll-of-polls method of collecting data. Other datasets relying on secondary worth mentioning is the world government indicators (WGI) from the World Bank or the Corruption Perceptions Index (CPI) from Transparency International. Most of the barometers rely on their own sources. V-Dem argues that using experts might reduce biases: citizens interact with only certain types of public officials, and consolidated democracies with stable institutions will have corruption which is removed from citizen lives (McMann et al., 2016). For each question-year observation, V-Dem uses five experts of a total of 2000 experts assisting in data collection, 3/5 of these are nationals or residents.

Aggregation of the index is based on several steps, including Item response theory (IRT) methods, the previously mentioned factor analysis and finally, averaging. The measure of executive corruption (v2x_execorr) is made by combining the latent factor scores of the measures of executive bribery (v2exebribe) and the measure of executive embezzlement (v2exembez). The measure of public corruption (v2x_pubcorr) is made similarly based on the measures of public sector bribery (v2excrptps) and embezzlement (v2exthftps). The corruption index (v2x_corr) is made by averaging the executive corruption index (v2x_execorr), the public sector corruption index (v2x_pubcorr), the indicator for legislative corruption (v2lgcrrpt), and the indicator for judicial corruption (v2jucorrdc). For countries without a legislature, the values are imputed using the average of the remaining measures.

Furthermore, McMann et al. (2016) discuss the geographic and temporal coverage of their dataset. V-Dem has avoided country selection bias by including every country. Since experts are asked the same questions for every year surveyed (from 1900-2018), the data is suitable for time-series analysis. This contrasts several other ways of measuring corruption, for example,

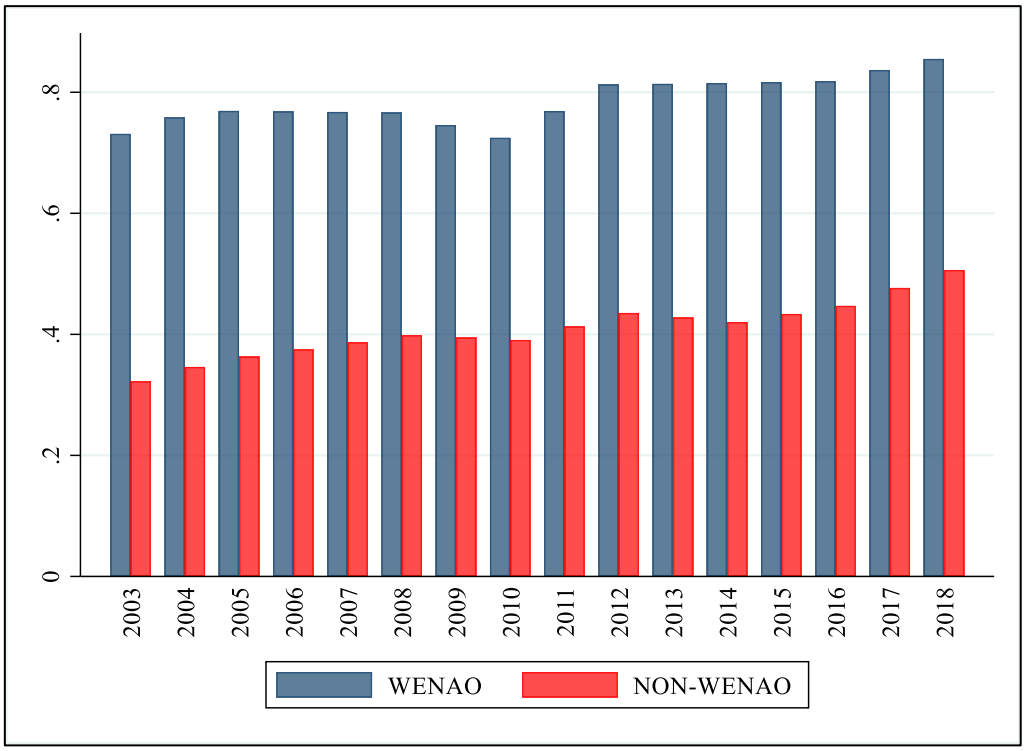
government audits which will be limited in both a temporal and geographical sense. Kraay and Kaufmann (2002) find that half of the variance in the WGI dataset is due to changes in the sources and coding rules (McMann et al., 2016).

The last tool is assessing convergent validity. That is, how well your dataset measures shared realities. Approaches can be both quantitative and qualitative. McMann et al. (2016) find that their measures are comparative to reviewed case studies, also for historical data. However, it is a theoretical possibility that experts are relying on the same qualitative studies for evaluating past years. V-Dem data generally correlates highly with other measures. However, V-Dem coders seem to rate countries less corrupt than the WGI. Exceptions are for example, Qatar, which V-Dem rates as more corrupt than WGI, while Lithuania is rated as less corrupt than in the WGI. Deviations like these might be attributed to the fact that McMann et al. (2016) find a bias toward liberal democracies among their coders, in addition to a tendency among female coders to rate countries as more corrupt. However, these are some of the few biases found by the authors.

3.3.2 Main independent variable – The e-government development index

For the operationalisation of e-government, I rely on the United Nations E-government survey and their biannual E-government development index, also known as the EGDI. The index strives to capture "the use of information and communications technologies (ICTs) to transform the public sector by enhancing its efficiency, effectiveness, accountability, inclusiveness, trustworthiness and supporting people's participation and engagement" (United Nations, 2018, p. 270). It is based on a survey of all 192 member states and ranges from zero to one, where higher scores represent better e-government relative to other nations and are not meant as an absolute measure of e-government quality. The survey is generally conducted biannually and started in 2001. However, the available data ranges from 2003-2018 and has seen an average increase globally since its inception. I now turn to discuss the validity of this measure applying the V-Dem evaluative framework.

Figure 3.1 Annual average score of the e-government development index (EGDI) for WENAO and NON-WENAO countries[3]



As shown in figure 3.1, the e-government index is steadily increasing over the time period with the highly advanced WENAO countries nearing the maximum EGDI value of 1. Still, it is worth noticing that non-WENAO countries are lagging, even if the trend is indicating that those nations are catching up. The dip in the EGDI scores, especially for the richer countries after 2008 indicate that maintaining advanced e-government systems is something governments must find room for in their already tight budgets.

The index is composed of the weighted average of normalised scores of three components of e-government: "(i) the scope and quality of online services quantified as the Online Service Index (OSI); (ii) the status of the development of telecommunication infrastructure or the Telecommunication Infrastructure Index (TII); and (iii) the inherent human capital or the Human Capital Index" (United Nations, 2018, p. 199). Before composing these measurements into one index, they are Z-score standardised. Standardising each measure before combining them, results in an indicator where " 'equal weights' truly means 'equal importance'" (United Nations, 2018, p. 199).

[3] WENAO = Western Europe, North America, Oceania

Following McMann et al. (2016), I will now discuss this index based on the three tools presented in the discussion above, namely: 1) the assessment of the content validity, 2) data generation and 3) convergent validity, beginning with the first tool and its subcomponents listed earlier when discussing the dependent variable.

As previously demonstrated, I found no universally agreed upon and simple definition of e-governance. This makes evaluating concept resonance or face validity a daunting task. Definitions have also evolved in the UN e-government surveys since the first edition from 2001. In the first e-government benchmarking, e-governance was defined broadly as "a tool for information and service provision", but the definition has been slightly altered since then (United Nations, 2018, p. 220). Definitional changes have sharpened the distinction between the general use of ICT by governments and the delivery of public services on the web, but have not affected the data collection methodology (Roche, 2017, p. 10). If we compare the subcomponents included in the index, that is the Online Service Index (OSI), the Telecommunication Infrastructure Index (TII) and the Human Capital Index (HCI), with these definitions, there is a continuous conceptual connection between the definitions above and the three measures. Although it is challenging to find any explicit theoretical reasoning behind combining precisely these measures in the e-government reports, Whitmore (2012) applies Bayesian factor analysis and find that the three composite measures better capture the underlying concept of E-government when combined, than separately.

For a concept and its measure to be useful, it should be deemed meaningful by its intended audience. The EGDI is intended for policy makers, academia, civil society, private sector and other actors with interest for e-government and ICTs for development. In other words, a broad audience. By analysis of citations, Roche (2017, p. 21) suggests that the e-government index has "steadily increased its influence world-wide". He further argues that "Discussion of e-Government as an area of important study … began in the 1960s. What is surprising is that the amount of scientific activity concerning e-Government started to increase rapidly from approximately the same time the United Nations introduced the index" (Roche, 2017, p. 21). A recent survey also shows that the index is considered "very useful" by approximately 75% of the member states (Roche, 2017, p. 22).

The concept of e-government is somewhat unclear and contributes to the difficulty of trusting the validity of the e-government measure. Uncertainties regarding what the EGDI is measuring is enhanced by the secrecy surrounding its survey questions, which is only partly public in the newest e-government report and mainly secret in earlier editions (Roche, 2017). However, of

the questions revealed in the 2018 report, it is possible to get an impression on what survey respondents consider essential. Examples are ranging from "Ability to apply for business licenses online" to "Existence of support for all languages" or "Information about gender equality" (United Nations, 2018, pp. 206-207). In sum, it seems like the quality of online government services is gauged through an evaluation of front-end services. Bannister (2007) criticise e-government benchmarks for not considering the back-end infrastructure, which also is critical in considering the scope and quality of e-government. However, the EGDI and its subcomponent the OSI, does not claim to measure much more than front-end user-centred services, but survey questions are partially kept secret. I would also argue that following the discussion above, the EGDI is sufficiently good at simplifying the murkiness of e-government, although stronger theoretical foundations are needed. The increasing popularity of the measure show that researchers and policy makers with high stakes in developing quality e-government services seem to trust the measure, thus increasing fecundity (Roche, 2017).

Bannister (2007) argue that in comparison with other more commercial measures of e-government, the EGDI is the most comparable over time and space. Roche (2017, p. 21) agrees stating that the "inherent methodology of the Index has shown over the years to be flexible enough to be modified so as to take into consideration how the world (technology; government policies; services) has changed, but in a way that does not destroy its year-to-year comparability for those nations making extensive use of it". The index has also been designed to make it universally applicable. That is, the survey material is designed in such a way that scores are not given with "state-of-the-art" technology in mind. The reason being that there is a vast gap in ICT utilisation between countries, and rating all countries compared to the most developed e-government systems, would result in too many nations receiving a score of 0.

The data used in the EGDI is a combination of original data and data gathered from other sources. Gathering data for the OSI 2018 edition is, for example done by 206 online United Nations Volunteer researchers from 89 countries covering a total of 66 languages. Volunteers are trained rigorously by e-government and online service delivery experts and guided by data team coordinators. Each country's e-services portal, e-participation portal is evaluated by at least two researchers. Subsequently, the data is transferred to data team reviewers who sends it to a senior reviewer (United Nations, 2018). Reviewers are also not nationals in an attempt at avoiding bias in the data collection process (Roche, 2017). This extra step with external reviewers is not done in the V-Dem data collection process, but the process is done by experts who are nationals or residents in the country under scrutiny. One weakness with the OSI is that

it does not include data, or at least not to my knowledge, about user satisfaction or knowledge about the services it is supposed to measure. E-government's "dirty little secret" is that we do not know much about the extent of how much citizens use these "shiny new tools" (Bannister, 2007, p. 181).

For the Human Capital Index, the data is sourced from the UNESCO-UIS or UNDP (United Nations, 2019). From 2014 the index consists not only of two variables but four. Past years have included adult literacy and gross enrolment ration, while after 2014, the two new components of expected years of schooling and mean years of schooling was added (United Nations, 2018, p. 204).

Data for the Telecommunication Infrastructure Index is provided by the International Telecommunication Union (ITU) (United Nations, 2019). Conceptually, this index has been unchanged since 2002. Since ICT changes rapidly, the index has also adapted to new realities. Three components have followed the index from the start: internet users, mobile phone subscriptions and fixed-telephone subscriptions have been used in the past. However, changes have been made. For example, in 2005, a measure of the online population was changed with fixed-broadband subscriptions. Another change was made in 2005 when the number of television-sets was dropped. Furthermore, the United Nations (2018) argues that there has been an improvement in data quality and coverage, which has led to fewer gaps in survey data. However, it is not clear whether they have found any way of improving earlier versions of the data.

While measuring something as the information society and specifically e-government, is inherently questionable (Bannister, 2007). Both limitations in time and resources prohibit me from comparing e-governance measures to case studies or analysing different e-government measures quantitatively. However, Bannister (2007) argue that at least the UN's EGDI and the e-government benchmark made by Accenture are consistent in their country ratings. Arguably, this consistency indicates that the two measures are tapping into the same underlying measure of e-government, and the convergent validity of the EGDI is acceptable.

3.4 Control variables and instruments

In this section, I will shortly discuss the control variables in my models. The variables included are theorised to influence predicted values of corruption. Controlling for these factors can provide greater credibility to my model by accounting for relevant alternative explanations.

3.4.1 Democracy

My measure of democracy is the v2x_polyarchy index from V-Dem. Polyarchy refers to Dahl's theoretical framework describing an electoral democracy and the index is aggregated by combining the five components of "Elected Officials, Clean Elections, Associational Autonomy, Inclusive Citizenship, and Freedom of Expression and Alternative Sources of Information separately" (Teorell, Coppedge, Lindberg, & Skaaning, 2019, p. 71). The index ranges from 0-1 and covers most years across 182 countries from 1900-2018 (Michael Coppedge et al., 2019; Teorell et al., 2019). Teorell et al. (2019) convincingly argue for the validity of their measure by undergoing the same rigorous investigations into their variable as described in the section on the corruption index. One of the main strengths of this variable is that it captures many vital concepts that are implied to be essential for e-government to work efficiently, such as freedom of expression and alternative sources of information. Elected officials and clean elections capture the importance of having not only the information to act on, but that real accountability is present.

3.4.2 Log GDP per capita in constant 2010 dollars

Data on gross domestic product per capita is downloaded from the World Bank. GDP per capita is measured annually and consists of the sum of gross value added by all actors in the economy, plus any product taxes, minus subsidies not included in the value of products (World Bank, 2019b). The World Bank has not included the degradation of natural resources or depreciation of assets produced in the economy. Volunteer labour is not included either or the distribution of wealth in society. Essentially GDP is an indicator of how well an economy is doing. Measuring GDP per capita in constant 2010 dollars enables comparability across years as opposed to nominal values affected by year-to-year inflation. The variable is measured annually.

Since GDP is skewed, and not normally distributed across time and countries, I have transformed the variable with the natural logarithm, reducing the skewness. Taking the log of GDP makes interpreting regression coefficients easier since we can understand them in percentage changes of Y. In contrast, the raw data is not necessarily easier to understand intuitively because of the large numbers involved. Percentages solve this problem.

3.4.3 Log Trade (% of GDP)

Trade is annually measured as the sum of exports and imports of goods and services as a share of GDP (World Bank, 2019c). I have undertaken a log transformation of this variable based on the same reasons as with GDP.

3.4.4 Log Natural Resources Rents (% of GDP)

Estimates are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents and forest rents annually (World Bank, 2019e). Rents are defined as "revenues above the cost of extracting the resources" (World Bank, 2019e). Due to the fixed supply of natural resources, profits are often driven more independently of market forces. That is, produced goods are more sensitive to supply and demand effects, since increased supply eventually will diminish demands and in drive profits towards zero. Estimates are measured annually. Log transformation is done to increase the normality of the distribution and ease interpretation.

3.4.5 Instrument - Log lightning strike density

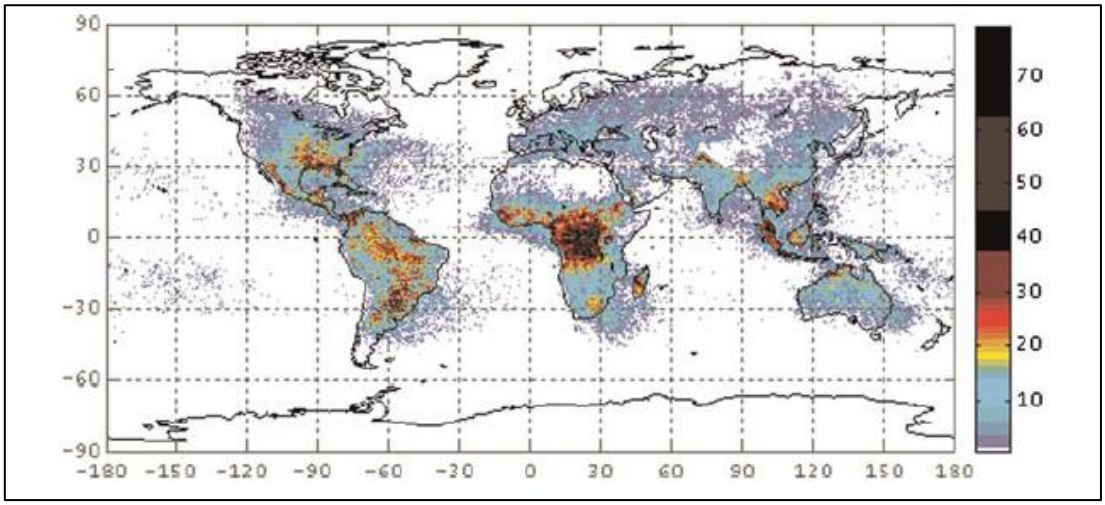
Lightning strike data is originally sourced by Dalgaard et al. (2011) from the National Aeronautics and Space Administration (NASA). Specifically, from the Optical Transient Detector (OTD). The OTD, a space-based sensor, was launched in 1995 and completed a 5-year mission. Values used in this data set is the 5-year average of lightning strikes per square kilometre. A potential problem with this lightning data is that it does not distinguish between cloud-to-cloud and cloud-to-ground strikes (Dalgaard et al., 2011). However, comparing OTD flash density data with U.S. National Lightning Detection Network ground lightning data shows a correlation above 0.95 (Dalgaard et al., 2011).

As discussed in the section "Estimation methods", Wooldridge (2013) presents two requirements for an instrument to be valid: instrumental exogeneity and instrument relevance. Exogeneity involves the instrument being uncorrelated with the error term. Relevance is achieved when the instrument z correlates with x but not directly with y and can be said to be "relevant for explaining variation in x" (Wooldridge, 2013, p. 492). Thus, the logic for implementing this as an instrument is straightforward. Digital technologies such as computers are sensitive to power disruptions and can be damaged by these events. A natural phenomenon causing power outages is lightning strikes. The damage on digital equipment increases the cost of dispersing technology and downtime reduces the efficiency of said technologies. Lightning strikes are therefore used as an instrument for internet dispersion. For this instrument to be

valid, Dalgaard et al. (2011, p. 390) point out that this would require "that there is *no* correlation between lightning density and changes in corruption before 1991 (the founding year of the World Wide Web)". Dalgaard et al. (2011) find no correlation between lightning and changes in corruption before 1991; it only exists after fulfilling the requirements for a valid instrument as described by Wooldridge (2013).

A potential weakness of this instrument is that it does not instrument e-government *per se*, but a measure of internet dispersion, an important driver for e-government. Another potential weakness is that with the global ubiquity of mobile devices such as battery-driven "laptops" or smartphones, a local power outage is not equated with the loss of internet access. Coverage of lightning strikes is not evenly distributed either. Log-transforming data might assign less weight to extreme values, but we should be aware that countries, especially in the Arctic will have significantly less lightning strike observations due to both natural reasons and the range of the OTD.

Figure 3.2: Average flash density (flashes per year per km^2) based on OTD data for the period April 12, 1995 to December 31, 1999 (Dalgaard et al., 2011, p. 407)



3.4.6 Internet and ICRG from Dalgaard et al. (2011)

For the replication of one of the models in Dalgaard et al. (2011), I use data provided by Thomas Barnebeck Andersen. The data contains the ICRG corruption index and a measure of internet users per 100 inhabitants from the ITU. The ICRG variable measures the difference between the years 1991 and 2005, which ranges from 0 to 6, where 6 indicates low corruption. It is very

important to note that I have reversed the variable in this dataset to make the sign go the same way as the V-Dem corruption index. Internet users are measured as the difference between 2005 and 2010. For further description, see (Dalgaard et al., 2011).

3.4.7 Urban population (% of the total population)

Data on the urban population as a percentage of the total population is downloaded from the World Bank. Originally the data is sourced by the World Bank from the United Nations Population Division. The variable measures how many out of a hundred people live in an area defined as urban (World Bank, 2019d).

3.4.8 Log individuals using the internet (% of the population)

Internet users are defined as individuals who have used the internet from any location in the last three months, including a wide range of devices such as computers, mobile phones, games machines, and digital TVs. I have obtained the data from the World Bank, but the original data is sourced from the International Telecommunication Union, World Telecommunication/ICT Development Report and database (World Bank, 2019a).

Internet users are included as a part of my robustness checks, since the EGDI has such a short T, and this variable includes observations as early as the beginning of the 1990s for some countries. Testing the same hypotheses with internet users is therefore interesting from a technical standpoint, since, as mentioned, different estimators have different needs in terms of data access. Theoretically, internet users and the EGDI should be reasonably similar, since it is already a part of the index and is. Researchers such as Dalgaard et al. (2011) (and common sense) argue that internet use is the main driver for e-government solutions. Moreover, we have seen that internet access is hypothesised as a way of empowering citizens by increasing information access and driving demand for investments in human capital (ibid). When access is unrestricted, the internet is also a way for civil society to organise to leverage government power and influence. Mathematically, the correlation between internet users and the linearly interpolated EGDI is also high at 0.84, and a factor analysis shows that they load on one underlying factor, but a low alpha score from a Cronbach's alpha suggests that they do not reliably measure this factor (StataCorp, 2017, p. 10). Both technically and theoretically, the measure of internet users is related to the EGDI, but they are not the same. I do believe, however, that robustness checks with this variable, could provide some idea of how an EGOV model with a larger T could look like.

3.5 Treatment of missing values

Ideally, we would always operate with complete datasets. However, missing values are usual. A value can be missing at random or due to non-random reasons (Mehmetoglu & Jakobsen, 2017). By running a table of missing values in Stata, the variable with the most missing values is the EGDI. Data is not missing at random, but due to the data being gathered biannually. Stata will by default, opt for listwise deletion when it encounters missing values. That is when a country misses a value on one or more variables, all values on that observation will be removed, resulting in a reduced sample (Mehmetoglu & Jakobsen, 2017). When data is missing completely at random, listwise deletion will not create a problem. However, that is not the case in this dataset. To increase the number of observations in my variable of interest, the EGDI, I opt for two solutions: linear interpolation and multiple imputation (Mehmetoglu & Jakobsen, 2017). The reason I opt for two methods is that the multiple imputation suite in Stata is incompatible several of the estimators I wish to use. Therefore, I run most of my regression models based on the linearly imputed variable, while regressions based on the multiple imputation will serve as a robustness check[4]. Hereafter, when I discuss the EGDI, it means the linearly interpolated version. Mehmetoglu and Jakobsen (2017) state that generally linear imputation results in underestimated standard errors, and if multiple imputation is not possible, we should opt for listwise deletion. The reason being that multiple imputation, in contrast to linear interpolation, seeks to mimic the error component in the existing values, not only be as close to the real value as possible, thus maintaining the variability of the original data (Mehmetoglu & Jakobsen, 2017).

4 Estimation Results

In the following chapter. I will review my main empirical findings from the regression models based on the hypotheses and estimation methods described in chapter 2 and 3. To reiterate, I had two main hypotheses H_1 and H_{2a} and two sub-hypotheses:

H_1: E-government causes lower levels of corruption

H_{2a}: Democracy causes lower levels of corruption

H_{2b}: Democracy has a mediating effect on e-government

[4] The results can be seen in appendix 8

H_{2c}: Democracy has a negative and curvilinear effect on corruption

I will start by presenting a table of my main models and move on to describe instrumental variable regression (IV), robustness checks and the Granger causality test before I briefly discuss the regression diagnostics. The models not included in this section will appear in the appendix, such as the FGLS-based estimates, and regressions on the WENAO and NON-WENAO sub-samples.

4.1 Regression models

Table 4.1: The effect of e-government on political corruption

	(1) DK	(2) DK	(3) DK	(4) DK	(5) DK	(6) DK
EGDI interpolated (t-1)		-0.211***	-0.226***	-0.0233	-0.0233	0.0695**
		(0.0257)	(0.0201)	(0.0177)	(0.0167)	(0.0303)
GDP per capita (t-1) (log)	-0.126***	-0.100***	-0.0989***	0.0192	0.0178	0.0171
	(0.00502)	(0.00587)	(0.00553)	(0.0112)	(0.0112)	(0.0106)
Trade (t-1) (log)	-0.00236	-0.0148**	-0.0152**	-0.0312***	-0.0350***	-0.0313***
	(0.00497)	(0.00569)	(0.00556)	(0.00442)	(0.00483)	(0.00460)
Natural resource rents (t-1) (log)	0.0374***	0.0226***	0.0220***	0.0146***	0.0161***	0.0158***
	(0.00334)	(0.00436)	(0.00410)	(0.00279)	(0.00289)	(0.00286)
Urban population (t-1)	0.00278***	0.00193***	0.00194***	0.000329	0.000202	0.000121
	(0.000144)	(0.000120)	(0.000127)	(0.000386)	(0.000344)	(0.000397)
Democracy (t-1)	-0.279***	-0.374***	-0.373***	-0.239***	-0.0188	-0.163***
	(0.0142)	(0.0315)	(0.0316)	(0.0477)	(0.0831)	(0.0408)
Democracy^2 (t-1) (log)					-0.245***	
					(0.0654)	
Year trend			0.00145*	0.00277***	0.00284***	0.00261***
			(0.000688)	(0.000500)	(0.000526)	(0.000542)
Democracy*EGDIi (t-1) (log)						-0.215***
						(0.0579)
Constant	1.485***	1.590***	-1.332	6.167***	6.308***	5.840***
	(0.0449)	(0.0192)	(1.388)	(0.972)	(1.028)	(1.052)
Observations	6,071	2,200	2,200	2,200	2,200	2,200
R-squared	0.546	0.648	0.649			
Number of groups	162	162	162	162	162	162
Time effects	NO	NO	YES	YES	YES	YES
Country fixed effects	NO	NO	NO	YES	YES	YES

Notes:
(1) Driscoll-Kraay robust standard errors in parentheses
(2) Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 4.1 displays the main regression models. All models (columns 1-6) are estimated using OLS with the Driscoll-Kraay robust standard errors (shown in parentheses). All independent variables are also lagged with one year. The first model in column 1 omits e-government allowing the control variables to exert their influence unaffected by the primary independent variable of interest. Controls work as expected, although trade is insignificant which. E-government enters the equation in column 2, and trade becomes statistically significant at the 0.05 level. The dependent variable and the controls now behave as expected according to theory and hypotheses. Substantively, one standard deviation (SD) of around 0.55 points increase in e-government would yield an approximately 20% of a standard deviation (0.05 points) decrease in perceived political corruption, keeping all else equal.

In real-world terms, this equates to going from a country like Bosnia Herzegovina to Colombia, which not necessarily have much significance in practical terms. The rest of the coefficients act as predicted with GDP per capita and trade both being negative and highly statistically significant at the 0.01 level, and a one SD increase in GDP of about 0.13 equates to a 28% of a SD decrease in corruption. Natural resource rents increase corruption as expected and are statistically significant at the 0.01 level. Substantively, a one SD increase in natural resource rents (0.24) will reduce political corruption by about 10% of a SD. Urban population is positively associated with corruption, in contrast to expectations, while democracy has a negative coefficient which is statistically significant at the 0.01 level. In sum, model 2 shows that low-corruption countries are rich and open democratic countries with a low reliance on natural resources and a developed e-government system. However, neither model one or two explicitly account for time or country fixed effect, and results might be driven by time trends and between effects. In column 3, a control for time trend is counted in. The coefficient for Year is positive and significant at the 0.1 level suggesting that corruption increases yearly. Generally, results do not change expressively. There is no changing of signs, and time trends do not seem to matter much for the overall output of the model.

Moving over to column 4, the first of three fixed effect estimations, which is a less efficient, possibly causing the loss of statistical significance in many variables. E-government loses its significance but keeps its sign. GDP per capita also loses significance and changes sign. Trade keeps its significance and sign in all models. Likewise, for natural resource rents, apart from in the first column. Urban population is no longer significant, but Democracy keeps its significance and sign, as it does in all columns. In specification 5, the quadratic term of

democracy enters the equation. The other variables do not change much but consider the two democracy coefficients. While the first coefficient of the two is not statistically significant, the second is, but it should not matter, because they are supposed to be interpreted in relation to each other and are collectively significant using an f-test. Furthermore, the first coefficient has a slightly lower value than the second, suggesting that the curve steepens with higher values of democracy.

Figure 4.1 Predictive Margins for the Squared Term of Democracy

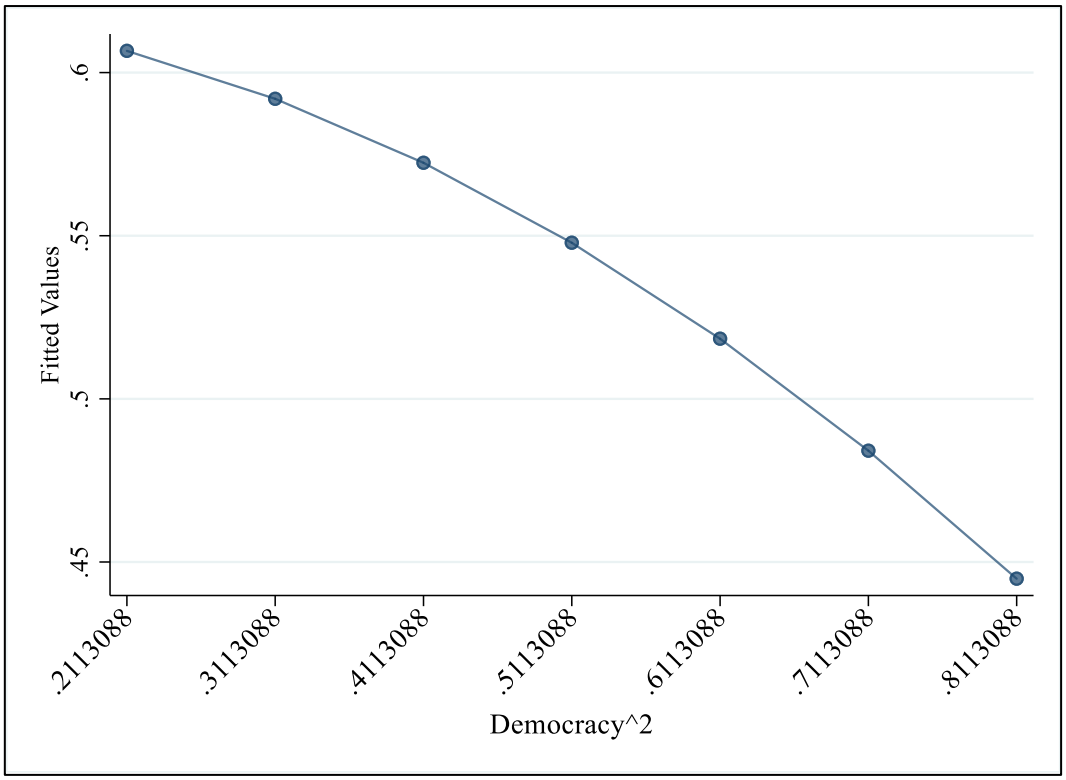
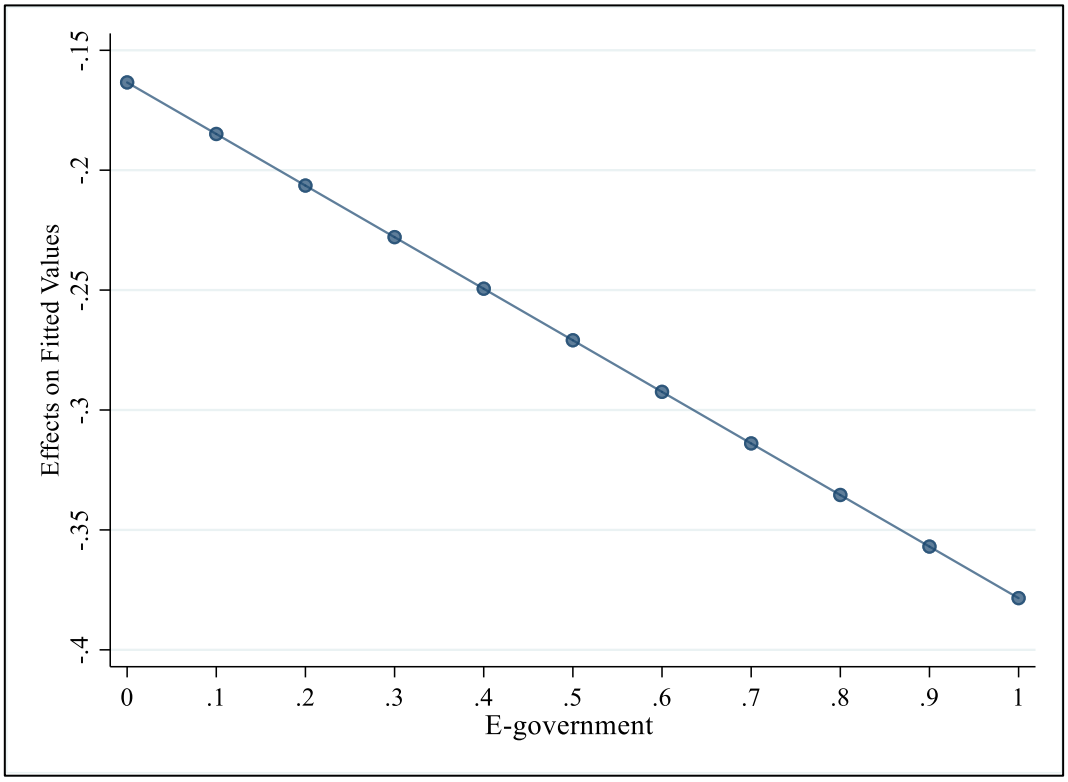


Figure 4.1 demonstrates that the quadratic term of democracy and how it steepens with higher values of the squared term of democracy, democracy^2, on the x-axis. This suggest that democracy's effect on corruption is higher with increasing democracy scores. All predictive margins are significant at the 0.01 level. These results lend strength to H_{2c} that democracy has a negative and curvilinear relationship with political corruption.

Figure 4.2 Average Marginal Effects of Democracy



The interaction term in column 6 is graphed in figure 4.2 and shows the effect of e-government on corruption mediated through democracy. That is, the effect of e-government seems to be enhanced with increased democratic accountability; with bigger values of democracy, the e-government variable increases its anti-corruption effect illustrated by the increasing negative values for the corruption score in the Y-axis for each 0.1 increase in e-government in the X-axis. The interaction term is strengthening the support for my hypotheses H_{2a} and H_{2b} that democracy will assist e-government in reducing corruption, and that democracy has an increasingly powerful effect on combating corruption.

4.2 Instrumental variable regression

Table 4.2: Instrumental variable regression

Instrument: Lightning density (log)	(1) Replication Dalgaard data	(2) Internet users	(3) EGDI interpolated	(4) Internet users	(5) EGDI interpolated
Internet users		-0.0150***		-0.0179***	
		(0.000723)		(0.00273)	
Internet users 0510	-0.00344***				
	(0.000687)				
EGDI interpolated			-1.450***		-7.317**
			(0.0474)		(3.526)
GDP per capita (t-1) (log)				0.0507*	0.566*
				(0.0281)	(0.334)
Trade (t-1) (log)				-0.00628	-0.101**
				(0.00751)	(0.0483)
Urban population (t-1)				0.00236***	0.00865***
				(0.000354)	(0.00328)
N. resource rents (t-1) (log)				-0.0234***	-0.190*
				(0.00872)	(0.107)
Democracy (t-1)				-0.330***	0.183
				(0.0253)	(0.287)
Constant	0.975***	0.827***	1.193***	0.371*	-1.130
	(0.0202)	(0.0144)	(0.0226)	(0.197)	(1.359)
Observations	5,782	4,129	2,622	3,734	2,306
R-squared	0.008	-0.499	0.382	0.166	-3.770
Time fixed effects	NO	NO	NO	YES	YES
Country fixed effects	NO	NO	NO	NO	NO

Notes:
(1) Robust standard errors in parentheses
(2) Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 4.2 shows the second and final stage of the instrumental variable regression calculated with heteroskedasticity robust standard errors. Column 1 contains the replication of the model found in Dalgaard et al. (2011, p. 414), while columns 2 through 5, the non-replication data is used. The findings in column 1 differ somewhat from the findings in Dalgaard et al. (2011, p. 414). Since specifications and statistical software are unlike, results are not expected to be identical. In my estimations, the coefficient's size is both larger and has gained statistical significance, which it does not have in the original version with a coefficient of -0.00011[5] and a standard error of (-0.006).

[5]Note that the coefficient in column 1 is calculated based on the reversed values of the original ICRG variable. That is, in the Dalgaard study, positive values correspond to increased corruption control. In this study a negative coefficient means less corruption. This is done to make a negative sign on the coefficients indicate lower corruption levels in both the ICRG and the V-Dem based regression models.

Column 2 is a replication of column 1 but deviates by using annual V-Dem and ITU data instead of the difference between two points in time. Measuring differences between 1-12 years of corruption using the `ds` prefix for the variable in Stata yielded no statistically significant results. The annual data behave similarly to the differenced in column 1, suggesting that the instrument might be valid to other datasets.

Moving over to column 3, the endogenous variable to be instrumented is no longer internet users, but the interpolated e-government index, EGDI. Lightning strikes were initially intended as an instrument for internet users, but it was at the same time hypothesised to pick up side-effects of internet dispersion such as e-government and higher demand for human capital, both central to the EGDI (Dalgaard et al., 2011, p. 415). The LM value is still well above 10, suggesting a strong instrument relevance, and the coefficients and significance values are as expected.

In column 4 and 5, the control variables are introduced in the equation in addition to time-fixed effects. Internet users are still statistically significant and follow the anticipated sign. For the controls, GDP per capita is now positively associated with corruption together with the urban population control. Natural resource rents are negatively associated with corruption, while democracy is also negative and significant. Column 5 generally show the same pattern when considering coefficient direction and significance. A notable deviation from the reduced model is a substantial increase in the size of the EGDI coefficient.

While the Kleibergen-Paap underidentification test for instrument relevance shows a LM statistic well above 10 for models in column 1 through 4, it rapidly drops when introducing controls. From column 3 to 4, it drops from approximately 375 to 49 and further sink to 4.4 in column 5 suggesting that internet is of weak relevance as an instrument for EGDI when controls are introduced. Kennedy (2005) proposes a plethora of causes that potentially explain the unexpected sign of GDP per capita, but weak instrumental relevance and another specification issue such as a lacking quadratic term might be one of them. Besides, measuring the difference of corruption from different years reverts the sign of GDP per capita, but drains the model's significance, which makes sense considering the persistence of corruption.

4.3 Robustness checks

Table 4.3: Robustness checks

	(1) DK	(2) NW	(3) DK	(4) NW	(5) DK	(6) NW	(7) NW
EGDI ipolated (t-1)	-0.0273	-0.0273					
	(0.0180)	(0.0297)					
Int users (t-1) (log)			-0.0202***	-0.0202***			
			(0.00358)	(0.00311)			
EGDI w. HCI (t-1)							-0.0736***
							(0.0180)
OSI ipolated (t-1)					-0.0502***	-0.0502***	
					(0.00834)	(0.0188)	
HCI ipolated (t-1)					0.0601**	0.0601**	
					(0.0224)	(0.0276)	
TII ipolated (t-1)					-0.00593	-0.00593	
					(0.0230)	(0.0252)	
GDPpc (t-1) (log)	0.0124	0.0124	-0.0633***	-0.0633***	0.0169	0.0169	0.0151
	(0.0117)	(0.0126)	(0.00895)	(0.0135)	(0.0113)	(0.0125)	(0.0111)
Trade (t-1) (log)	-0.0317***	-0.0317***	-0.0348***	-0.0348***	-0.0319***	-0.0319***	-0.0328***
	(0.00461)	(0.00862)	(0.00402)	(0.00561)	(0.00456)	(0.00852)	(0.00463)
NRRents (t-1) (log)	0.00814***	0.00814	0.0205***	0.0205***	0.00736**	0.00736	0.00774***
	(0.00252)	(0.00551)	(0.00260)	(0.00453)	(0.00259)	(0.00554)	(0.00247)
Urban pop (t-1)	0.000387	0.000387	-0.000686	-0.000686	0.000520	0.000520	0.000314
	(0.000415)	(0.000785)	(0.000409)	(0.000527)	(0.000432)	(0.000796)	(0.000452)
Democracy (t-1)	-0.243***	-0.243***	-0.234***	-0.234***	-0.242***	-0.242***	-0.243***
	(0.0472)	(0.0492)	(0.0363)	(0.0335)	(0.0455)	(0.0479)	(0.0465)
Constant	0	1.073***	1.209***	1.355***	0	1.029***	0
	(0)	(0.0885)	(0.0840)	(0.0740)	(0)	(0.0873)	(0)
Observations	2,200	2,200	3,678	3,678	2,200	2,200	2,200
Number of groups	162		162		162		162
Time fixed effects	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES

Notes:
(1) Newey-West or Driscoll-Kraay standard errors in parentheses
(2) Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 4.3 display seven columns with models serving as robustness checks by including alternative model specifications. I have included different dependent variables and added FE-based estimations with the Newey-West (NW) autocorrelation and heteroscedasticity robust standard errors. The motivation being to see whether the main regression models yield the same results as alternative specifications. For example, the NW method does not account for general types of spatial dependence, like the DK method (Hoechle, 2007). Column 1 and 2 contain the specifications using the e-government independent variable with fixed effects and the DK and NW standard errors, respectively. One notable change from table 1 is that the DK model now

contains a time dummy, instead of a linear time trend. Possibly due to collinearity, the constant is therefore excluded from the DK specifications employing the EGDI slightly altering the coefficients from table 1. However, the coefficients will be identical to the ones produced by the NW method when including the year and country dummies (FE). Columns 3 and 4 are also using the fixed effect estimation method with the DK and NW standard errors, but here the EGDI is replaced by internet users. Column 5 and 6 is included to show the decomposed EGDI in action using the DK and NW method, respectively. Column 7 is added to demonstrate a potential problem with my model specification and the EGDI.

At first glance, it is evident that the EGDI models still struggle with finding statistically significant effects. Like in table 1, it is trade and democracy that seem to be the most stable, keeping their sign and statistical significance across all model specifications. Notice that the DK method seem to provide smaller standard errors than the NW estimations, and is probably due to the DK method's short panel and spatial correlation robust properties. However, the smaller standard errors in the DK estimates do not seem to have any significant impact on p-values except for the natural resource rents variable when using the EGDI.

Columns 3 and 4 use the internet users dependent variable, which provides 1478 additional data points and increased variability resulting in an increased number of significant coefficients. Signs and significance are now in line with expectations. Substantial effects are reasonable as well. In column 3 a one SD increase of in internet users of 1.3 equates to a reduction of 38.5% of a SD (0.07) in perceived political corruption. Increasing democracy by one SD (0.08) results in a reduction of corruption equal to a 26.7% of a SD decrease in corruption holding all else equal.

The decomposed e-government development index is introduced in column 5 and 6, allowing us to study the individual effect of each component. Notice that the number of observations is now reduced to the same as in the first columns since we are relying on the EGDI data. Control variables are therefore almost identical in terms of signs and significance as in columns 1 and 2, and I will, therefore, skip the controls and focus on the subcomponents of the EGDI: The Online Service Index (OSI), Human Capital Index (HCI) and the Telecommunication Index (TII). Furthermore, the components have the same coefficients and similar standard errors in both the DK and NW estimation, and I will, therefore, only calculate substantive effects from the DK model. A one standard deviation increase in the OSI of 0.10 results in a reduction in corruption of 11% of a standard deviation (0.48). Counterintuitively, the HCI is positively associated with corruption, and is probably due to some collinearity issue, since the correlation

between HCI and political corruption is -0.5. Nevertheless, increasing the HCI by one SD, according to this model, increases corruption by approximately 8% of a standard deviation. On the other hand, the TII is not significant, but increasing it by on SD equates to a reduction in political corruption of 1% of a SD. In sum, it seems like the online service index is driving the results of the e-government index. The surprising coefficient of the HCI might be cause for concern for my other models. In column 7, I run a model with a modified EGDI where the HCI is omitted. It is created by combining the iOSI and the iTII and dividing them by 2 and removing the potentially problematic HCI results in a significant modified EGDI. By increasing the modified EGDI by one SD gives a small 1.3% of a SD reduction in political corruption.

4.4 Granger Causality

Table 4.4: Granger causality

Corruption → E-government			E-government → Corruption	
Lags	F statistic	Prob > F	F statistic	Prob > F
Lag 1	(1,14) = 2.05	0.1742	(1,14) = 0.54	0.4764
Lag 2	(2,13) = 11.74	0.0012	(2,13) = 0.00	0.9978
Lag 3	(3,12) = 3.57	0.0471	(3,12) = 0.53	0.6730
Lag 4	(4,11) = 5.87	0.0088	(4,11) = 0.49	0.7404
Lag 5	(5,10) = 2.13	0.1222	(5,10) = 1.43	0.2935

Table 4 shows the Granger causality test. For lags 2, 3 and 4, political corruption seems to *Granger cause* e-government, but e-government never *Granger causes* political corruption at a lag of 2, 3 and 4, although f-statistics are not very large, this counter the intuitive and theoretical relationship between these two concepts. Although, the e-government data does not Granger cause corruption in any way, these ambiguous results are not, however, very surprising considering the lack of consistency in findings from previous research.

4.5 Diagnostics

Running diagnostics on the regression models is a way of checking if the models comply with the modified Gauss-Markov assumptions. When working with longitudinal data, several already mentioned concerns are raised regarding the model's validity and reliability. The results of these tests are often a good indication of how to specify our models. Running the `xtserial`

command confirms the presence of autocorrelation (Drukker, 2003), while the `xttest3` confirms heteroscedasticity (Greene, 2003), and the `xtcd2` (Pesaran, 2015) spotted spatial dependency suggesting that standard errors should be adjusted using the DK method. `Xttest0` (Breusch & Pagan, 1980) show the presence of random effects suggesting running an FGLS RE, while an F-test on the time and country dummies detected time and country fixed effects. Running a "dfuller" test on the variables rejecting the null hypotheses that at least one panel contains a unit root, rendering a lagged dependent variable unnecessary (Choi, 2001). Predicting Cook's D values using show no values above 1, but several above 4/N. I choose to follow Mehmetoglu and Jakobsen (2017) in running the model based on the cut-off point of 1. The main reason being that a cut-off above 4/N would render the number of observations small, which is demonstrated in appendix number 6. Running an OLS allows for using the user-written program `regcheck` by Mehmetoglu and Jakobsen (2017) show that there is a potential issue with the functional form evident by the significant RESET test.

5 Discussion

In the previous chapter, I presented the result from my data analysis from a technical perspective. Nonetheless, raw numbers add little meaning by themselves, only analysis grounded in theory and previous research answer the question whether the regressions models strengthen or weaken the support for the hypotheses. In this chapter, I will review the hypotheses and methods to evaluate the empirical analysis. The following discussion is structured after the hypotheses, starting with the main hypotheses that e-government will reduce corruption, followed by a debate whether democracy leads to lower corruption and how different levels of democracy effects e-government efficiency and corruption. Uncertainties surrounding causality and which direction it points in the research is also emphasised. The hypotheses were:

H_1: E-government causes lower levels of corruption

H_{2a}: Democracy causes lower levels of corruption

H_{2b}: Democracy has a mediating effect on e-government

H_{2c}: Democracy has a negative and curvilinear effect on corruption

5.1 Does e-government cause lower corruption?

As ICTs rapidly become more widespread, the need to know more about the real-world effects of digital solutions in society and government becomes increasingly essential. E-government initiatives are hailed by practitioners and policymakers alike as a way of improving government through a networked structure of interconnectivity, aid in service delivery, efficiency and effectiveness, interactivity, decentralisation, transparency, and accountability (Yildiz, 2007). One often discussed effect of e-government is its presumed ability to prevent political corruption. Previous anecdotal research shows that e-government solutions have successfully reduced corruption when e-government measures are implemented correctly. Quantitative studies also indicate that e-government and internet use reduce corruption. However, general purpose technologies like the internet and its derivatives change rapidly in unknown ways, and they may have a very different real-world meaning for corruption and democracy in 2003 and 2018. As data availability is rapidly increasing and e-government research is less practitioner-oriented and gradually more cognisant of potential pitfalls of e-government, asking the question of whether e-government improves government and cause less corruption is a valid one. More importantly is knowing how to maintain inclusive institutions.With the ever-increasing availability of data, I employed time-series methods to a global panel to critically investigate the claims of e-government's anti-corruption capabilities.

The answer whether the empirical results strengthened or weakened the support for the hypotheses is not clear. The answer lies in a grey area between an explicit confirmation and rejection. There are several explanations for this, but they can be categorised into two related but distinct classes: technical explanations and substantial explanations. In most models, the EGDI is behaving as expected in relation to corruption control, even after the control variables are allowed to exert influence. Following these models only, we get a strong indication that e-government is a viable anti-corruption tool globally, even after controlling for some of the most commonly agreed predictors of corruption. However, in all but a couple of the tables, including the robustness checks, the EGDI loses its significance when introducing fixed effects, suggesting from a technical standpoint that the within variance is too small to detect any effects after accounting for time-constant omitted variables. Although Hoechle (2007) argues for the Driscoll-Kraay standard error's large N, small T robustness, losing significance with a short panel comes as no surprise. In the discussion on estimators, I have already warned against potential issues with the FE estimator on short panels with slowly changing variables increasing the danger of committing type II errors failing to reject a false null hypothesis (Beck, 2001;

Nickell, 1981; Wilson & Butler, 2007). Moreover, the loss of statistical significance in the fixed effects models potentially suggests that much of the variance in the data potentially stems from between country effects. That is, low-corruption-high-e-government countries drive the results in the pooled models in column 1 through 3 in the main models, while the within variance is too small to yield any significant results. The general increase in global average EGDI scores since 2003 can be part of an explanation that is showing that a global improvement of e-government and not necessarily be followed by a decrease in corruption within countries. Another technical explanation is that the statistically significant effects are obscured by influential units. However, introducing internet users as a proxy for e-government increases the number of observations drastically and yields results that are expected based on previous research and theoretical arguments. Internet use is just related to the concept of e-government and might have very different effects on corruption than e-government. Presumably, internet use is more related to a bottom-up control of corruption, while e-government systems are implemented from the top-down, implicating that the state implementing it must have pluralistic institutions allowing for openness and accountability. That is, e-government as a corruption tool is probably more dependent on a democratic system than internet use. As seen in column 6 in the main models, the inclusion of an interaction term between e-government and corruption allows e-government to regain its significance, but with a changed sign indicating that in countries with a democracy score of 0 e-government serves to increase corruption. More on this further below.

Another concern related to the validity of the e-government index (EGDI), is the lack of openness surrounding its survey materials, making the real validity of the index even harder to gauge. Furthermore, the EGDI proxies the number of e-government users by including the TII, but it can't with certainty account for e-government's "dirty little secret"; that we do not know much about the extent of how much citizens actually use these "shiny new tools" (Bannister, 2007, p. 181). A technical concern with the index arises when studying the individual effect of each index component, that is the OSI, TII and HCI. It becomes apparent that the components in the index might drive the results in unexpected directions. As seen in the robustness tests, the only component driving the results is the OSI. This contrasts the findings made by Elbahnasawy (2014) showing that both the TII and the OSI drive the results, while he did not find the HCI to be significant either. Although the TII does not significantly drive any results in my robustness test, the internet users variable basically measures the same concept and is consistent in finding a negative relationship with corruption. The significant OSI is hard to

interpret because of the lack of methodological transparency from the United Nations. One explanation is that the OSI is indicating the transparency and usability of a country's e-government services which in turn leads to corruption control. The countries with a higher score on the OSI might also be countries that are already open and have institutions that are willing to exert limitations on their own power. Human capital is seen as an important anti-corruption factor in much literature, however neither these data nor Elbahnasawy (2014) find statistical evidence suggesting it is important for reducing corruption. Theoretically, this makes little sense since educated citizens are presumably better at obtaining and processing complex information which is not a liability when interacting with government agencies. However, a possible explanation is that many countries have both highly developed educational systems, but at the same time have institutions rampant with corruption. Also, if e-government systems are intuitively designed, higher educations might not matter much for its usability.

Granting the results produced by the FE models, the benefit of the doubt accepting that the within effect is insignificant and sometimes even positive, a series of substantial interpretations arise. Across countries, unevenly distributed political, economic and societal capacities reinforce the digital divide and consequently obstacles for successful e-government implementations. According to Heeks' (2002), there are three archetypes of failure created by different reality gaps: hard-soft, private-public, country context gaps. To reiterate: Hard-soft gaps arise when the rationality and rules of technology meet the reality and social context in which the technology operates. An example is bureaucrats resisting implementation of e-government systems or learn how to circumvent the system's rules to continue with corrupt behaviour. Private-public gaps occur when technology is transferred from the private sector to the public sector without accounting for differences in cultures and capacities. Lastly, country context gaps rise when e-government systems are transferred between developed and developing countries – New Delhi is not New York.

Accounting for potential reality gaps and developmental differences, the lack of expected results in the regressions are easier to accept since the FE estimator account for the unobserved and relatively time constant country contexts that take a long time to change. In other words, there are gaps between and within countries that are driven by slowly changing factors and not even e-government is a quick-fix anti-corruption measure because of the persistence of corruption and its material and institutional determinants. Thus institutions also matter for outcomes of e-government. But what constitutes good institutions and how do they come about? What constitutes a good institution in the discussion on democracy and development, seem to

be determined by many of the same factors determining the outcome of e-government initiatives. Success rate seem to be highly dependent on contextual factors such as access to resources, trained personnel, and not at least, political will. As a humble warning for policymakers: these factors are seemingly needed to be of a certain quality before investing in costly digitisation projects. The fact that investment in ICT projects presents corruption opportunities in themselves illustrates the uphill battle against corruption and institutional inertia (Charoensukmongkol & Moqbel, 2014). According to my data, e-government is therefore likely to be of greatest benefit to advanced democracies with the necessary resources, demonstrated by the fixed effect regressions using subsamples consisting of solely 'Western' nations and one including the rest. More prosperous nations are presumably in the innovation forefront of e-government systems, making it likelier that technologies are adapted to their public sector cultures, and if a system is shown to be a bad fit, they have resources to alter or replace the defunct technology, effectively minimising risks of public-private gaps having a detrimental impact. Birds of a feather flock together, and reality gaps will be less costly to overcome in and between similar cultures and nations, especially western ones. Institutionally these nations are also more responsive partly due to their already low corruption and inclusivity.

Because of contextual factors having such a significant impact on the effect of e-government, the results indicate that the relationship between e-government and corruption control is highly dependent on other factors such as institutional contexts. E-government might reduce corruption, but on a global scale, it might not be the first tool of choice when the time comes to choose the right tools to reduce corruption.

Furthermore, my data does not provide a clear answer to which way the causal effect goes. By employing a Granger causality test it is possible to assess the temporal relationship between two variables, to see if our concepts are subjects to the theorised causal process: is our X-variable occurring before Y in time? Beware that the Granger causality is not true causality in a sense and should be interpreted with caution. After all, the rooster *granger causes* the sunrise every morning, but it obviously does not make any sense to argue that the rooster is both necessary and sufficient for the sunrise occurring. In this data the Granger causality test provided ambiguous answers. Corruption granger causes corruption in lags 2 through 4, while the opposite is not the case, e-government does not granger cause corruption. Although intuitively corruption should not cause e-government, one possible explanation might be that countries with high corruption would be motivated to acquire e-government solution for

themselves. Implementing new ICT solutions in large organisations takes time, maybe even 2 to 4 years?

The motivation behind employing an instrumental variable regression is often to get rid of problems with measurement errors, and time varying fixed effects. But most importantly, the IV regression implemented with a strong instrument should get rid of endogeneity concerns and allow the researcher to infer a certain degree of causality. The results of the IV regressions show that lightning strikes strongly instrument internet users and the EGDI which in turn are associated with corruption control. The exogeneity of the instrument allows us to infer some causality to the initial claim that e-government reduces corruption. While Dalgaard et al. (2011) make a strong case for lightning strike density as an instrument for internet users, they never intended it to directly instrument e-government which was only mentioned as a side-effect of internet dispersion, and the instrument validity should be taken with a small grain of salt. Dalgaard et al. (2011) also used a difference-in-difference design looking at changes in corruption and changes in internet users. This dataset does not, however, find any significant results when differencing one or more years, but the IV regression behaves as expected according to the hypotheses when ignoring this and instead considering the levels in corruption suggesting that in contrast to the granger results, internet and e-government reduces corruption when removing some of the bias from the estimator. However, introducing the control variables in the IV regression models, makes the instrument significantly weaker, suggesting that the instrument is possibly not as robust as initially hoped. The large change in the coefficients, especially for the EGDI implicates that the instrument is weak and that results, at least in the multivariate regression, should be interpreted with caution.

Having considered both technical and substantial possibilities, H_1 is weakened when compared to previous quantitative research which usually find that e-government and internet use clearly reduces corruption by increasing transparency and accountability allowing the principal citizen to exert increased control on the agent civil servant and politician. My findings are more ambiguous leaving me to abstain from a clear rejection of the null hypothesis. Institutional determinants and other contextual factors seem to matter more than some research has indicated.

5.2 The effect of democracy on corruption

E-government is perceived as a way of increasing government transparency and minimising the amount of decentralised information and information asymmetry that exist in the relationship

between citizen and government services. Providing rational citizens with essential information should reduce the principal-agent problem. Two essential components omitted from this theory: the necessity of real accountability and the incentive to report corruption. Democracy may be able to provide the cure for that; horizontal, vertical and societal democratic accountability provide residents with the leverage needed to combat corrupt street bureaucrats and higher-ups in the political pecking order. The relationship between the measure of democracy and corruption was, therefore hypothesised to be both negative and significant. While research argues that while democracy reduces corruption, in transitions periods democracies might provide more corruption opportunities than full-fledged autocracies since even the introduction of free and fair elections might present corruption opportunities (McMann et al., 2017). Based on that assumption, I introduced the quadratic term of democracy to capture this effect. Lastly, to more explicitly test whether the effect of e-government is dependent on democratic accountability, I added an interaction term between the two inspired by Kock and Gaskins (2014). The motivation for the extensive focus on democratic institutions is inspired by the previous research on e-government and corruption which has approached the topic with technological determinism, not adding much emphasis on the idea that e-government can be turned against the citizenry's best interest, rotating the initially stated principal-agent problem on its head to reduce individual and civil society freedom.

In all the models, democracy's effect on corruption control is found to be one of the most consistent of all variables included, adding to the credibility to H_{2a}: *Democracy causes lower levels of corruption*. The finding suggests that the theoretical arguments are sound, and that democratic accountability is contributing to combating corruption. Furthermore, the curvilinear effect is also found to be significant. However, when graphed, the curvilinear effect is seemingly weak when using a fixed effect model, in contrast to the much clearer curve when omitting fixed effects. Still, the finding supports the notion that building democratic institutions is not an immediate solution to corruption problems and is dependent on factors such as election quality (McMann et al., 2017). Another possible explanation for the increase in perceived corruption in transitional democracies is that increasing openness makes corruption visible and impact expert evaluations of corruption levels. Increases in corruption in lower levels of corruption might, therefore, indicate a growing awareness of corruption, which might trigger further anti-corruption measures from individuals and non-governmental organisations alike. I, therefore, conclude that *H_{2c}: Democracy has a negative and curvilinear effect on corruption*

has gained strength. The mediating effect of democracy on e-government is statistically reliable in the regression models adding significance to even the fixed effect estimation.

While democratic responsiveness contributes to reducing corruption, an important question remains to be answered; does e-government mean the same in autocracies and democracies? To test this mediating relationship democracy has between e-government corruption, I added an interaction term to the baseline model. The interaction term between e-government and democracy suggest that democracy is essential to the effect of e-government, a finding that is, to the best of my knowledge, under communicated in the literature. In countries with a theoretical score of 0 on democracy, the e-government variable was positively associated with corruption, while the effect of e-government increased in strength with increases in the score on democracy. Reasons for this may be many, but in the discussion on e-government in low-democracy settings, several examples of how governments restrict information access on the internet is presented. Honourable mentions include the outlawing of the spread of "pessimism" online in Turkey or the pre-installed firewall in ICTs sold in China - not to mention the highly controversial e-government G2C "social credit system" proposed implemented in 2020 by the Chinese authorities. Though countries restrict information freedom, their citizens still might have similar e-government services as in far more democratic countries like electronic registration of taxes or electronic applications for old-age pensions. However, restricting internet access essentially chokes of any societal accountability, rendering both e-government and internet weak tools for citizens confronted with corruption. Although there exist ways of circumventing firewalls, it is probably a privilege enjoyed by the technically apt, and not the case for most internet users, rendering collective action unviable. Democracy scores and internet freedoms are also likely endogenous factors, meaning that the one predicts the other, possibly reinforcing the existing institutions in the country. That is, democracies might become more open as a result of opening the bureaucracy and cutting down the distance between the citizen and the public sector. On the other hand, autocratic regimes can increase surveillance and worsen the principal-agent problem for the citizen. Such a possibility is not unlikely following the discussion on inclusive and extractive institutions where the persistence and re-enforcing mechanisms of institutions is so strong that e-government just becomes another tool for the incumbent to use in the pursuit for power.

A possible wealth effect is also worth considering. Richer countries, whom are often (for now) democracies, are more likely to see a much smaller digital divide, meaning that the positive effects of digitisation benefit most citizens. This effect is demonstrated in appendix 3. On the

other hand, in less developed nations, abolishing the role of "street bureaucrats" and brick and mortar offices might not be the death of distance digital solutions aim at achieving. The opposite might be the result for people without the budget or infrastructure to reap the benefits of the computer age. The haves would then be the ones with extensive access to their government while the have nots are left with a more distanced relationship and the few times they have to interact with the state, bribes are likely to continue being paid because of the persistency of principal-agent problems that are maintained by keeping people in the dark.

In sum, both theoretical and empirical results lend credibility to hypotheses H_2a, H_{2b}, and H_{2c} that democratic responsiveness the efficiency of e-government as an anti-corruption tool. Furthermore, democracy is also shown to have a curvilinear relationship even in a fixed effect estimation, and the interaction term show how the effect of e-government is mediated through democratic accountability. The technological optimism and determinism that has characterised some literature is, therefore, doubtable.

6 Conclusion

Governments, international organisations and the research community agree that the extensive use of information and communication technology in the public sector, also known as e-government, presents an opportunity to enhance transparency, accountability and efficiency. These are traits that are recognised as contributing to lower corruption levels. Thus, a natural consequence of e-government should be a reduction in corruption and more inclusive institutions consequently contributing to development. Corruption is in many ways the antitheses to these concepts, as well as democracy and inclusive institutions. This paper therefore sat out to answer the question: *Does e-government cause lower levels of corruption?* Furthermore, by identifying a gap in the literature, I set out to examine an auxiliary hypothesis that democracy is essential to combating corruption, but especially crucial to the success of e-government as an anti-corruption tool

By employing a longer panel than previously available and a novel measure of corruption provided by the Varieties of Democracy project, the question is considered using rigorous time-series estimations. Based on the empirical results, the answer, however, is not clear. In contrast to earlier research, econometric results are ambiguous with regards to the effect of e-government on its own. To explore the effect of different levels of democracy I followed McMann et al. (2017) and employed a quadratic term to see how different levels of democracy impact corruption levels. Inspired by Kock and Gaskins (2014) I specified a model using an

interaction term between e-government and democracy allowing me to study the effect of e-government at different levels of democracy. All three specifications of democracy, that is the linear specification, the quadratic and the interaction term yielded the results expected by theory and previous research.

For the purely linearly specified models, the e-government variable loses its effect when employing the fixed effect estimator. I analyse the absence of expected results for the e-government models alone by separating the explanations in two categories: technical and substantial reasons. The technical perspective on the loss of significant findings in the fixed effects (FE) models is that it occurs due to a lack of data resulting in a low within variation in each country. Institutions are slowly changing so the same goes for corruption. If we ignore the risk of committing a type II error, by failing to reject the false null hypothesis, and grant the FE results the benefit of the doubt, we can understand the finding in more substantial ways. As an example, even investing in the infrastructure needed to implement e-government systems presents an opportunity for corruption, suggesting that countries with pervasive corruption are fighting an uphill battle, meaning that an increase in the EGDI can have opposite effects for developed and developing nations.

However, considering the risk of omitted variable bias, adding the interaction term between e-government and democracy yield widely different results, even using the FE estimator. This suggest that democracy is impacting high and low-democracy countries differently. In countries with low scores democracy, e-government worsen corruption. On the other hand, the higher the score on democracy, the more effective e-government becomes at tackling corruption problems. This result suggest that e-government and its related technologies does not automatically improve institutions suggested by much of the early optimism in e-government literature (Dada, 2006). I argue that the impact of ICT on institutions and corruption is therefore highly path dependent on pre-existing institutional qualities. Consequently, good institutions, such as those found in advanced democracies are further strengthened by e-government. In contrast, autocracies can cement their power by using pervasive surveillance, firewalls and other measures to keep citizens uninformed and unorganised without any possibility or incentive to report corruption.

E-government research is a field in rapid development and in continuous need of further exploration. Technologies are changing fast, and as a consequence the implications of implementing them are varying annually – something that needs to be continuously conveyed to stakeholders seeking to develop e-government. Once again, it is apparent that technologies

are neither inherently good nor bad, and studies on these concepts are in other words highly relevant for both citizens and democracy. The last 20 years of development in ICT is a blatant proof of this, both in positive and negative terms with the development of highly efficient e-government solutions in countries like Norway, but there is no lack of examples of mass surveillance and censorship, even in democracies. The results from this study show that e-government and ICTs are most effective in countries with already strong institutions, while countries seeking to employ e-government as a solution to widespread corruption, should particularly careful to not import systems they are not capable of maintaining or handling due to weak institutions or the lack of infrastructure. Previous research on Switzerland by Misuraca et al. (2010) indicate that the most effective e-government systems are "home grown" and therefore adapted to local needs. However, the discussion on the digital divide demonstrate that not all countries have the capabilities needed to develop their own e-government solutions nor the capacity to readily adopt state-of-the-art systems from richer countries. Nations with a more moderate development level looking to up their use of technology in the public sector should therefore seek to find what comparable nations do identify best practices, and not necessarily jump directly to implementing the state-of-the-art technology. In countries where ICTs are only recently becoming widespread, digital divides may exclude the have nots. Incremental steps are therefore warranted to account for institutional lag. Most importantly, though, is that e-government projects are implemented in a way that supports the development of inclusive institutions by maintaining the respect for individual rights.

I conclude that international organisations, researchers, policy makers and other stakeholders need to consider the widely different implications e-government has in different institutional contexts, and that there is no universal solution for improving institutions and consequently boosting development. Future research should therefore focus on studying subgroups of countries something space limitations has hindered this thesis from exploring in depth. Elbahnasawy (2014) also encourages research on the effect of e-government on different types of corruption such as grand and petty, something that can be done with the V-Dem corruption index. The study of ICT has been around in various forms since the 1960s, however, adequate theory building and empirical studies have been scant (Bannister & Connolly, 2015). Political scientists are now seemingly catching up to the practitioner-oriented literature and can further add theoretical knowledge to the field of e-government. Moreover, measures of e-government seem to still be constructed somewhat arbitrarily because of an apparent lack of theoretical ballast and naturally because of the rapidly changing nature of the concept. Researchers,

including the writer of this text, therefore have to a certain degree blindly trust that the operationalisation of the e-government measures are theoretically valid, mostly based on the reasoning that this and that variable is "widely cited". This is also one of the main weaknesses with this study, namely the lack of a long time series and strong measures. Furthermore, studies on the use of e-government could provide useful insights, since few studies have data on how widely used these "shiny new" e-government services are. However, the United Nations E-participation index is a potential useful tool to this end. Methodical discussions aside, my results indicate that e-government has a potential for improving institutions, be there inclusive or extractive ones, but e-government does not seem to have a uniform impact on all countries, and should be considered with scrutiny as a tool for development.

7 References

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail : the origins of power, prosperity, and poverty*. London: Profile Books.

Achen, C. H. (2000). Why lagged dependent variables can suppress the explanatory power of other independent variables. *Ann Arbor, 1001*(2000), 48106-41248.

Achen, C. H. (2005). Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. *Conflict Management and Peace Science, 22*(4), 327-339. doi:10.1080/07388940500339167

Al-Ali, S. (1991). Technological dependence in developing countries: A case study of Kuwait. *Technology in Society, 13*(3), 267-277. doi:https://doi.org/10.1016/0160-791X(91)90003-F

Andersen, K. N., Henriksen, H. Z., Medaglia, R., Danziger, J. N., Sannarnes, M. K., & Enemærke, M. (2010). Fads and Facts of E-Government: A Review of Impacts of E-government (2003–2009). *International Journal of Public Administration, 33*(11), 564-579. doi:10.1080/01900692.2010.517724

Andersen, T. B. (2009). E-Government as an anti-corruption strategy. *Information Economics and Policy, 21*(3), 201-210. doi:https://doi.org/10.1016/j.infoecopol.2008.11.003

Baldersheim, H., Øgård, M., & Haug, A. V. (2008). *Mot den virtuelle kommunen : studier i e-demokrati og e-forvaltning*. Bergen: Fagbokforl.

Baltagi, H., Badi. (2005). *Econometric Analysis of Panel Data* (3rd ed.). West Sussex: John Wiley & Sons Ltd.

Bannister, F. (2007). The curse of the benchmark: an assessment of the validity and value of e-government comparisons. *International Review of Administrative Sciences, 73*(2), 171-188. doi:10.1177/0020852307077959

Bannister, F., & Connolly, R. (2015). *The great theory hunt: Does e-government really have a problem?* (Vol. 32).

Beck, N. (2001). Time-Series–Cross-Section Data: What Have We Learned in the Past Few Years? *Annual Review of Political Science, 4*(1), 271-293. doi:10.1146/annurev.polisci.4.1.271

Beck, N. & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American political science review, 89*(3), 634-647.

Beck, N. & Katz, J. N. (2011). Modeling Dynamics in Time-Series–Cross-Section Political Economy Data. *Annual Review of Political Science, 14*(1), 331-352. doi:10.1146/annurev-polisci-071510-103222

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods, 3*(1), 133-153.

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly, 27*(3), 264-271. doi:https://doi.org/10.1016/j.giq.2010.03.001

Bhatnagar, S. (2003). E-government and access to information. *Global corruption report, 2003*, 24-32.

Box, G. E., & Draper, N. R. (1987). *Empirical model-building and response surfaces*: John Wiley & Sons.

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The review of economic studies, 47*(1), 239-253.

Buckman, G. (2008). *Globalization Tame it or scrap it?* S.l.]: S.l. : NBN International.

Bueno de Mesquita, B., & Smith, A. (2012). *Dictator's Handbook* (Vol. 1). New York, NY: PublicAffairs.

Charoensukmongkol, P., & Moqbel, M. (2014). Does Investment in ICT Curb or Create More Corruption? A Cross-Country Analysis. *Public Organization Review, 14*(1), 51-63. doi:10.1007/s11115-012-0205-8

Charron, N., Dahlström, C., & Lapuente, V. (2015). *Measuring Meritocracy in the Public Sector in Europe: a New National and Sub-National Indicator* QoG Working Paper Series. Political Science The Quality of Government Institute Gothenburg.

Chinn, M. D., & Fairlie, R. W. (2007). The determinants of the global digital divide: a cross-country analysis of computer and internet penetration. *Oxford Economic Papers, 59*(1), 16-44. doi:10.1093/oep/gpl024

Choi, I. (2001). Unit root tests for panel data. *Journal of international money and Finance, 20*(2), 249-272.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, I., Staffan, Teorell, J., Altman, D., . . . Ziblatt, D. (2019). V-Dem Codebook v9. Retrieved from

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., . . . Ziblatt, D. (2019). *V-Dem [Country-Year/Country-Date] Dataset v9*

Dada, D. (2006). The failure of e-government in developing countries: a literature review. *The Electronic Journal of Information Systems in Developing Countries, 26*(1), 1-10.

Dahlström, C., Lapuente, V., & Teorell, J. (2012). The merit of meritocratization: Politics, bureaucracy, and the institutional deterrents of corruption. *Political Research Quarterly, 65*(3), 656-668.

Dalgaard, C.-J., Bentzen, J., Selaya, P., & Andersen, T. B. (2011). Does the Internet Reduce Corruption? Evidence from U.S. States and across Countries. *The World Bank Economic Review, 25*(3), 387-417. doi:10.1093/wber/lhr025

Della Porta, D., & Vannucci, A. (1999). *Corrupt exchanges : actors, resources, and mechanisms of political corruption*. New York: Aldine de Gruyter.

Dobel, P. (1999 [1978]). The Corruption of a State In D. Della Porta & A. Vannucci (Eds.), *Corrupt Exchanges: actors, resources and mechanisms of political corruption* (pp. 1-33). New York: Aldine de Gruiter

Doong, S. H., & Ho, S.-C. (2012). The impact of ICT development on the global digital divide. *Electronic Commerce Research and Applications, 11*(5), 518-533. doi:https://doi.org/10.1016/j.elerap.2012.02.002

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *Stata Journal, 3*(2), 168-177.

Elbahnasawy, N. G. (2014). E-government, internet adoption, and corruption: an empirical investigation. *World Development, 57*, 114-126.

Fukuyama, F. (2012[1992]). *The End of History and the Last Man*: Penguin Books.

Greene, W. H. (2003). *Econometric analysis*: Pearson Education India.

He, Z. (2000). Corruption and anti-corruption in reform China. *Communist and Post-Communist Studies, 33*(2), 243-270. doi:https://doi.org/10.1016/S0967-067X(00)00006-4

Heeks, R. (2002). Information Systems and Developing Countries: Failure, Success, and Local Improvisations. *The Information Society, 18*(2), 101-112. doi:10.1080/01972240290075039

Heeks, R. (2003). *Most e-government-for-development projects fail: how can risks be reduced?* (Vol. 14): Institute for Development Policy and Management, University of Manchester ….

Hobbes, T. (2000 [1651]). Leviathan. In W. Ebenstein & A. Ebenstein (Eds.), *Great Political Thinkers* (6 ed., pp. 364-379). New York: Harcourt Brace: Wadsworth.

Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *Stata Journal, 7*(3), 281.

Homburg, V. (2008). *Understanding E-Government: Information Systems in Public Administration*: Routledge Ltd.

Huntington, S. P. (1968). Modernization and Corruption. In J. A. Heidenheimer, M. Johnston, & T. V. LeVine (Eds.), *Political Corruption: A Handbook* New Brunswick Transaction Publishers

Jain, A. K. (2001). Corruption: A Review. *Journal of Economic Surveys, 15*(1), 71-121. doi:http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291467-6419/issues

Kaufmann, D., & Wei, S.-J. (1999). *Does" grease money" speed up the wheels of commerce?* Retrieved from

Kennedy, P. E. (2005). Oh no! I got the wrong sign! What should I do? *The Journal of Economic Education, 36*(1), 77-92.

Kim, S., Kim, H. J., & Lee, H. (2009). An institutional analysis of an e-government system for anti-corruption: The case of OPEN. *Government Information Quarterly, 26*(1), 42-50. doi:https://doi.org/10.1016/j.giq.2008.09.002

King, G., & Roberts, M. E. (2015). How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It. *Political Analysis, 23*(2), 159–179.

Kock, N., & Gaskins, L. (2014). The Mediating Role of Voice and Accountability in the Relationship Between Internet Diffusion and Government Corruption in Latin America and Sub-Saharan Africa. *Information Technology for Development, 20*(1), 23-43. doi:10.1080/02681102.2013.832129

Kolstad, I., & Wiig, A. (2009). Is Transparency the Key to Reducing Corruption in Resource-Rich Countries? *World Development, 37*(3), 521-532. doi:https://doi.org/10.1016/j.worlddev.2008.07.002

Kraay, A., & Kaufmann, D. (2002). *Growth without governance*: The World Bank.

Kumar, R., & Best, M. L. (2006). Impact and Sustainability of E-Government Services in Developing Countries: Lessons Learned from Tamil Nadu, India. *The Information Society, 22*(1), 1-12. doi:10.1080/01972240500388149

Laffont, J.-J., & Martimort, D. (2002). *The theory of incentives : the principal-agent model*. Princeton, N.J: Princeton University Press.

Lee, J. (2010). 10year retrospect on stage models of e-Government: A qualitative meta-synthesis. *Government Information Quarterly, 27*(3), 220-230. doi:https://doi.org/10.1016/j.giq.2009.12.009

Leys, C. (1965). What is the Problem about Corruption? *The Journal of Modern African Studies, 3*(2), 215-230.

Lio, M.-C., Liu, M.-C., & Ou, Y.-P. (2011). Can the internet reduce corruption? A cross-country study based on dynamic panel data models. *Government Information Quarterly, 28*(1), 47-53. doi:https://doi.org/10.1016/j.giq.2010.01.005

Lührmann, A., Gastaldi, L., Grahn, S., Lindberg, I., Staffan, Maxwell, L., Mechkova, V., . . . Pillai, S. (2019). Democracy Facing Global Challenges In *V-DEM ANNUAL DEMOCRACY REPORT 2019*.

McMann, K., Pemstein, D., Seim, B., Teorell, J., & Lindberg, S. (2016). Strategies of validation: Assessing the Varieties of Democracy corruption data.

McMann, K., Seim, B., Teorell, J., & Lindberg, I., Staffan. (2017). *Democracy and Corruption: A Global Time-Series With V-Dem Data* Retrieved from

Mehmetoglu, M., & Jakobsen, T. (2017). *Applied Statistics Using Stata* London: Sage Publications Ltd.

Mejias, U. A., & Vokuev, N. E. (2017). Disinformation and the media: the case of Russia and Ukraine. *Media, Culture & Society, 39*(7), 1027-1042. doi:10.1177/0163443716686672

Mills, M. (2008). Globalization and Inequality. *European Sociological Review, 25*(1), 1-8. doi:10.1093/esr/jcn046

Milner, H. V. (2006). The Digital Divide: The Role of Political Institutions in Technology Diffusion. *Comparative Political Studies, 39*(2), 176-199. doi:10.1177/0010414005282983

Misuraca, G., Rossel, P., & Glassey, O. (2010). Overcoming barriers to innovation in E-Government. In P. G. Nixon, V. N. Koutrakou, & R. Rawal (Eds.), *Understanding E-Government in Europe*. New York: Routledge

Nam, T. (2018). Examining the anti-corruption effect of e-government and the moderating effect of national culture: A cross-country study. *Government Information Quarterly, 35*(2), 273-282.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica, 49*(6), 1417-1426. doi:10.2307/1911408

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Ojha, A., & Palvia, S. (2012). E-Government and the Fight Against Corruption: Conceptual Modeland Five Case Studies from India. *Journal of Information Technology Case and Application Research, 14*(4), 11-29. doi:10.1080/15228053.2012.10845710

Park, H. M. (2011). Practical guides to panel data modeling: A step by step analysis using Stata.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews, 34*(6-10), 1089-1117.

Petersen, M. A. (2008). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies, 22*(1), 435-480. doi:10.1093/rfs/hhn053

Pevehouse, J. C., & Brozek, J. D. (2008). Time-Series Analysis In J. M. Box-Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford Handbook of Political Methodology* (pp. 456-474). Oxford: Oxford University Press.

Philip, M. (2015). The definition of political corruption In P. M. Heywood (Ed.), *Routledge Handbook of Corruption* (pp. 17-30). London: Routledge.

Reece, B. (2006). E-Government Literature Review. *Journal of E-Government, 3*(1), 69-110. doi:10.1300/J399v03n01_05

Roche, E. M. (2017). *Un E-Government Survey for the Period 2001-2016 - External Ex Post Facto Evaluation*.

Rose-Ackerman, S. (1978). *Corruption : a study in political economy*. New York etc.: Academic Press.

Rose-Ackerman, S. (2006). Introduction and Overview In S. Rose-Ackerman (Ed.), *International Handbook on the Economics of Corruption* Cheltenham: Edward Elgar.

Rothstein, B., & Teorell, J. (2015). Causes of Corruption. In P. M. Heywood (Ed.), *Routledge Handbook of Political Corruption* (pp. 79-94). London: Routledge.

Skog, O.-J. (2004). *Å forklare sosiale fenomener: en regresjonsbasert tilnærming* (2 ed. Vol. 7): Gyldendal norsk forlag.

StataCorp. (2017). *Stata Multivariate Statistics: Reference Manual*: Stata Press Publication.

Stepan, M., & Mokry, S. (2015). *E-Government in China: Establishing a Social Credit System in China – Realist Dystopia?*

Teorell, J., Coppedge, M., Lindberg, S., & Skaaning, S.-E. (2019). Measuring Polyarchy Across the Globe, 1900–2017. *Studies in Comparative International Development, 54*(1), 71-95. doi:10.1007/s12116-018-9268-z

Todaro, M. P., & Smith, S. C. (2015). *Economic Development* (12 ed.). Harlow, UK: Pearson Education.

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

United Nations. (2018). *E-Government Survey 2018*. Retrieved from New York:

United Nations. (2019). Frequently Asked Questions Retrieved from https://publicadministration.un.org/egovkb/en-us/FAQ

Uslaner, E. M. (2015). The Consequences of Corruption. In P. M. Heywood (Ed.), *Routledge Handbook of Political Corruption* (pp. 199-211). London: Routledge

Voigt, S. (2013). How (Not) to measure institutions. *Journal of Institutional Economics, 9*(1), 1-26. doi:10.1017/S1744137412000148

Warren, M. E. (2015). The meaning of corruption in democracies In P. M. Heywood (Ed.), *Routledge Handbook of Political Corruption* (pp. 42-55). London: Routledge

Whitmore, A. (2012). A statistical analysis of the construction of the United Nations E-Government Development Index. *Government Information Quarterly, 29*(1), 68-75. doi:https://doi.org/10.1016/j.giq.2011.06.003

Wilson, S. E., & Butler, D. M. (2007). A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications. *Political Analysis, 15*(2), 101-123. doi:10.1093/pan/mpl012

Wirtz, B. W., & Daiser, P. (2018). A meta-analysis of empirical e-government research and its future research implications. *International Review of Administrative Sciences, 84*(1), 144-163. doi:10.1177/0020852315599047

Wooldridge, J. M. (2013). *Introductory econometrics : a modern approach* (5th ed., international ed. ed.). S.l.: South-Western, Cengage Learning.

World Bank. (2019a). Metadata individuals using the internet. Retrieved from https://data.worldbank.org/indicator/it.net.user.zs

World Bank. (2019b). Metadata on GDP per capita constant 2010 $. Retrieved from https://databank.worldbank.org/data/reports.aspx?source=world-development-indicators#divMetadataCtl

World Bank. (2019c). Metadata on trade Retrieved from https://databank.worldbank.org/data/reports.aspx?source=world-development-indicators#divMetadataCtl

World Bank. (2019d). Metadata on urban population (% of total). Retrieved from https://databank.worldbank.org/data/reports.aspx?source=2&series=NY.GDP.TOTL.RT.ZS&country=

World Bank. (2019e). Metadata total natural resources rents (% of GDP). Retrieved from https://data.worldbank.org/indicator/NY.GDP.TOTL.RT.ZS?end=2017&start=2017&view=map

Yildiz, M. (2007). E-government research: Reviewing the literature, limitations, and ways forward. *Government Information Quarterly, 24*(3), 646-665. doi:https://doi.org/10.1016/j.giq.2007.01.002

Yildiz, M. (2012). Big questions of e-government research. *Information Polity: The International Journal of Government & Democracy in the Information Age, 17*(3/4), 343-355. doi:10.3233/IP-2012-000284

Zuppo, C. M. (2012). Defining ICT in a boundaryless world: The development of a working hierarchy. *International journal of managing information technology, 4*(3), 13.

8 Appendix

APPENDIX 1: Descriptive statistics

Variables	Observations	Minimum	Maximum	Mean	Standard Deviation
Political corruption	9227	.006	.976	.495	.293
Public corruption	9272	.004	.979	.474	.296
Executive corruption	9272	.011	.978	.482	.298
EGDI	1512	0	.946	.456	.22
EGDI interpolated	2688	0	.946	.46	.217
EGDI multiple imp.	2460	0	.946	.46	.218
OSI+TII	2688	0	.955	.339	.241
TII interpolated	2688	0	.935	.276	.245
OSI interpolated	2688	0	1	.402	.263
HCI interpolated	2688	0	1	.714	.214
Democracy	9291	.008	.948	.427	.289
Urban population	9708	2.077	100	48.235	24.362
GDP per capita (log)	7706	4.88	11.641	8.134	1.523
Natural resource rents (log)	6748	0	4.433	1.517	1.139
Internet users (log)	4222	0	4.598	2.041	1.593
Lightning strike density (log)	9678	-3.992	3.686	1.638	1.323

APPENDIX 2: Countries included

WENAO COUNTRIES				
Australia	Finland	Italy	Norway	United Kingdom
Austria	France	Japan	Portugal	United States
Belgium	Germany	Luxembourg	Spain	
Canada	Greece	Netherlands	Sweden	
Denmark	Ireland	New Zealand	Switzerland	

NON-WENAO COUNTRIES				
Afghanistan	Comoros	Iraq	Mongolia	Solomon Islands
Albania	Costa Rica	Iceland	Myanmar	El Salvador
Angola	Cabo Verde	Jamaica	Mauritania	Sierra Leone
Argentina	Cyprus	Korea, Rep.	Mauritius	Suriname
Armenia	Czech Republic	Kyrgyz Republic	Namibia	Slovak Republic
Azerbaijan	Dominican Republic	Libya	Philippines	Eswatini
Bahrain	Ethiopia	Lao PDR	Pakistan	Timor-Leste
Bangladesh	Fiji	Kazakhstan	Nicaragua	Tajikistan
Barbados	Egypt, Arab Rep.	Sri Lanka	Papua New Guinea	Ukraine
Belarus	Georgia	Lebanon	Panama	Turkey
Benin	Ghana	Kuwait	Oman	Togo
Bhutan	Gabon	Lesotho	Poland	Uruguay
Bolivia	Estonia	Liberia	Peru	Tanzania
Bosnia and Herzegovina	Eritrea	Cambodia	Nepal	Tunisia
Botswana	Algeria	Kenya	Niger	Uganda
Brazil	Guyana	Jordan	Saudi Arabia	Turkmenistan
Bulgaria	Guinea	Maldives	Romania	Thailand
Burkina Faso	Equatorial Guinea	Lithuania	Russian Federation	Seychelles
Burundi	Haiti	Morocco	Singapore	Zimbabwe
Cameroon	Honduras	Moldova	Sudan	Uzbekistan
Central African Republic	Croatia	Mali	Senegal	Vanuatu
Chile	Guatemala	Mexico	Rwanda	South Africa
China	Cuba	North Macedonia	Mozambique	Zambia
Colombia	Ecuador	Madagascar	Malawi	Serbia
Congo, Dem. Rep.	Gambia, The	Iran, Islamic Rep.	Malaysia	Slovenia
Congo, Rep.	Guinea-Bissau	Israel	Nigeria	Chad
Cote d'Ivoire	Hungary	Latvia	Paraguay	Venezuela, RB
United Arab Emirates	India	Malta	Qatar	Vietnam

APPENDIX 3: Regressions on subsamples. WENAO: Western Europe, North America and Oceania

	(1) All countries	(2) Non-WENAO	(3) WENAO
EGDI interpolated (t-1)	-0.0233	0.000258	-0.0775***
	(0.0177)	(0.0229)	(0.0219)
GDP per capita (t-1) (log)	0.0192	0.0205*	0.278***
	(0.0112)	(0.0111)	(0.0258)
Trade (t-1) (log)	-0.0312***	-0.0316***	-0.0658***
	(0.00442)	(0.00449)	(0.0142)
Natural resource rents (t-1) (log)	0.0146***	0.0155***	-0.00516
	(0.00279)	(0.00307)	(0.00599)
Urban population (t-1)	0.000329	0.000705*	0.00410***
	(0.000386)	(0.000370)	(0.000823)
Democracy (t-1)	-0.239***	-0.243***	0.168**
	(0.0477)	(0.0477)	(0.0720)
Year trend	-0.00277***	-0.00345***	-0.00234***
	(0.000500)	(0.000625)	(0.000447)
Constant	6.170***	7.554***	1.663**
	(0.972)	(1.218)	(0.650)
Observations	2,200	1,878	322
Number of groups	162	139	23
Time effects	YES	YES	YES
Country fixed effects	YES	YES	YES

Driscoll-Kraay standard errors in parentheses
Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX 4: Grand and petty corruption

	(1) Political corruption index	(2) Executive corruption	(3) Public corruption
EGDI interpolated (t-1)	-0.0233	0.00987	-0.0158
	(0.0177)	(0.0302)	(0.0206)
GDP per capita (log) (t-1)	0.0192	0.0248	0.0460***
	(0.0112)	(0.0189)	(0.00708)
Trade (log) (t-1)	-0.0312***	-0.0397***	-0.0180***
	(0.00442)	(0.00650)	(0.00329)
Natural resource rents (log) (t-1)	0.0146***	0.0256***	0.00803
	(0.00279)	(0.00452)	(0.00476)
Urban population (t-1)	0.000329	0.000817	-0.000569
	(0.000386)	(0.000591)	(0.00116)
Democracy (t-1)	-0.239***	-0.335***	-0.182***
	(0.0477)	(0.0573)	(0.0420)
Year trend	-0.00277***	-0.00392***	-0.00252***
	(0.000500)	(0.000727)	(0.000386)
Constant	6.167***	8.414***	5.372***
	(0.972)	(1.422)	(0.700)
Observations	2,200	2,201	2,201
Number of groups	162	162	162
Time fixed effects	YES	YES	YES
Country fixed effects	YES	YES	YES

Driscoll-Kraay standard errors in parentheses
Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX 5: FGLS estimations

	(1) RE	(2) FE
EGDI interpolated (t-1)	-0.0715	-0.0424
	(0.0471)	(0.0491)
GDP per capita (log) (t-1)	-0.0467***	0.00243
	(0.0172)	(0.0228)
Trade (log) (t-1)	-0.0284***	-0.0280***
	(0.00673)	(0.00888)
Natural resource rents (log) (t-1)	0.00942	0.00393
	(0.00631)	(0.00664)
Urban population (t-1)	-0.00103	0.000501
	(0.000927)	(0.00126)
Democracy (t-1)	-0.297***	-0.280***
	(0.0993)	(0.106)
Constant	1.272***	0.769***
	(0.102)	(0.195)
Observations	2,199	2,199
R-squared		0.190
Number of countries	162	162
Time fixed effects	NO	YES
Country fixed effects	NO	YES

Robust standard errors in parentheses
Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX 6: FGLS estimations with and without high-leverage observations

	(1) All observations	(2) Cook<4/_N	(3) Cook>4/_N
EGDI interpolated (t-1)	-0.0273	-0.0561**	0.0994***
	(0.0223)	(0.0228)	(0.0321)
GDP per capita (log) (t-1)	0.0124	0.0116	-0.0379***
	(0.0104)	(0.0114)	(0.0143)
Trade (log) (t-1)	-0.0317***	0.0157**	-0.0553***
	(0.00417)	(0.00732)	(0.00484)
Natural resource rents (log) (t-1)	0.00814*	0.00853*	0.0121**
	(0.00421)	(0.00456)	(0.00532)
Urban population (t-1)	0.000387	0.00139*	0.000800
	(0.000758)	(0.000791)	(0.00102)
Democracy (t-1)	-0.243***	-0.268***	-0.109***
	(0.0172)	(0.0189)	(0.0235)
Constant	0.672***	0.451***	1.050***
	(0.0945)	(0.108)	(0.142)
Observations	2,200	1,560	640
R-squared	0.184	0.199	0.373
Number of countries	162	134	75

Standard errors in parentheses
Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX 7: Regressions with non-transformed variables

	(1) DK	(2) DK	(3) DK	(4) DK	(5) DK	(6) DK
EGDI interpolated (t-1)		-0.355***	-0.377***	-0.010	-0.010	0.034
		(0.024)	(0.019)	(0.017)	(0.017)	(0.028)
GDP per capita (log) (t-1)	-0.000***	-0.000***	-0.000***	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Trade (log) (t-1)	-0.000***	-0.000***	-0.000***	-0.000*	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Natural resource rents (log) (t-1)	0.005***	0.002***	0.002***	0.000	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban population (t-1)	-0.001***	-0.000	0.000	0.001***	0.001***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Democracy (t-1)	-0.300***	-0.374***	-0.373***	-0.254***	-0.139	-0.169***
	(0.016)	(0.025)	(0.025)	(0.051)	(0.115)	(0.044)
Democracy^2 (log) (t-1)					-0.130	
					(0.083)	
Democracy*EGDIi (log) (t-1)						-0.248***
						(0.059)
Constant	0.763***	0.973***	0.969***	0.000	0.000	0.812***
	(0.020)	(0.016)	(0.011)	(0.000)	(0.000)	(0.068)
Observations	6,076	2,200	2,200	2,200	2,200	2,200
R-squared	0.547	0.654	0.655			
Number of groups	162	162	162	162	162	162
Time fixed effects	NO	NO	YES	YES	YES	YES
Country fixed effects	NO	NO	NO	YES	YES	YES

Driscoll-Kraay standard errors in parentheses
Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

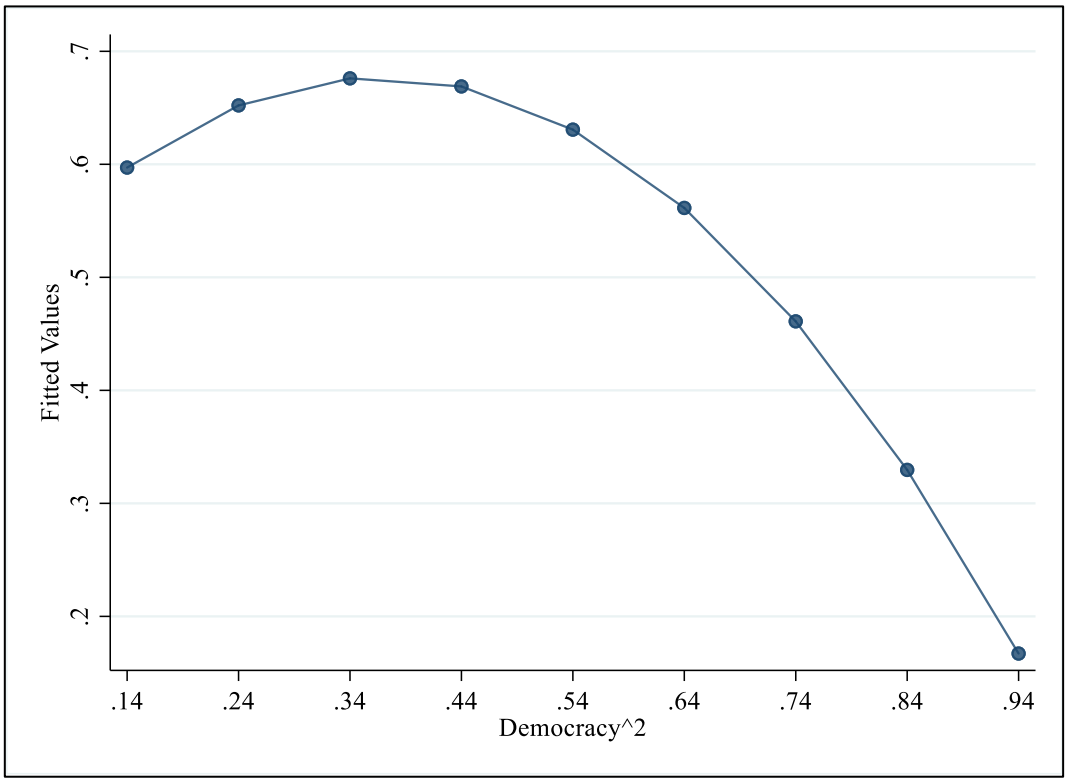
APPENDIX 8: Missing values treated with multiple imputation

	(1) FE	(2) FE	(3) RE
EGDI (t-1)	-0.006	-0.006	-0.015
	(0.017)	(0.017)	(0.017)
GDP per capita (log) (t-1)	0.012	0.012	-0.043***
	(0.023)	(0.022)	(0.016)
Trade (log) (t-1)	-0.031**	-0.031**	-0.031***
	(0.013)	(0.012)	(0.009)
Natural resource rents (log) (t-1)	0.008	0.008	0.014**
	(0.007)	(0.007)	(0.007)
Urban population (t-1)	0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)
Democracy (t-1)	-0.243**	-0.243**	-0.263***
	(0.099)	(0.095)	(0.090)
Constant	1.072***	0.666***	1.226***
	(0.161)	(0.205)	(0.108)
Observations	2,200	2,200	2,200
Number of countries		162	162
Estimator	OLS	FGLS	FGLS
Time fixed effects	YES	YES	YES
Country fixed effects	YES	NO	YES

Robust standard errors in parentheses

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

APPENDIX 9: Predictive margins for the squared term of Democracy with non-transformed variables



APPENDIX 10: Coefficients for figure 4.1 Showing the predicted effect of the squared term of democracy on political corruption

	(1)
Democracy^2	Political corruption
1bn._at	0.597***
	(0.0104)
2._at	0.652***
	(0.00635)
3._at	0.676***
	(0.00384)
4._at	0.669***
	(0.00362)
5._at	0.631***
	(0.00469)
6._at	0.561***
	(0.00566)
7._at	0.461***
	(0.00614)
8._at	0.330***
	(0.00614)
9._at	0.167***
	(0.00587)
Observations	2,200

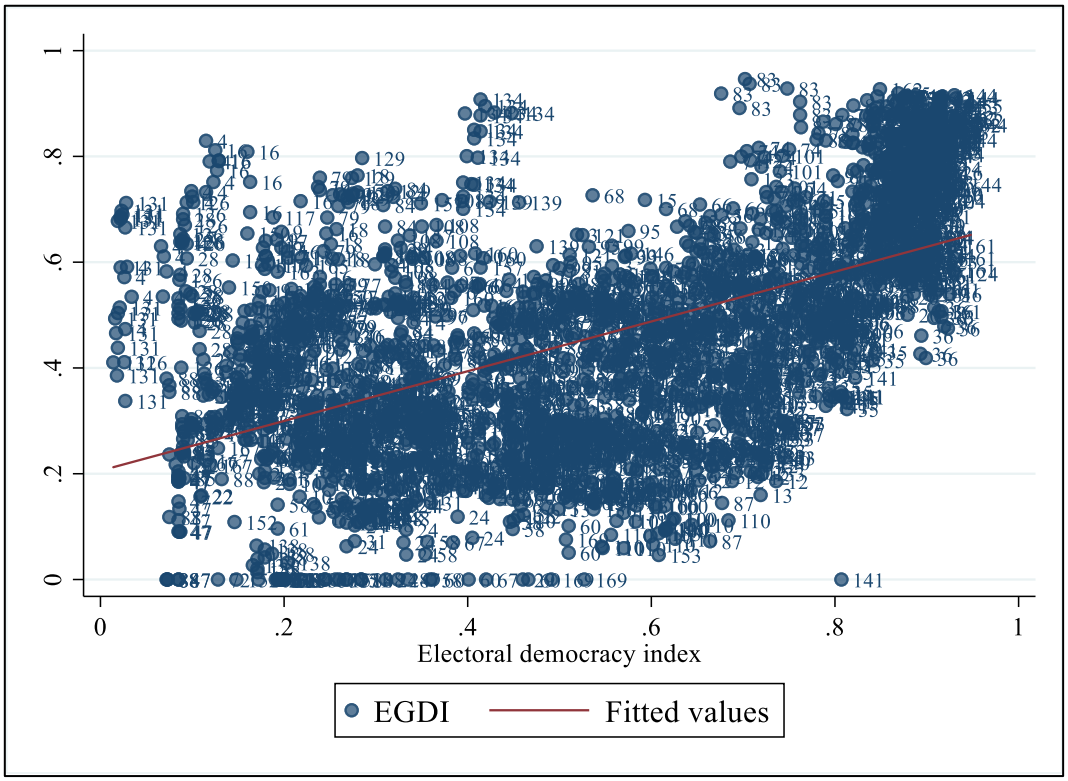
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX 11: Coefficients for figure 4.2 showing the change in y/change in x ratio

EGDI	(1) Democracy
1bn._at	-0.163***
	(0.0408)
2._at	-0.185***
	(0.0410)
3._at	-0.206***
	(0.0421)
4._at	-0.228***
	(0.0438)
5._at	-0.249***
	(0.0463)
6._at	-0.271***
	(0.0493)
7._at	-0.292***
	(0.0527)
8._at	-0.314***
	(0.0566)
9._at	-0.335***
	(0.0607)
10._at	-0.357***
	(0.0651)
11._at	-0.378***
	(0.0697)
Observations	2,200

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

APPENDIX 12: Scatterplot showing relationship between the EGDI and Democracy. Notice outliers 16 (Bahrain), 134 Singapore and 131 Saudi Arabia



APPENDIX 13: Scatterplot showing relationship between the EGDI and GDP per capita (log). Notice countries 61 Equatorial Guinea and 88 Libya