

# Do FOI laws and open government data deliver as anti-corruption policies?

## Evidence from a cross-country study

Mária Žuffová

(accepted at Government Information Quarterly)

29 April 2020

### Abstract

In election times, political parties promise in their manifestos to pass reforms increasing access to government information to root out corruption and improve public service delivery. Scholars have already offered several fascinating explanations of why governments adopt transparency policies that constrain their choices. However, knowledge of their impacts is limited. Does greater access to information deliver on its promises as an anti-corruption policy? While some research has already addressed this question in relation to freedom of information laws, the emergence of new digital technologies enabled new policies, such as open government data. Its effects on corruption remain empirically underexplored due to its novelty and a lack of measurements. In this article, I provide the first empirical study of the relationship between open government data, relative to FOI laws, and corruption. I propose a theoretical framework, which specifies conditions necessary for FOI laws and open government data to affect corruption levels, and I test it on a novel cross-country dataset. The results suggest that the effects of open government data on corruption are conditional upon the quality of media and internet freedom. Moreover, other factors, such as free and fair elections, independent and accountable judiciary, or economic development, are far more critical for tackling corruption than increasing access to information. These findings are important for policies. In particular, digital transparency reforms will not yield results in the anti-corruption fight unless robust provisions safeguarding media and internet freedom complement them.

Keywords: freedom of information; open government data; digital transparency; accountability; corruption; media and internet freedom; cross-country analysis

# 1 Introduction

Political parties have different positions on social and economic policies. They disagree on foreign relations or environmental priorities and many other issues. However, they all endorse policies that increase government transparency. Elections after elections, parties on the opposite ends of the political spectrum, have been pledging to adopt new and more ambitious transparency policies than those that were adopted by their predecessors. In 1997, the UK's Labour party promised to "clean up politics" and pass Freedom of Information Act (FOIA) (1997). Thirteen years later, the UK's Conservative Party used the very same words and pledged it would "clean up politics" by giving people a right to government data (2010: 65). Manifesto pledges are not purely symbolic (McMillan, 2018) and indeed the Labour government passed FOIA in 2000, and the 2010 to 2015 Conservative and Liberal Democrat coalition government advanced open government data publication. The UK government is one of many governments adopting transparency policies. In the past few decades, the openness has become a global norm with governments creating international coalitions to join forces in endeavours to increase access to government information. The UK and US government, leading these efforts, emphasised that transparency is an essential means to better public service delivery<sup>1</sup> and succeeded in setting it as a priority on the global agenda<sup>2</sup>.

Research exploring why governments adopt policies increasing access to government information is rich (Banisar, 2004; Berliner, 2012, 2016; Berliner & Erlich, 2015; McClean, 2011; Michener, 2010; Roberts, 2006; Schnell, 2017; Scrollini, 2015; Shkabatur & Peled, 2016). It offers different perspectives, including the role of international non-governmental organisations, political competition or domestic institutional structures as important factors affecting when and what quality FOI laws, open data or other transparency policies are adopted. However, studies that investigate if governments accomplish with transparency policies what they purport to accomplish are few. Several empirical studies explored the relationship between FOI laws and corruption levels and found inconclusive evidence (Adu, 2018; Nam, 2012; Peisakhin & Pinto, 2010; Relly & Schwalbe, 2013). While Lindstedt and Naurin (2010) observed that greater transparency is associated with lower corruption, Vadlamannati and Cooray (2017) and Costa (2013) demonstrated that the adoption of FOI laws increases corruption perception in the short term, in particular in countries with free media. Using Romania as a case study, Schnell (2014, 2017) showed that even when FOI laws are adopted for external or internal signalling, they have serious consequences for political actors and are reforms that are difficult to withdraw once passed. Comparable studies about the impact of open government data policies are lacking. The research on open data had centred mostly on its potential benefits (Noveck, 2009, 2017, 2018), potential barriers of its adoption (Janssen, Charalabidis, & Zuiderwijk, 2012; Martin, 2014; Moore, 2011; Parycek, Schöllhammer, & Schossböck, 2016), and open data use (Zuiderwijk & Janssen, 2015; Zuiderwijk, Janssen, & Dwivedi, 2015).

This quantitative cross-country study is the first to empirically investigate the relationship between open government data and corruption levels relative to FOI laws. Drawing on Bentham's (1999, 2001) political thought, principal-agent theory (Coase, 1990;

Dilulio, 1994) and research on the transparency – corruption link (Fox, 2007; Lindstedt & Naurin, 2010; Peisakhin & Pinto, 2010), I develop a theoretical framework which I subsequently test on a novel dataset. I argue that alike to FOI laws, open government data alone are not sufficient to tackle such a complex problem as corruption is. I identify media and internet freedom as critical conditions that need to be present for these transparency policies to modify corruption levels. I also control for other factors that are essential in the anti-corruption fight, in particular, free and vibrant civil society, independent and accountable judiciary, electoral freedom and economic development. I test this theoretical framework on a cross-sectional dataset compiled for this study.

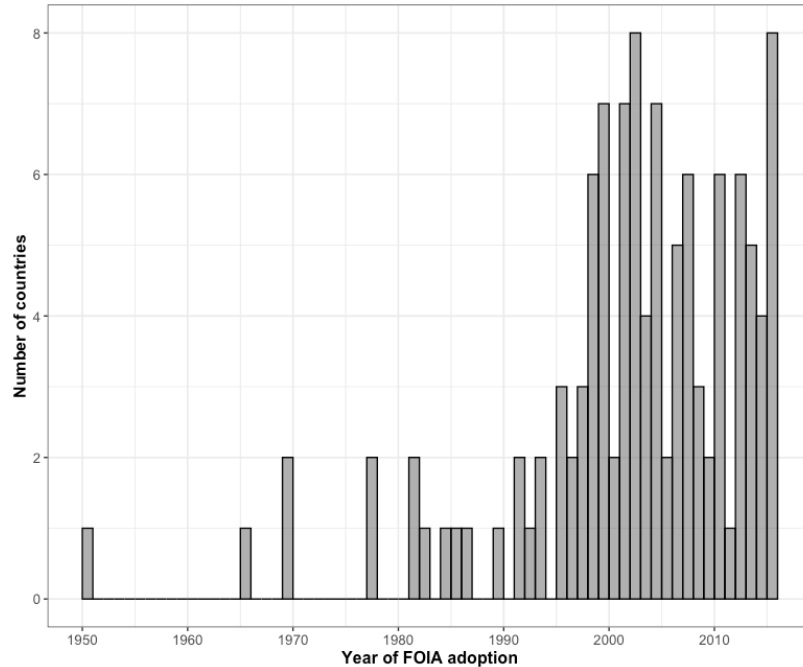
The key findings suggest that both FOI laws and open government data are significantly associated with corruption only in the interaction with media and internet freedom. Media and internet freedom are important moderators modifying the relationship between FOI laws, open data and corruption. In countries that enjoy high levels of media and internet freedom, an increase in transparency is significantly associated with a decline in corruption. From the theoretical perspective, these results further validate Lindstedt’s and Naurin’s framework and demonstrate its versatility for estimating the effect of different measures of transparency on corruption. Second, the findings suggest that enthusiasm for transparency policies as a powerful anti-corruption means, shared by political leaders and anti-corruption activists, should be restrained until more research into their impact in different regime types is conducted. In countries with restricted conditions for media operation and severe internet censorship, transparency policies might serve as smokescreen boosting the international image of governments as being transparent and progressive, shifting attention from substantial domestic policy issues. The findings also suggest that some data might matter more than others for tackling corruption. I find an association between land ownership data and corruption levels. Lastly, I demonstrate that factors other than access to government information or data are associated with corruption levels, in particular, free and fair elections, judicial independence and accountability, and economic development.

## **2 Distinctions between FOI laws and open government data**

The quest for more transparent governments is not new (Cucciniello, Porumbescu, & Grimmelikhuijsen, 2017). The fundamentals of contemporary debates about transparency have their roots in the political thought of Jeremy Bentham. He emphasised the importance of the unrestricted flow of information for compliance with moral and social norms and proposed several measures that would enable people to monitor actions of their rulers (Bentham, 1999, 2001; Schofield, 2006). He also argued that the press is necessary for the public to be able to hold the government to account, as it amplifies important political information. At present, political scientists and economists commonly agree that access to government information is a necessary condition for democracy (Stiglitz, 1999a) and development (Sen, 1999). The normative assumption that the public needs to be well-informed about the government’s actions to be able to hold it accountable was exhibited in the global spread of FOI laws in the

20<sup>th</sup> century (see Figure 2-1). The jurisprudence of supranational courts for human rights, such as the Inter-American or European Court for Human Rights has acknowledged that right of access to government information is a human right (ECtHR, 2009, 2013)<sup>3</sup>.

Figure 2-1: FOIA adoption rate from 1950 to 2016



Note: For effective visualisation of the figure, the first FOIA adopted in Sweden in 1766 was removed. The figure starts with the Finnish Act on the Openness of Government Activities, which was passed in 1951.

With new digital technologies, increasing internet penetration and levels of digital literacy, information has become available to masses. The exponential growth of computer processing power enabled governments to collect and manage an unprecedented volume of information and the concept of open government data to proliferate. Internet activists and scholars define open data as any data which is “complete, primary, timely, accessible, machine-processable, non-discriminatory, non-proprietary, and license-free” (No author, 2007). In more layman’s terms, this means that anyone can use, modify and share data for any purpose. Making government data available has been seen as the government’s obligation (Kitchin, 2014) but also as an opportunity (Mayer-Schonberger & Cukier, 2013). The public should be able to, subject to reasonable exceptions, benefit from anything that has been created using public resources, data included (Kitchin, 2014). Once the data is in the public domain, it creates prospects for multi-disciplinary mass collaboration and creation of new services and products (Surowiecki, 2005) and new opportunities for civic participation (Noveck, 2009). “Data, data everywhere”, announced the headline of the Economist’s special report in 2010<sup>4</sup>. The discourse in the media and that of civil society has been full of superlatives, referring to data as a new oil or gold and claiming it will be a new fix for long-standing social and political problems, such as corruption (OECD, 2016; Santiso & Roseth, 2017; Vrushi & Hodess, 2017), climate change, and famine (Laperrière, 2019).

There are several critical distinctions between information provided through FOI requests and made available as open data on governments' dedicated platforms. First, who makes the decision what information and when will be disclosed is different for the two. Why does it matter? This distinction has obvious consequences for what information is eventually made available and whether it is relevant for government accountability. As for FOI requests, the public authority is usually obliged to disclose within legally set time limits any requested information which is of public interest and not a subject of exemptions from FOIA. If the authority withholds information unlawfully, a requester has several legal means to challenge this decision. Proactive publication of open government data relies on the authority's goodwill. Fung (2013) defined this new digitally-enabled transparency as the information on tap in contrast to the information on demand. Shkabatur (2012) referred to the publication of government datasets as discretionary transparency. Both these definitions are eloquently accurate, as it is at the government's discretion to decide what datasets it will make publicly available. Noveck (2017) defined the information provided through FOI requests as ex-post and as open data as ex-ante. She argues that this shift to ex-ante puts more emphasis on collaboration and innovation rather than monitoring and accountability. She proposed that the ex-ante publication makes open data less adversarial tactics than FOI. Therefore, it is not surprising that governments have generally welcomed the shift towards greater proactive disclosure of information as a more cooperative and less oppositional approach.

However, collaborations require reciprocity and equality. O'Neill (2006) argued that open data represents a heavily one-sided way of interaction where the government directs the information flow. The question then stands, if governments are in a position of information suppliers and decide what data will be published, what reason do they have to release any information that will put them into a bad light? Samaha (2006: 918), drawing from Stiglitz, argued that governments will always prefer to disclose "information that makes the administration look public-spirited, effective, and efficient, but withhold information to the contrary". With open data, governments can easily do it as there is no legal requirement for them to respond to public requests for specific datasets.

There are other distinctions between the access to government information under the FOI regime, and proactive publication of open data and some of them are outlined in Table 2-1 using the UK as an example. However, the distinction above, I argue, is relevant for the potential of these transparency policies to deliver as effective anti-corruption measures. While potentially high political costs are attached to the adoption of FOI laws or asset disclosures for high-level officials as Schnell (2017) demonstrated, the adoption of open data policies presents a lesser risk for politicians pushing them forward. Compared to FOI laws, open data policies can sustain to be a cheap anti-corruption lip service without major consequences for political actors. High-corruption countries, in particular, might be tempted to adopt them for signalling purposes or to attract international funding, but then might not be able or willing to sustain a regular publication of datasets relevant for the government oversight (Shkabatur & Peled, 2016).

Table 2-1: Key distinctions between FOI and open government data in the UK

	FOI	open data	
Accessibility in terms of time	statutory time limits	immediate	once published
Availability of an offline equivalent	yes	no	
Legal framework	FOIA/FOISA	RPSI*	
Legal safeguard of information provision	public interest test	none	
Availability of appeal mechanisms	yes	no	
Oversight body	yes	no	

\* The Reuse of Public Sector Information Regulations 2015

### 3 FOI laws, open data and corruption: Identifying the missing links

The view of transparency as a means to the accountable government goes back to Jeremy Bentham’s political thought<sup>5</sup>. Bentham believed that information flows are at the heart of citizen-government relations and if unrestricted, can serve as an essential government’s corrective mechanism. From a principal-agent perspective, increasing access to government information decreases information asymmetry between the government (agent) and the public (principal). As a result, it creates more favourable conditions for citizens (principals) to hold their elected representatives accountable (Dilulio, 1994) and make well-informed choices and decisions (Heald, 2006; Stiglitz, 1999b). Many scholars already demonstrated the importance of the access to information for public service delivery, accountability, but also a political competition (Berliner, 2014; Berliner & Erlich, 2015; Besley & Burgess, 2002; Reinikka & Svensson, 2005). Hence, Stiglitz’s (1999a) claim that information is a public good benefiting all members of society is not in the slightest an exaggeration.

This study uses the theoretical framework introduced by Lindstedt and Naurin (2010) to explain the relationship between open government data, relative to FOI laws, and corruption. The definition of corruption for purposes of this study includes petty corruption among citizens in addition to the abuse of a public office for private gain (Mauro, 1995). Lindstedt and Naurin (2010) proposed that increasing access to information makes undertaking fraudulent behaviour more complicated and dangerous. For instance, if the FOI legislation is robust and well-implemented and enforced, then political leaders engaged in fraudulent behaviour have to make an additional effort to cover it up in a more sophisticated way. It might discourage them from corrupt practices altogether. In this case, greater access to government information or data serves to prevent corruption. However, when investigative journalists use this information, it can also act as a corrective measure pointing out to already committed infringements and calling for justice. Either way, drawing from principal-agent theory (Coase, 1990; Dilulio, 1994), there is an assumption that greater access to government information or data will be associated with lower levels of corruption.

Lindstedt and Naurin (2010) argued that, in addition to access, two other conditions – the publicity and accountability condition – must be fulfilled for the information to affect

corruption. The publicity condition means that citizens (principals) are able to receive the information that has been released about the misbehaviour of their representatives (agents). The ability of information reception does not stand here only for having technology but also having skills (both technical and literacy competencies). As Margetts (2006: 203) eloquently wrote, “digital government can only be transparent to a digital society”. However, the available research on open data (Birchall, 2015; Dai & Li, 2016; Lourenço, 2015; Worthy, 2015a) suggests that only a small fraction of the population is able to access, and analyse, and thus, benefit from published government datasets. Most citizens do not have time, capacity, and interest to browse the government’s accounts at their leisure (Munro, 2015; Wheeler, 2012; Worthy, 2015a, 2015b), leaving “any army of armchair auditors”<sup>6</sup> too few in numbers or poorly equipped to win any major battle. Likewise, the legal jargon in FOI requests might act as a discouragement to submit a request. The time frames for responding also affect the size and composition of the audience and result in non-egalitarian distribution of information provision (Pozen, 2017). Therefore, the mere availability of government information or data is insufficient to tackle corruption (Fox, 2007; Kolstad & Wiig, 2009; Sharafutdinova, 2010).

Lindstedt and Naurin (2010) posited that mediators, such as mass media or civil society organisations, are required to transmit important messages to the public. First, for the media to be able to fulfil their intermediary role, media penetration has to be high (Besley & Burgess, 2002; Ferraz & Finan, 2008; Reinikka & Svensson, 2005). As the contemporary media is complex, this includes but is not limited to the levels of news circulation, TV and radio transmission, or internet penetration. Second, the media act as watchdogs. If political leaders have to operate in the free and pluralistic media environment, and they want to retain their positions, they cannot afford to be identified with fraud, nepotism or any other serious misconduct. If media disseminates any reputation-harming information, their chances for re-election are very likely to drop. For example, Ecker, Glinitzer, & Meyer (2016) found that voters who perceive corruption in a country as high punish incumbents. More importantly, voters seem to punish corrupt politicians when they have sufficiently concrete and reliable information about corruption (Winters & Weitz-Shapiro, 2013).

However, the information flow is possible only in the media environment with unrestricted conditions for operation, where media are free from censorship, restrictive libel laws or editorial interference from the government or private owners. Scholars have demonstrated that governments are more responsive to their constituents in countries with higher levels of media freedom (Besley & Burgess, 2002; Besley, Burgess, & Prat, 2002; Besley & Prat, 2006; Brunetti & Weder, 2003; Chowdhury, 2004; Djankov, McLiesh, Nenova, & Shleifer, 2003; Ferraz & Finan, 2008; Norris, 2008; Reinikka & Svensson, 2005; Solis & Antenangeli, 2017; Svensson, 2005; Yazaki, 2017). However, government ownership or control is not the only challenge journalists face that may affect editorial content (Hallin & Mancini, 2012; Voltmer, 2013; Waisbord, 2002). If journalists face danger for doing their daily work, such as attacks and harassment, they might self-censor themselves. In line with the above, I argue that unless the publicity condition is fulfilled, access to politically important information through FOI laws and open data will not be a game-changer in fighting corruption.

That said, the advancements in digital technologies have not only changed the government information strategies and the amount of government information in the public domain, but it has also transformed media. In addition to traditional media, now citizens can actively participate in news creation and dissemination. Therefore, if a citizen spots a dubious tender in the publicly available procurement data, she can inform about it on social media, her blog or through her newsletter. The role of social media as information sources should not be underestimated, given the largest platforms have billions of monthly active users<sup>7</sup> (Esteban, 2019). However, traditional media still act as key gatekeepers of citizen-generated content in non-democratic countries (Ali & Fahmy, 2013) as well as democratic ones (Hermida & Thurman, 2008). In non-democratic regimes, it is unlikely that citizen-generated content disapproving the government actions will make it to the print or screens of national television. However, if the approach to the internet censorship in these countries is more relaxed than to media censorship, citizens might be still able to disseminate politically important information through social media and use messaging applications for connecting and mobilising others and coordinating further actions to voice their discontent. Scholars documented the use of social media by citizens and activists in different non-democratic regimes for what essentially was a citizen political journalism, e.g. during the Iranian Green Movement that emerged after the rigged 2009 presidential election, 2011 Egyptian and Libyan revolutions against the oppression of their dictatorial leaders (Ali & Fahmy, 2013; Howard & Hussain, 2013b; Tufekci & Wilson, 2012).

However, as the use of social media for citizen journalism becomes more sophisticated, so do government strategies to counteract any dissent. Some governments, such as the most skilled internet censor China, are able to react immediately to the undesirable content and remove it and disrupt online networks. However, others are less successful in controlling the network, which leaves them with having to fight the protesters on the streets or shut the whole population off from the internet, which in some cases, e.g. in Egypt under Mubarak, led to even more forceful protests (Howard & Hussain, 2013a). If the government does not have strong control over the internet and its response in the streets is violent, even though citizen-generated content will unlikely appear in traditional media, it can easily reach the international audience and elicit a response from other governments through their foreign policy. Therefore, in some contexts, it is possible that even if the media is heavily censored, citizens are still able to share information online.

Lastly, the accountability condition means that some redress mechanisms are in place and enable that corrupt agents are held accountable (Lindstedt & Naurin, 2010). A free and fair election is the most basic accountability mechanism, but many scholars argue that although it is a necessary one, it is not sufficient for different reasons. First, the elections happen once in a few years only, leaving political leaders to choose to be unresponsive in between the elections. Second, voters, given they possess all critical information, have still only one vote to hold accountable several public officials for their performance in many different fields. Also, as Smulovitz and Peruzzotti (2000) rightly conclude, voting is an individual and private act, and voters do not coordinate with each other to punish a specific politician by voting him out of the office.



While electoral democracy serves as an accountability mechanism despite its limitations, constitutional equality before the law presents a stronger accountability mechanism. Information obtained through FOI requests and greater access to government datasets may help to uncover irregularities in awarding contracts, favouring certain suppliers, but it can serve as an effective anti-corruption measure only if those responsible are sanctioned. On the contrary, knowledge about corrupted behaviour and impunity may lead to a higher perception of corruption, greater social acceptance of it and higher public distrust. Many scholars argue that prosecutors and judges play a significant role in pursuing the rule of law and government accountability. Nonetheless, to be able to play this role well, they must be independent and decide cases “free from the coercion, blandishments, interference, or threats from governmental authorities or private citizens” (Rosenn, 1987: 7). Larkins (1996) emphasises that judicial independence guarantees equality before the law for everyone, members of the government included.

Based on the theoretical framework presented above, I propose a set of hypotheses about the relationship between open government data, relative to FOI laws, and corruption:

**h1a:** Countries with greater availability of open government data are more likely to exhibit lower levels of corruption than countries with limited availability.

**h2a:** The relationship between open government data and corruption is conditional upon the quality of media (and internet) freedom.

**h3a:** The relationship between open government data availability and corruption is conditional upon the degree of judicial accountability and independence.

**h1b:** Countries with a long tradition of FOI laws are more likely to exhibit lower levels of corruption than those where FOI legislation was introduced just recently.

**h2b:** The relationship between FOI laws and corruption is conditional upon the quality of media (and internet) freedom.

**h3b:** The relationship between FOI laws and corruption is conditional upon the degree of judicial accountability and independence.

While I expect the effects of both FOI laws and open data on corruption to be conditional on media and internet freedom, and judicial accountability and independence, I argue that there will be a difference between the strength and size of the effects of open government data, relative to FOI laws. As access to government information under the FOI regime is stronger than through the publication of government datasets, I also expect its effects on corruption will be greater.

## 4 Data and measurement

Since data on the availability of open government data exists only for years from 2013 to 2016<sup>8</sup> and the variation of data on corruption levels is minimal within this period; longitudinal analysis was excluded as a possible estimation method. Due to data limitations, instead, cross-sectional analysis using the latest available data from 2016 was employed.

#### 4.1 Dependent variable

Low levels of corruption are a sign of governments which are responsive to their constituents and responsibly governs public resources. In this study, I use the Control of Corruption (CoC) score<sup>9</sup> from the Worldwide Governance Indicators as a measure of corruption. It is an aggregate and robust measure compiled from a variety of indices, such as Afrobarometer, Latinobarometro, Transparency International's Global Corruption Barometer (GCB) and Corruption Perception Index (CPI), World Economic Forum Global Competitiveness Report and others. The estimate ranges from -2.5 to 2.5. A negative score indicates high, and a positive score indicates low levels of corruption. In 2016, corruption was highest in Equatorial Guinea with the -1.81 estimate and lowest in New Zealand with the 2.28 estimate.

CoC score captures both petty and grand forms of corruption, as well as state capture by elites and private interests. It is a robust measure that can withstand the criticism that is often raised of the components of CoC individually (Andersson & Heywood, 2009; Heywood & Rose, 2013). Measures of corruption perception, in particular, CPI is criticised mostly for its expert nature. Scholars argued that expert evaluators might only repeat common knowledge about particular countries and perpetuate labelling them corrupt or non-corrupt even when the situation has already changed (Heywood & Rose, 2013; Rose & Mishler, 2010). Other scholars (Mills, 2017; Sharafutdinova, 2010) made an important point that in hybrid regimes, (false) accusations of corruption are often used to gain an advantage or eliminate political competition, which results in driving the overall public perception of corruption in a country up. CPI also captures grand corruption mostly, i.e. cases of overpriced tenders in the state administration, but does not reflect petty corruption among citizens, for example, to access essential public services. However, as available research in the field demonstrated, although grand and petty corruption is connected more broadly, discrepancies between them might be substantial in some cases. While citizens might have a high perception of corruption if they are commonly informed about poor management of public resources, they might still have good access to public services without relying on bribing (Heywood & Rose, 2013; Rose & Mishler, 2010). At the same time, however, indices measuring actual corruption suffer from flaws of other types. For instance, GCB, the largest world-wide public opinion survey on corruption, is likely to suffer from the compromising effects of social desirability bias. Respondents might not tell about the bribes they had to pay. They might be reluctant to speak the truth, fearing potential repercussions for their dishonest conduct or simply not wanting to be associated with socially unacceptable behaviour (Arnold & Feldman, 1981; Chung & Monroe, 2003). Using CoC for the analysis is, thus, a reasonable choice, as aggregating several measures of corruption can partially compensate for their individual imperfections.

#### 4.2 Independent and control variables

The availability of open government data is measured as the Implementation sub-index of Open Data Barometer, an index designed by the World Wide Web Foundation in 2013. It is a continuous variable measuring the availability of key government datasets in open formats. The score scale ranges from 0 to 100. In 2016, the lowest achiever was Mali with an 8.67 score, and the highest achiever was the UK with a 91 score. Open Data Barometer's sub-index

examines the availability of detailed government budget data, spending data, data on public contracts, company registration data, international trade data, map and land ownership data, national environmental statistics, public transport timetable data, crime statistics, detailed census data, legislation and election results data, education performance data and health sector performance data in open formats (World Wide Web Foundation, 2016).

The differences in the open government data availability score between individual datasets are substantial. The highest mean of data availability score for all countries is for census data ( $\mu = 53.91$ ,  $Md = 65$ ,  $n = 115$ ), the lowest mean is for spending data ( $\mu = 11.87$ ,  $Md = 5$ ,  $n = 115$ ). As can be observed in the appendix in Figure 1-1, democratic regimes release more data on average. However, all countries are reluctant to release data on contracts, company registration data, and land ownership data, i.e. data that are of utmost importance for monitoring the use of public resources and exposing corrupt behaviour.

Open Data Barometer offers broad geographical coverage, and from available open data indices, has the most stable methodology over time. Nonetheless, it has its limitations too. While the measure evaluates against the open data criteria, i.e. whether data is machine-readable and reusable, openly licensed, regularly updated, provided free of charge, Open Data Barometer does not assess reliability, accuracy or quality of the data. Therefore, taking Open Data Barometer at face value assumes trust that the government has been collecting and managing its data rigorously and also providing it in its entirety in the public domain.

However, this might not always be the case. Some governments might be tempted to censor or doctor the data that could reveal their poor performance in certain areas. In summer 2019, Brazilian President Jair Bolsonaro suggested that the data published by the National Institute for Space Research, the government agency in charge of monitoring deforestation, will require prior approval from the administration following the publication of figures that showed 88% rise in deforestation in a year<sup>10</sup>. Also, the example of Tanzania which recently amended its Statistics Act to restrict disseminating any statistical data distorting or discrediting the official statistics, and thus limited their citizens' ability to question official government data, may suggest why cautiousness about data quality is substantiated (World Bank, 2018). Other infamous cases in the past also suggest that there is a reason for scepticism about reliability, quality or entirety of data provided by the government.

The tradition of FOI laws is measured as the number of years since FOI legislation has been adopted and enforced. It is a continuous variable, ranging from 0 for countries that do not have FOIA in place to 250 for the Swedish FOIA that has been adopted in 1766. The length of the FOI law tradition in a country is a useful proxy for measuring the openness of public authorities and the implementation of FOI laws. For instance, Scandinavian governments that were among the firsts to adopt FOI laws are well-known for their positive approach to transparency, which they perceive primarily as public access to government information (Grønbech-Jensen, 1998). Having decades-long experience with FOI laws is also likely to affect their implementation and enforcement positively. Measuring the impact of FOI laws through the length of their existence has, of course, limitations. For instance, it can be conflated with the overall democratic tradition in a country. However, in light of the absence of a cross-country measure of the quality of the implementation of FOI laws, the tradition of

FOI laws is the best available measure. Although a measure of the quality of the FOI laws' design exists (RTI Rating), the practice can be miles away from the letter of the law. In other words, de jure FOIA does not necessarily reflect de facto FOIA. For instance, based on the RTI Rating, Afghanistan, classified as a not free country, has the most comprehensive FOI legal framework. With the elections accompanied by frauds, frequent violent assaults on journalists, and high levels of corruption, it is unlikely that common bureaucratic application of the new FOI law holds to its standards written in the law.

Table 4-1: Top twenty-five RTI Rating countries by their FH status

	RTI Rating	FH Status
1. Afghanistan	139	Not free
2. Mexico	136	Partly free
3. Serbia	135	Free
4. Sri Lanka	131	Partly free
5. Slovenia	129	Free
6. Albania	127	Partly free
7. India	127	Free
8. Croatia	126	Free
9. Liberia	124	Partly free
10. El Salvador	122	Free
11. Sierra Leone	122	Partly free
12. South Sudan	120	Not free
13. Tunisia	120	Free
14. South Africa	119	Free
15. Vanuatu	119	Free
16. Maldives	116	Partly free
17. Azerbaijan	115	Not free
18. Saint Kitts and Nevis	115	Free
19. Antigua	113	Free
20. Kenya	113	Partly free
21. Nepal	113	Partly free
22. Macedonia	112	Partly free
23. Ethiopia	111	Not free
24. Nicaragua	111	Partly free
25. Moldova	110	Partly free

Note: RTI – Right to Information, FH – Freedom House

For measuring media freedom, I use the Freedom House's Freedom of the Press Index and the number of murdered journalists in a given country in a given year. The Freedom of the Press Index is based on experts' survey responses. Experts, usually media scholars assess legal, political and economic media environment in a particular country. The range of the score goes from 0, indicating that a country enjoys the highest levels of media freedom to 100, which suggest heavy restrictions and censorship. In 2016, the top achiever was Norway, and the lowest achiever was North Korea (Freedom House, 2017).

The Committee to Protect Journalists have collected and published data on assaults on journalists and the press for more than 30 years. They keep evidence of killed and imprisoned journalists, as well as journalists who went missing. The Committee also collects data on the causes of the death, whether it was a murder or journalist died in crossfire or combat, or on a dangerous assignment. Since the death of journalists in conflict zones is unlikely to be related to corruption, I included only murders in the analysis. This variable has additional explanatory power, as Pearson's correlation coefficient between Freedom of the Press Index and murders of journalists was very low (0.15).

In addition to media freedom, internet freedom is also considered. I use the Freedom House's Freedom on the Net Index, which is also based on experts' survey responses. It assesses obstacles to the internet access, limits on content including technical filtering and blocking of websites, self-censorship, and use of social media for civic engagement, and user rights violations online but also offline repercussions for online activities, e.g. harassment or detention. The score range goes from 0, indicating that a country enjoys the highest levels of internet freedom to 100, which suggest severe internet censorship. In 2016, the top achiever was Iceland and Estonia, and the lowest achiever was China (Freedom House, 2016).

Civil society variable is measured as an additive score created from three variables from the Varieties of Democracy (V-Dem) dataset: v2cseeorgs, v2csreprss and v2csprtcpt (Pemstein, 2015). The range is from 0, indicating severely restricted civil society to 11 indicating free civil society. The first variable v2cseeorgs measures the extent of government's control over CSOs entry and exit, and its range is from 0 which represents monopolistic control to 4 which indicates that the government does not hinder formation and operation of CSOs. The second variable v2csreprss measures the extent of the government's repression of CSOs. It is an ordinal variable with five levels where 0 indicates severe repression of civil society aiming for its elimination and 4 means CSOs are free to organise without fearing repercussions. The last variable v2csprtcpt quantifies the extent of people's engagement in CSOs, and its range is from 0 where associations are government-sponsored, and engagement is often compulsory to 3 where CSOs are numerous and active in diverse agenda, and people are involved at least in one of them.

Judicial accountability and independence are also measured as an additive score created from two variables from the Varieties of Democracy (V-Dem) dataset: v2juacct and v2jucorrdc (Pemstein, 2015). The range is from 0, indicating low judicial accountability and independence to 8 showing the opposite. The first variable v2juacct measures judicial accountability, i.e. how often are judges removed from their posts or otherwise disciplined if they are found responsible for a serious misdemeanour. It is an ordinal variable with five levels (never = 0, seldom = 1, about half of the time = 2, usually = 3, always = 4). The second variable v2jucorrdc assesses judicial independence, i.e. how often do individuals or businesses pay bribes to speed up or delay the process, or to obtain a favourable judicial decision. It also is an ordinal variable with five levels (always = 0, usually = 1, about half of the time = 2, not usually = 3, never = 4) ordered in reverse order.

Table 4-2: Descriptive statistics

Variable	N	Mean	St. Dev.	Min	Max
Control of Corruption	193	-0.1	1	-1.8	2.3
Corruption Perception Index	172	42.8	19.3	10	90
Bribery rates	110	22.9	17.1	0	77
Open government data	115	32.4	18.7	6	91
Budget data	115	40.4	32.2	5	95
Spending data	115	11.9	21.8	0	95
Data on contracts	115	23.7	25.6	0	95
Data on companies	115	20.5	26.4	0	100
Land ownership data	115	13.6	19.3	0	90
Years of FOIA	196	9.2	20.5	0	250
Press freedom	195	48.5	23.8	9	97
Murdered journalists	196	0.1	0.3	0	2
Civil society	171	7.8	2.9	0	11
Judiciary	171	4.3	1.8	0	8
Free and fair elections	171	59.6	30.1	0	98
GDP (in 1000 USD)	190	13.3	20.5	0.25	165

Free and fair elections are also measured using a variable from the V-Dem dataset, which defines free and fair elections as the elections free from registration fraud, systematic irregularities, government intimidation of the opposition, vote-buying, and election violence (Pemstein, 2015). The variable is an aggregate measure consisting of other V-Dem indicators. It is an interval variable ranging from 0 to 1, converted to 0 to 100 scale for better interpretability of the results. In the literature, the free and fair election is described as the essential government accountability mechanism (Smulovitz & Peruzzotti, 2000).

GDP per capita data (in 1000 USD) from the World Bank is used as a control variable as previous research in the field has demonstrated that the level of economic development and growth matters for an absence of corruption. GDP can also serve as a proxy for assessing wider living standards and literacy, as it is associated with other development measures, such as the Human Development Index (HDI).

### 4.3 Estimation method and addressing endogeneity concerns

Linear models were used to test for a significant effect of FOI laws and open government data and other factors on corruption levels. I controlled for the presence of free and fair elections, and gross domestic product per capita. All final models for estimating the associations between FOI laws, open government data and corruption levels were visually checked (diagnostic plots) for normal distribution of residuals, constant variation (homoscedasticity), influential observations and independence of residuals. In addition, Shapiro-Wilk (a formal test of the normal distribution of residuals) and Breusch-Pagan (a formal test of heteroskedasticity) tests were conducted. These formal tests, together with a visual examination of diagnostic plots, suggested that the models meet the assumptions of general linear models. For a detailed

description of these tests and diagnostic plots, see Appendix 1.2. All analyses were conducted in R (v3.5.1).

Previous research on transparency and corruption often mentions potential endogeneity problems. Scholars emphasised that many corruption measures, in this analysis CPI as an element of CoC, in particular, are constructed through eliciting opinions of experts, which is an imperfect measure. Although experts are knowledgeable of country contexts, Peisakhin and Pinto (2010) argued they might fail to untie the nature of the causal relationship between policy and changes in behaviour. Endogeneity problems may occur, i.e. corruption may well be caused by factors other than an identified predictor, but which are correlated to that predictor. This might lead to biased ordinary least squares (OLS) estimators for the effect of the predictor, as they will absorb the effect of omitted variables. Reverse causality and measurement errors may also cause endogeneity. For example, one may argue that less corrupt countries, might self-select themselves to pass FOI laws and open up government data. Thus, the causal relationship between FOI laws and open government data on one hand and corruption levels on the other might well run in the opposite direction. While this is plausible, and less corrupt countries might opt-in for more transparency, more corrupt countries might be motivated equally if not even more, to pass but not properly implement FOI laws or increase access to government data to signal they are too transparent and accountable. Schnell (2017) demonstrated that this was the case in Romania, where levels of corruption are high. Scholars emphasised that, in particular, open data can easily be misused for these purposes as they do not require a hard-accountability commitment (Peixoto, 2013; Yu & Robinson, 2012).

On the empirical level, the issue of endogeneity is addressed by identifying the instruments and thus accounting for the effect of omitted variables, potential reverse causality and measurement errors (Angrist & Pischke, 2015). If an exogenous instrument for open government data measure can be identified and open government data maintains a significant coefficient when instrumented in the second stage of the instrumental variable (IV) estimation, it can be concluded that open government data affects corruption rather than vice versa. Sovey and Green (2011) argued that if instrument variables are not formed through random assignment, as is the case in the experiments, the good practice for observational studies is to provide substantial theoretical justification for why the selected instrument is uncorrelated with the error term, and hence with the dependent variable. At the same time, it should be a good predictor of the endogenous independent variable in question.

I identified an instrumental variable that fulfils criteria – membership in the Open Government Partnership (OGP). The OGP’s role is to secure actionable commitments from national governments to increase transparency, accountability, and public participation. Governments draft national action plans to formulate their commitments. So far, more than one-sixth of them were related to open data. I argue that the membership in OGP affects domestic politics and contribute to the global diffusion of open data policies since OGP has monitoring mechanisms in place to assess how governments fulfil their pledges and implement policies to which they committed. As a result, open data commitments in most countries eventually translate into higher availability of government datasets. Nonetheless, at the same time, the membership in OGP does not make countries free from corruption. The CoC

estimates for OGP member countries ranged in 2016 from -1.03 score in Nigeria to 2.28 score in New Zealand. In fact, out of 75 OGP member countries in 2016, the majority (42) had a negative CoC estimate, i.e. were considered rather corrupt than non-corrupt. Also, the eligibility criteria for OGP membership are relaxed and aspiring countries can formally meet them relatively easily, which is why I argue that the OGP membership is, to a great extent, exogenous and detached from other variables. In 2016, some OGP members did not represent democratic regimes, infringed political rights and civil liberties, and were known for restricting journalists in their work. Some did not have FOI laws.

## 5 Results

The results are presented as follows. The associations of open government data and corruption and FOI laws and corruption measured as CoC, and the moderating effect of press and internet freedom are discussed below (see Table 5-1, Table 5-2, Table 1-1). All models are estimated using ordinary least square and open data models also two-stage least square regression to address potential endogeneity problems. Based on the theoretical framework proposed above, the following models were estimated:

**Model 1:**  $\text{Corruption} = \alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \epsilon$

**Model 2:**  $\text{Corruption} = \alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \beta_{\text{open data} \times \text{press freedom}} + \epsilon$

**Model 3:**  $\text{Corruption} = \alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \beta_{\text{open data} \times \text{judiciary}} + \epsilon$

**Model 4:**  $\text{Corruption} = \alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \beta_{\text{FOI years} \times \text{press freedom}} + \epsilon$

**Model 5:**  $\text{Corruption} = \alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \beta_{\text{FOI years} \times \text{judiciary}} + \epsilon$

### 5.1 Open data and the moderating effects of press and internet freedom

First, I will describe the results from models estimating the effect of open government data on corruption levels. As can be observed from Table 5-1, except for model 3, the findings are consistent with h1a hypothesis that greater availability of government data is significantly associated with lower levels of corruption. In model 1, holding all other predictors constant, for every score increase in open government data score, CoC score increases by 0.007 scores, i.e. levels of corruption are lower.



Table 5-1: The associations between open government data and corruption (OLS)

Control of corruption			
	Model 1	Model 2	Model 3
Open data	0.007* (0.003)	0.022*** (0.006)	-0.004 (0.007)
Years of FOIA	0.002 (0.002)	0.001 (0.002)	0.002 (0.002)
Press freedom	-0.004 (0.004)	0.008 (0.005)	-0.003 (0.004)
Murdered journalists	-0.269* (0.113)	-0.204 (0.110)	-0.266* (0.112)
Civil society	-0.073** (0.025)	-0.070** (0.024)	-0.072** (0.025)
Judiciary	0.216*** (0.032)	0.212*** (0.031)	0.153** (0.051)
Free and fair elections	0.008** (0.002)	0.010*** (0.002)	0.008** (0.002)
GDP	0.020*** (0.003)	0.017*** (0.003)	0.019*** (0.004)
Open data*press freedom		-0.0004** (0.0001)	
Open data*judiciary			0.002 (0.001)
Constant	-1.132** (0.400)	-1.833*** (0.443)	-0.936* (0.416)
N	110	110	110
R <sup>2</sup>	0.879	0.890	0.882
Adjusted R <sup>2</sup>	0.870	0.880	0.872
Residual Std. Error	0.364 (df = 101)	0.349 (df = 100)	0.362 (df = 100)
F Statistic	92.019*** (df = 8; 101)	90.216*** (df = 9; 100)	83.295*** (df = 9; 100)

\* p < .05; \*\* p < .01; \*\*\* p < .001

Interestingly, freedom of the press was also not significantly associated with corruption in any model. However, in model 1 and 3, the importance of media freedom for uncovering and tackling bribe and corruption are demonstrated through the variable measuring the numbers of murdered journalists. Every additional murder of a journalist is significantly associated with an increase in corruption levels (decrease in CoC score) by almost 6% in both models ( $p < 0.05$ ). In a plain language, the risk of death linked to investigative reporting might put off other journalists from investigating and informing about corruption scandals and help corruption to flourish further.

Also, all models (1-3) revealed the importance of the independent and accountable judiciary and the presence of free and fair elections and economic growth for tackling corruption. Holding all other predictors constant, for every score increase in the independence of judicial system CoC score increases by 0.15-0.22 scores, which on the scale from -2.5 to 2.5

represents a substantive increase (4%). GDP per capita is also a statistically significant predictor of corruption with high-income countries being associated with lower corruption. Every additional 1000 USD of GDP per capita is associated with approximately 0.02 increase on the CoC scale. An increase in free and fair elections score is also significantly associated with lower corruption levels. The relationship between free civil society and corruption levels goes surprisingly in the opposite direction, i.e. freer civil society is significantly associated with higher levels of corruption.

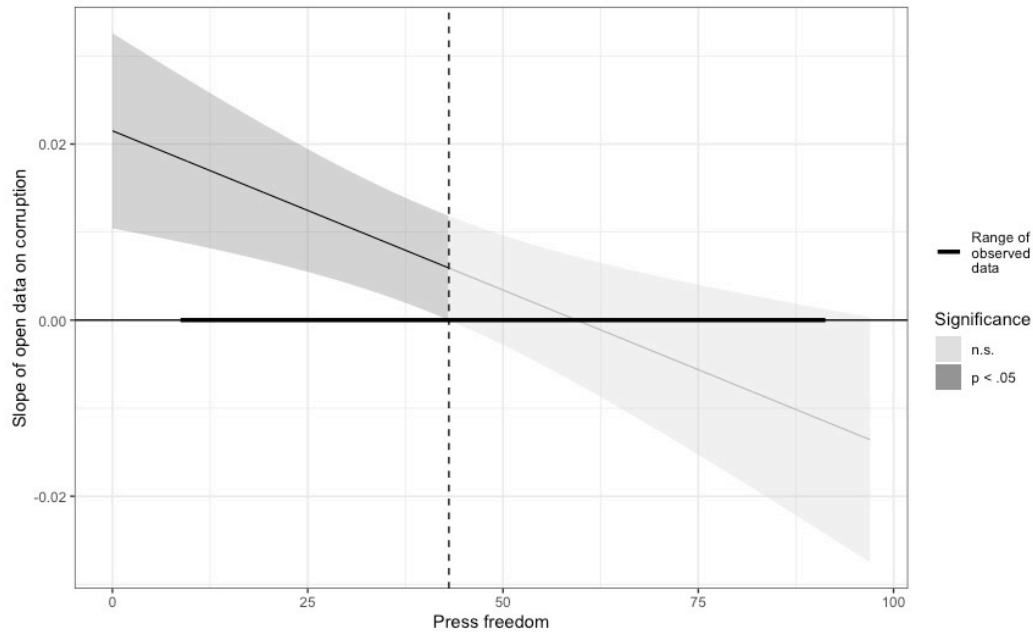
Models with added interaction effect (model 2 and 3) yield further interesting results. The moderating effect of press freedom (model 2) modifies the size and strength of the association between open government data and corruption and is consistent with the proposed theory. The association of open data and corruption is greater in the model with press freedom as a moderator. In model 2, for every score increase in open data availability, CoC score increases by 0.02 scores (robust standard error = 0.006,  $p < 0.001$ ), i.e. levels of corruption go down. Substantively speaking, this is a 0.4% increase on the scale of CoC (from -2.5 to 2.5). As can be observed from the interaction effect in Table 5-1, for every decrease in the press freedom score, the effect of open data on corruption diminishes by 0.0004. To unpack the nature of moderating effect, I also conducted a simple slopes analysis, which helps to determine whether the relationship between open data availability and corruption is significant for specific values of the press freedom (Hayes, 2018). I set the values of the press freedom score at 10 (free press, e.g. Norway, Sweden), 30 (moderate government or corporate influence, e.g. Italy), 50 (high levels of violence against journalists or state control, e.g. Albania and Bosnia and Herzegovina) and 90 (not free, e.g. Saudi Arabia). For the press freedom score of value 10, the coefficient is 0.02 ( $p < 0.001$ ), for the score of value 30, it decreased to 0.01 ( $p < 0.001$ ). However, for the press freedom score of value 50 as well as 90, the relationship between open data and corruption is no longer significant.

As the Johnson-Neyman plot illustrates more precisely in Figure 5-1, open data availability has no effect on corruption when the press freedom score is higher than 43.08. This finding is consistent with the proposed theory as well as Lindstedt's and Naurin's (2010) results. It means that greater open data availability positively affects corruption in countries which enjoy a free press, i.e. news coverage is vigorous, the safety of journalists is assured, the government interference in media is marginal, and the press is not imperilled by the interests of powerful economic groups. At a level of diminishing press freedom score to 43.08, open data is no longer significantly associated with corruption levels.

This finding is also interesting in light of Schnell's (2014, 2017) work. Using the case study of Romania, she demonstrated that even when the government adopts transparency and anti-corruption policies, in particular FOI laws and asset disclosures for high-level officials, as cheap talk, they might still "generate costs for decision-makers" and they did in Romania (2017, p. 420). She argued that while the pressure from the EU to adopt and retain adopted transparency and anti-corruption policies was an important factor, domestic actors were at least equally vital, and every attempt to weaken or retract these policies was followed by public outrage. The findings above aptly complement Schnell's research and propose an argument that this public outrage owes to a great extent to the free press. In countries with heavily

state-controlled media or otherwise violated journalistic rights, journalists might not be able to fully benefit from adopted transparency and anti-corruption policies and their reporting might be heavily censored, leaving the public uninformed or mislead about corruption cases.

Figure 5-1: Conditional effect of open data on corruption as function of press freedom



As can be seen in model 3 in Table 5-1, the presence of judiciary does not modify the association between open government data and corruption, and thus the findings cannot confirm h1c hypothesis. Also, in model 3, the individual effect of open data on corruption vanish.

In addition to media freedom, the conditional effects of open data on corruption as a function of internet freedom were explored. The moderating effects of internet freedom are stronger than that of media freedom. As can be observed from Table 1-1 in the Appendix, the moderating effect of internet freedom (model 2) modifies the size and strength of the association between open government data and corruption. For every score increase in open data availability, CoC scores increases by 0.04 scores (robust standard error = 0.01,  $p < 0.001$ ), i.e. levels of corruption go down. Substantively speaking, this is a 0.8% increase on the scale of CoC (from -2.5 to 2.5). With deteriorating internet freedom by one point, the effect of open data on corruption diminishes by 0.001. Simple slopes analysis also shows that the relationship between open data availability and corruption is significant only for values of internet freedom either lower than 39.64 or higher than 85.86. Thus, in countries with decent internet freedom (the range includes countries such as Mexico, Tunisia or Ukraine), the increase in open data availability is associated with lower levels of corruption.

I also included membership in OGP as an instrumental variable and used two-stage least squares regression to estimate the effect of open government data on corruption and address potential endogeneity problems. By using the instruments, I get a variance of open government data measure that is not correlated with the confounding variable. The results are presented in Table 1-2 in the appendix. The F-test statistic on instruments is 12.993 ( $p < 0.001$ ), interacted with press freedom is 6.284 ( $p < 0.001$ ), and interacted with judiciary is

15.761 ( $p < 0.001$ ). Therefore, the null hypothesis that instruments are weak can be rejected. However, the Wu-Hausman test is statistically significant only in model 1 and 3 ( $p < 0.05$ ), suggesting that IV estimates are more consistent than OLS estimates in these models. However, in model 2, the Wu-Hausman test was not significant, and thus the null hypothesis that both OLS and IV estimates are consistent cannot be rejected for this model.

I also explored the associations between different types of government datasets and corruption levels. The score of data availability is an aggregate score and consists of scores for different government datasets. Therefore, it was possible to explore how the datasets that are particularly crucial for monitoring how the government manages public resources, such as data on government spending, contracts or land ownership data are related to corruption levels. The results in Table 1-4 in the appendix show that some datasets essential for monitoring public resources, in particular, land ownership data has indeed a statistically significant effect on corruption. I refrained from proposing directional hypotheses, as the literature on the link between data on the government's management of public money and corruption has been inconclusive in this regard<sup>11</sup>. I find that an increase in land ownership data is associated with lower levels of corruption. However, this effect is small, with a coefficient of 0.005 (robust standard error = 0.002) statistically significant at  $p < 0.05$ . Similarly, the evidence about the direction of the relationship between the availability of data on land ownership and corruption is inconclusive (Benjamin, Bhuvaneswari, Rajan, & Manjunatha, 2007). Lastly, the effect of free and fair elections, journalists' murders, civil society, independent and accountable judiciary and GDP on corruption levels remained statistically significant as in the majority of models.

## 5.2 FOI laws and the moderating effects of press and internet freedom

Table 5-2 presents the model without interaction effect and models where interaction effect between the number of FOI years and press freedom was added (model 4a, 4b and 5). In the first model, the tradition of FOI legislation, measured as the number of years since FOIA was adopted in a country, was not significantly associated with corruption levels. Therefore, the results are inconsistent with h1b hypothesis that countries with a long tradition of FOI laws are more likely to exhibit lower levels of corruption than those where FOI legislation was introduced just recently. This is an interesting non-finding, contrary to the intuitive assumptions. These results might be indicative of the problems with the measurement discussed above. While FOI law tradition might be associated with the quality of the law implementation, a more accurate measure of FOI implementation might be more appropriate.

The diagnostic plots of residuals for model 4a\*, in particular, the Residuals vs. Leverage plot, reveals that Sweden, with the longest tradition of FOI legislation in the world, is an influential observation because it falls beyond Cook's distance of 1 and thus, might affect the slope of the regression line and have an effect on the parameter estimates (see Figure 1-8 in Appendix). Therefore, in addition to model 4a where Sweden was included, a model without

---

\* Model 4a: Corruption =  $\alpha + \beta_{\text{open data}} + \beta_{\text{FOI years}} + \beta_{\text{press freedom}} + \beta_{\text{murdered journalists}} + \beta_{\text{civil society}} + \beta_{\text{judiciary}} + \beta_{\text{free election}} + \beta_{\text{GDP}} + \beta_{\text{FOI years}} * \beta_{\text{press freedom}} + \epsilon$

Sweden was estimated (4b). As can be observed from Table 5-2, the inclusion of Sweden in the dataset influence the parameter estimates. The results of model 4a are consistent with some of my hypotheses, but the moderating effect of press freedom is significantly stronger when Sweden is excluded from the analysis (4b).

Table 5-2: The associations between FOI laws and corruption (OLS)

Control of corruption				
	Model 1	Model 4a	Model 4b	Model 5
Open data	0.007* (0.003)	0.008** (0.003)	0.007* (0.003)	0.007* (0.003)
Years of FOIA	0.002 (0.002)	0.005* (0.002)	0.017** (0.006)	-0.007 (0.012)
Press freedom	-0.004 (0.004)	-0.003 (0.004)	0.00004 (0.004)	-0.004 (0.004)
Murdered journalists	-0.269** (0.113)	-0.253* (0.112)	-0.245* (0.110)	-0.267* (0.114)
Civil society	-0.073** (0.025)	-0.070** (0.025)	-0.067** (0.024)	-0.073** (0.025)
Judiciary	0.216*** (0.032)	0.208*** (0.032)	0.214*** (0.032)	0.203*** (0.036)
Free and fair elections	0.008** (0.002)	0.008** (0.002)	0.008*** (0.002)	0.008** (0.002)
GDP	0.020*** (0.003)	0.019*** (0.003)	0.017*** (0.004)	0.020*** (0.003)
Years of FOIA*press freedom		-0.0003* (0.0001)	-0.0005** (0.0002)	
Years of FOIA*judiciary				0.001 (0.002)
Constant	-1.132** (0.400)	-1.191** (0.395)	-1.379*** (0.399)	-1.083** (0.407)
N	110	110	109	110
R <sup>2</sup>	0.879	0.884	0.885	0.880
Adjusted R <sup>2</sup>	0.870	0.874	0.874	0.869
Residual Std. Error	0.364 (df = 101)	0.358 (df = 100)	0.353 (df = 99)	0.365 (df = 100)
F Statistic	92.015*** (df = 8; 101)	85.058*** (df = 9; 100)	84.244*** (df = 9; 99)	81.480*** (df = 9; 100)

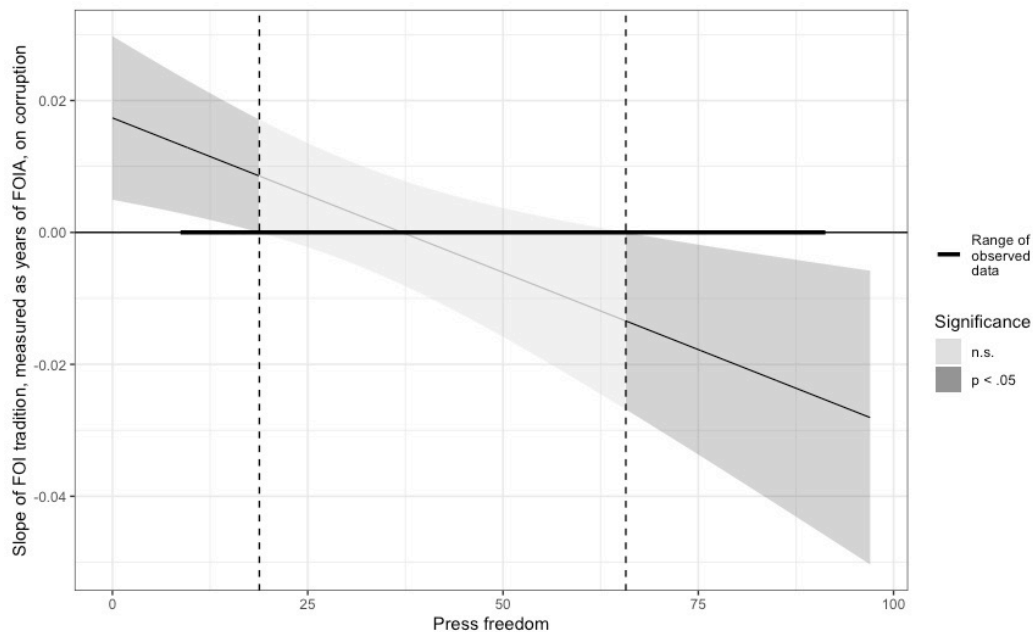
\* p < .05; \*\* p < .01; \*\*\* p < .001

The association between FOI tradition and corruption is significant only when interactions between FOI years and press freedom are added. In model 4a, for every additional year of FOI legislation in place, CoC score increases by 0.005 scores (robust standard error = 0.002, p < 0.05), i.e. levels of corruption go down. In model 4b, these associations are even stronger. For every additional year of FOI law in place, CoC score increases by 0.017 scores (robust standard error = 0.006, p < 0.01), i.e. levels of corruption lower. As can be observed

from the interaction effect in Table 5-2, for every decrease in the press freedom score, the effect of FOI tradition on corruption diminishes by 0.0005.

A simple slopes analysis was again conducted to better understand the complexion of the moderating effect of press freedom on the relationship between FOI tradition and corruption. I followed the same procedures as in the case of exploring the moderating effect of press freedom on open data slopes. I set the values of press freedom score at 10, 30, 50 and 90 scores. The results were different. The FOI tradition is significantly associated with corruption levels only in countries with excellent press freedom score, such as Canada, Norway or Switzerland or opposite, in countries with an extremely poor score, such as Iraq, Malaysia or Qatar. Else, the relationship was not significant. For example, for the press freedom score of value 10, the coefficient is 0.01 ( $p < 0.001$ ), but for the score of value 30, it is insignificant. The relationship between FOI tradition and corruption is also insignificant at press freedom score of value 50. The Johnson-Neyman plot shows that the effect of press freedom on the FOI tradition-corruption association is not significant when press freedom score is outside of the interval 18.81 – 65.72. Once the press freedom score decreases below 65.72, the association becomes significant again.

Figure 5-2: Conditional effect of FOI tradition on corruption as function of press freedom



FOI tradition is significantly associated with decreasing corruption (higher CoC score) in countries with press freedom score lower than 18.81, i.e. countries with unrestricted conditions for media operation. However, in contrast with open data, the tradition of FOI legislation becomes significantly and negatively associated with corruption levels (CoC score) once press freedom score is higher than 65.72 (very high level of censorship). The interaction between FOI laws tradition and internet freedom was not significant. The interaction effect between the number of years with FOI laws in place and independence and accountability of judiciary was also not significant (model 5).

Overall, the findings mostly support previous research on transparency policies and their link to corruption, i.e. the effects being conditional upon media freedom (Besley & Burgess, 2002; Besley et al., 2002; Besley & Prat, 2006; Brunetti & Weder, 2003; Chowdhury, 2004; Djankov et al., 2003; Ferraz & Finan, 2008; Norris, 2008; Reinikka & Svensson, 2005; Solis & Antenangeli, 2017; Svensson, 2005; Yazaki, 2017). The results add to the previously accumulated knowledge in the field and demonstrate that access to free and pluralistic media and internet might be crucial for any transparency measure to be effective also as an anti-corruption measure. It might not be enough for the information and data to be out in the public domain; the information needs to be acted upon. Investigative journalists and civic activists who use FOI laws or government datasets to obtain information, act as intermediaries, who are able to extract crucial pieces from the abundance of the information and interpret them to the public. If their working conditions worsen, become dangerous or compromised by different types of censorships, it might allow corruption to flourish, as it will go unnoticed.

## 6 Conclusions and discussion

Political leaders have always pledged transparency. In the past, many built their election campaigns on the promise to pass FOI laws. Recently, they made a case for open government data as a new digitally-enabled transparency policy that will help to eradicate mismanagement of public funds. Since then, governments worldwide have launched national open data portals and published thousands of datasets. This study is one of the first to investigate the relationship between open data, relative to FOI laws, and corruption levels. The results suggest that both FOI laws and open data are significantly associated with corruption levels only in the interaction with press freedom. The relationship between open data and corruption levels is also conditional on the levels of internet freedom. In particular, the effect of open data on corruption decreases with diminishing media and internet freedom. These results are interesting with respect to the theory as well as policies. First, from the theoretical perspective, they validate Lindstedt's and Naurin's framework and demonstrate its versatility for estimating the effect of different measures of transparency on corruption. Second, the findings suggest that enthusiasm for open data as a powerful anti-corruption means, shared by political leaders and anti-corruption activists should be approached with caution and the relationship between open data and corruption should be further investigated. The present results suggest that media and internet freedom modify the relationship between transparency policies and corruption levels. While in countries that enjoy high levels of media and internet freedom, the increase in open data availability or longer tradition of FOI laws is significantly associated with lower corruption levels, these associations become insignificant once media and internet freedom decrease to a specific threshold. This suggests that transparency policies might be less effective for addressing corruption in conditions where journalists or discontented citizens are censored. For instance, Mexico has adopted several transparency policies and is one of the leaders in open data. However, at the same time, its rate of killings of investigative journalists (Committee to Protect Journalists 2018) makes the profession extremely dangerous and might

discourage journalists from using government information and data indicating that powerful politicians or entrepreneurs are involved in corruption. In countries with restricted conditions for media operation, open government data and other transparency policies might serve as smokescreen boosting the international image of governments as being transparent while shifting attention from substantial domestic policy issues. The analysis has also shown that not all types of government datasets might matter for the anti-corruption fight equally. I find an association between land ownership data and corruption levels only.

There are several limitations to this study. It offers insights into the relationship between open data, relative to FOI laws, and corruption in a certain point of time. Due to the unavailability of longitudinal data on open data availability, quantitatively exploring a causal link is difficult. That said, qualitative case studies using, for example, process tracing techniques to add inferential leverage, could further build on the present findings and explore the impact of specific transparency policies and the role of media and both offline and online civic activism for their success.

This study also takes at face value that government-held information and data is accurate. However, the information made available by governments might be defective in manifold ways, intentionally or unintentionally. While I provided several examples when the accuracy of government information was questionable, it is not within the remits of this study to address this limitation. That said, further research into the credibility of government information more generally, and the debate about the right of access to accurate government information and data is needed. Another caveat of the analysis is that the used measures are imperfect. While measures are only proxies of researched phenomena, some reflect them better than others. For instance, as already raised, the FOI laws tradition can be conflated with democratic tradition to some extent as non-democratic countries started to pass FOI laws only recently. This measure does not tell us much about the implementation of the law. While an excellent piece of legislation might be implemented poorly, imperfect legislation might well deliver good levels of openness. Nonetheless, other than to acknowledge this limitation, a little can be done to address it, as there are no cross-country measures of the quality of FOIA implementation available because they are notoriously difficult to develop.

## **Acknowledgements**

A previous version of this article was presented at the 2018 EPSA Conference and the 2019 ECPR General Conference. I thank the participants at these events as well as Heinz Brandenburg, Zac Greene, and Ben Worthy for their invaluable feedback. I am also grateful for the excellent comments and suggestions provided by three anonymous reviewers and editor. All remaining errors and omissions are my own.

## **Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.



## References

- Adu, K. K. (2018). The paradox of the right to information law in Africa. *Government Information Quarterly*, 35(4), 669–674. <https://doi.org/10.1016/j.giq.2018.10.003>
- Ali, S. R., & Fahmy, S. (2013). Gatekeeping and citizen journalism: The use of social media during the recent uprisings in Iran, Egypt, and Libya. *Media, War and Conflict*, 6(1), 55–69. <https://doi.org/10.1177/1750635212469906>
- Andersson, S., & Heywood, P. M. (2009). The politics of perception: Use and abuse of transparency international's approach to measuring corruption. *Political Studies*, 57(4), 746 – 767. <https://doi.org/https://doi.org/10.1177/0010414013489081>
- Angrist, J. D., & Pischke, J. S. (2015). Mastering 'metrics: The path from cause to effect. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton, Oxford: Princeton University Press. <https://doi.org/10.1093/erae/jbv011>
- Arnold, H. J., & Feldman, D. C. (1981). Social Desirability Response Bias in Self-Report Choice Situations. *Academy of Management Journal*, 24(2), 377–385. <https://doi.org/10.2307/255848>
- Bac, M. (2001). Corruption, connections and transparency: does a better screen imply a better scene? *Public Choice*, 107, 87 – 96. <https://doi.org/10.1023/A:1010349907813>
- Banisar, D. (2004). The Freedominfo.org Global Survey: Freedom of Information and Access to Government Record Laws around the World.
- Benjamin, S., Bhuvaneswari, Rajan, P., & Manjunatha. (2007). Bhoomi: 'E-Governance', or, an Anti-Politics Machine Necessary to Globalize Bangalore?
- Bentham, J. (1999). Political Tactics. In M. James, C. Blamires, & C. Pease-Watkin (Eds.) (1st ed., p. 267). Oxford: Oxford University Press.
- Bentham, J. (2001). Writings on the poor laws. (M. Quinn, Ed.) (1st ed.). Oxford: Clarendon Press.
- Berliner, D. (2012). Institutionalizing Transparency: The Global Spread of Freedom of Information in Law and Practice. University of Washington. Retrieved from [https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/21770/Berliner\\_washington\\_0250E\\_10985.pdf?sequence=1&isAllowed=y](https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/21770/Berliner_washington_0250E_10985.pdf?sequence=1&isAllowed=y)
- Berliner, D. (2014). The political origins of transparency. *Journal of Politics*, 76(2), 479–491. <https://doi.org/10.1017/S0022381613001412>
- Berliner, D. (2016). Transnational advocacy and domestic law: International NGOs and the design of freedom of information laws. *Review of International Organizations*, 11(1), 121–144. <https://doi.org/10.1007/s11558-015-9228-6>
- Berliner, D., & Erlich, A. (2015). Competing for transparency: Political competition and institutional reform in Mexican states. *American Political Science Review*, 109(1), 110–128. <https://doi.org/10.1017/S0003055414000616>
- Besley, T., & Burgess, R. (2002). The political economy of government responsiveness: theory and evidence from India. *The Quarterly Journal of Economics*, 117(4), 1415–1451. <https://doi.org/10.1162/003355302320935061>
- Besley, T., Burgess, R., & Prat, A. (2002). Mass media and political accountability. In R.

- Islam (Ed.), *The right to tell: the role of mass media in economic development* (1st ed., pp. 45–60). Washington, DC: World Bank Institute. Retrieved from <http://documents.worldbank.org/curated/en/957661468780322581/pdf/multi0page.pdf>
- Besley, T., & Prat, A. (2006). Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96(3), 720–736. <https://doi.org/10.1257/aer.96.3.720>
- Birchall, C. (2015). “Data.gov-in-a-box”: Delimiting transparency. *European Journal of Social Theory*, 18(2), 185–202. <https://doi.org/10.1177/1368431014555259>
- Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, 87(7–8), 1801–1824. [https://doi.org/10.1016/S0047-2727\(01\)00186-4](https://doi.org/10.1016/S0047-2727(01)00186-4)
- Chowdhury, S. K. (2004). The effect of democracy and press freedom on corruption: An empirical test. *Economics Letters*, 85(1), 93–101. <https://doi.org/10.1016/j.econlet.2004.03.024>
- Chung, J., & Monroe, G. S. (2003). Exploring Social Desirability Bias. *Journal of Business Ethics*, 4(4), 291–302. <https://doi.org/10.1023/A:1023648703356>
- Coase, R. H. (1990). *The Firm, the Market, and the Law*. Chicago: University of Chicago Press.
- Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, 19(1), 15–18. <https://doi.org/10.1080/00401706.1977.10489493>
- Costa, S. (2013). Do freedom of information laws decrease corruption? *Journal of Law, Economics, and Organization*, 29(6), 1317–1343. <https://doi.org/10.1093/jleo/ews016>
- Cucciniello, M., Porumbescu, G. A., & Grimmeliikhuijsen, S. (2017). 25 years of transparency research: Evidence and future directions. *Public Administration Review*, 77(1), 32–44. <https://doi.org/10.1111/puar.12685>
- Dai, J., & Li, Q. (2016). Designing audit apps for armchair auditors to analyze government procurement contracts. *Journal of Emerging Technologies in Accounting*, 13(2), 71–88. <https://doi.org/10.2308/jeta-51598>
- Dilulio, J. D. (1994). Principled agents: the cultural bases of behavior in a federal government bureaucracy. *Journal of Public Administration Research and Theory*, 4(3), 277–318. <https://doi.org/10.1093/oxfordjournals.jpart.a037210>
- Djankov, S., McLiesh, C., Nenova, T., & Shleifer, A. (2003). Who owns the media? *The Journal of Law and Economics*, 46(2), 341–382. <https://doi.org/10.1086/377116>
- Ecker, A., Glinitzer, K., & Meyer, T. M. (2016). Corruption performance voting and the electoral context. *European Political Science Review*, 8(3), 333–354. <https://doi.org/10.1017/S1755773915000053>
- ECtHR. Case no. 31475/05, *Kenedi v. Hungary* (2009). Retrieved from <http://hudoc.echr.coe.int/eng?i=001-92663>
- ECtHR. Case No. 48135/06, *Youth Initiative for Human Rights v. Serbia*. (2013). Retrieved from <http://hudoc.echr.coe.int/eng?i=001-120955>
- Esteban, O.-O. (2019). The rise of social media. Retrieved March 15, 2020, from <https://ourworldindata.org/rise-of-social-media>
- Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians: the effects of Brazil’s publicly

- released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2), 703–745.  
<https://doi.org/10.1162/qjec.2008.123.2.703>
- Fox, J. (2007). The uncertain relationship between transparency and accountability. *Development in Practice*, 17(4–5), 663–671. <https://doi.org/10.1080/09614520701469955>
- Freedom House. (2016). Freedom on the Net 2016 - Silencing the Messenger: Communication Apps Under Pressure. Washington, DC. Retrieved from <https://freedomhouse.org/report/freedom-net/2016/silencing-messenger-communication-apps-under-pressure>
- Freedom House. (2017). Freedom of the Press 2017 Methodology. Retrieved May 14, 2018, from <https://freedomhouse.org/report/freedom-press-2017-methodology>
- Fung, A. (2013). Infotopia: unleashing the democratic power of transparency. *Politics & Society*, 41(2), 183–212. <https://doi.org/10.1177/0032329213483107>
- Grønbech-Jensen, C. (1998). The scandinavian tradition of open government and the European union: Problems of compatibility? *Journal of European Public Policy*, 5(1), 185–199. <https://doi.org/10.1080/13501768800000091>
- Hallin, D. C., & Mancini, P. (2012). Comparing media systems beyond the western world. *Comparing Media Systems beyond the Western World*. New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9781139005098>
- Hayes, A. F. (2018). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach (2nd ed.). New York, NY: Guilford Press.
- Heald, D. (2006). Varieties of transparency. In C. Hood & D. Heald (Eds.), *Transparency The Key to Better Governance?* (pp. 25–43). Oxford: Oxford University Press. <https://doi.org/10.5871/bacad/9780197263839.001.0001>
- Hermida, A., & Thurman, N. (2008). A clash of cultures. *Journalism Practice*, 2(3), 343–356. <https://doi.org/10.1080/17512780802054538>
- Heywood, P. M., & Rose, J. (2013). Close but no cigar: the measurement of corruption. *Journal of Public Policy*, 34(3), 507–529. <https://doi.org/10.1017/S0143814X14000099>
- Howard, P. N., & Hussain, M. M. (2013a). Authoritarian Responses and Consequences. In *Democracy's Fourth Wave? Digital Media and the Arab Spring*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199936953.003.0001>
- Howard, P. N., & Hussain, M. M. (2013b). Digital Media and the Arab Spring. In *Democracy's Fourth Wave? Digital Media and the Arab Spring*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199936953.003.0001>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption Barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268. <https://doi.org/10.1080/10580530.2012.716740>
- Kitchin, R. (2014). The data revolution: big data, open data, data infrastructures & their consequences. London: SAGE Publications Ltd. <https://doi.org/10.4135/9781473909472>
- Kolstad, I., & Wiig, A. (2009). Is Transparency the Key to Reducing Corruption in Resource-Rich Countries? *World Development*, 37(3), 521 – 532. <https://doi.org/10.1016/j.worlddev.2008.07.002>
- Laperrière, A. (2019, June). Open data and food security. *The Ecologist*. Retrieved from

- <https://theecologist.org/2019/jun/07/open-data-and-food-security>
- Larkins, C. M. (1996). Judicial Independence and Democratization: A Theoretical and Conceptual Analysis. *American Journal of Comparative Law*, 44(4), 605–626. <https://doi.org/10.2307/840623>
- Lindstedt, C., & Naurin, D. (2010). Transparency is not enough: Making transparency effective in reducing corruption. *International Political Science Review*, 31(3), 301–322. <https://doi.org/10.1177/0192512110377602>
- Lourenço, R. P. (2015). An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly*, 32(3), 323–332. <https://doi.org/10.1016/j.giq.2015.05.006>
- Margetts, H. (2006). Transparency and digital government. In C. Hood & D. Heald (Eds.), *Transparency The Key to Better Governance?* (pp. 197–207). Oxford: Oxford University Press. <https://doi.org/10.5871/bacad/9780197263839.001.0001>
- Martin, C. (2014). Barriers to the open government data agenda: taking a multi-level perspective. *Policy & Internet*, 6(3), 217–240. Retrieved from <http://doi.wiley.com/10.1002/1944-2866.POI367>
- Mauro, P. (1995). Corruption and Growth. *The Quarterly Journal of Economics*, 110(3), 681–712. <https://doi.org/10.2307/2946696>
- Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. London: John Murray Publishers.
- McClean, T. (2011). *Shackling Leviathan: a comparative historical study of institutions and the adoption of Freedom of Information*. London School of Economics and Political Science.
- McMillan, F. (2018). *The Meaning of Mandates: Party Competition and the Programme-to-Policy Linkage*. University of Strathclyde.
- Michener, G. (2010). *The surrender of secrecy: explaining the emergence of strong access to information laws in Latin America*. University of Texas, Austin. Retrieved from <https://repositories.lib.utexas.edu/handle/2152/ETD-UT-2010-05-1112>
- Mills, L. C. (2017). Catching the ‘big fish’: The (ab)use of corruption-related prosecutions across sub-Saharan Africa. *Development Policy Review*, 35, 160–177. <https://doi.org/10.1111/dpr.12326>
- Moore, M. (2011). The limits of transparency. *Political Quarterly*, 82(4), 506–508. <https://doi.org/10.1111/j.1467-923X.2011.02248.x>
- Munro, R. (2015, April 7). Poor data is preventing armchair auditors holding government to account. *The Guardian*. Retrieved from <https://www.theguardian.com/society/2015/apr/07/data-armchair-auditors-government-departments-performance-indicators>
- Nam, T. (2012). Freedom of information legislation and its impact on press freedom: A cross-national study. *Government Information Quarterly*, 29(4), 521–531. <https://doi.org/10.1016/j.giq.2012.03.003>
- No author. (2007). *The Annotated 8 Principles of Open Government Data*. Retrieved January 29, 2018, from <https://opengovdata.org/>

- Norris, P. (2008). The fourth estate. In 1 (Ed.), *Driving democracy: do power-sharing institutions work?* (p. 320). New York/Cambridge: Cambridge University Press.
- Noveck, B. S. (2009). *Wiki government: how technology can make government better, democracy stronger, and citizens more powerful*. Washington, DC: Brookings Institution Press. <https://doi.org/10.5860/choice.47-2862>
- Noveck, B. S. (2017). Rights-Based and Tech-Driven: Open Data, Freedom of Information, and the Future of Government Transparency. *Yale Human Rights and Development Law Journal*, 19(1), 1–45.
- Noveck, B. S. (2018). Open Data: The Future of Transparency in the Age of Big Data. In D. E. Pozen & M. Schudson (Eds.), *Troubling Transparency: The History and Future of Freedom of Information* (1st ed., pp. 206–225). New York, NY: Columbia University Press. <https://doi.org/10.1017/jli.2019.7>
- O'Neill, O. (2006). Transparency and the ethics of communication. In C. Hood & D. Heald (Eds.), *Transparency The Key to Better Governance?* (pp. 75–90). Oxford: Oxford University Press. <https://doi.org/10.5871/bacad/9780197263839.001.0001>
- OECD. (2016). *Compendium of good practices on the publication and reuse of open data for Anti-corruption across G20 countries: Towards data-driven public sector integrity and civic auditing*. Paris. Retrieved from <https://www.oecd.org/gov/digital-government/g20-oecd-compendium.pdf>
- Parycek, P., Schöllhammer, R., & Schossböck, J. (2016). “Each in their own garden”: Obstacles for the implementation of open government in the public sector of the German-speaking region. In *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance (ICEGOV '15-16)* (pp. 291–300). <https://doi.org/10.1145/2910019.2910105>
- Peisakhin, L., & Pinto, P. (2010). Is transparency an effective anti-corruption strategy? Evidence from a field experiment in India. *Regulation and Governance*, (4), 261 – 280. <https://doi.org/10.1111/j.1748-5991.2010.01081.x>
- Peixoto, T. (2013). The uncertain relationship between open data and accountability: A response to Yu and Robinson’s *The New Ambiguity of “Open Government.”* *UCLA Law Review Discourse*, 60, 200–213.
- Pemstein, D. et al. (2015). *V-Dem Codebook*.
- Pozen, D. E. (2017). *Freedom of Information Beyond the Freedom of Information Act*. *University of Pennsylvania Law Review*, 165, 1097–1158.
- Reinikka, R., & Svensson, J. (2005). Fighting corruption to improve schooling: evidence from a newspaper campaign in Uganda. *Journal of the European Economic Association*, 3(2–3), 259–267. <https://doi.org/10.1162/jeea.2005.3.2-3.259>
- Relly, J. E., & Schwalbe, C. B. (2013). Watchdog journalism: India’s three largest English-language newspapers and the Right to Information Act. *Asian Journal of Communication*, 23(3), 284–301. <https://doi.org/10.1080/01292986.2012.729149>
- Roberts, A. S. (2006). *Blacked out: Government secrecy in the information age*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9781139165518>
- Rose, R., & Mishler, W. (2010). *Experience versus perception of corruption: Russia as a test*

- case. *Global Crime*, 11(2), 145 – 163. <https://doi.org/10.1080/17440571003669175>
- Rosenn, K. S. (1987). Protection of Judicial Independence in Latin America. *The University of Miami Inter-American Law Review*, 1(19), 1–35.
- Samaha, A. (2006). Government Secrecy, Constitutional Law, and Platforms for Judicial Elaboration. *UCLA Law Review*, 53(4), 909–997.
- Santiso, C., & Roseth, B. (2017). Data disrupts corruption. *Stanford Social Innovation Review*, 15(2), 50–55.
- Schnell, S. (2014). International Policy Diffusion and Domestic Policy-Making- the Case of Two Transparency and Anticorruption Policies in Romania. The Faculty of The Columbian College of Arts and Sciences of The George Washington University.
- Schnell, S. (2017). Cheap talk or incredible commitment? (Mis)calculating transparency and anti-corruption. *Governance*.
- Schofield, P. (2006). Publicity, responsibility, and the architecture of government. In Oxford University Press (Ed.), *Utility and Democracy: The political thought of Jeremy Bentham* (pp. 250–271). Oxford. <https://doi.org/10.1093/acprof:oso/9780198208563.001.0001>
- Scrollini, F. M. (2015). Right-to-information arenas: Exploring the right to information in Chile, New Zealand and Uruguay. London School of Economics. Retrieved from <http://etheses.lse.ac.uk/3361/>
- Sen, A. (1999). *Development as Freedom*. Oxford: Oxford University Press.
- Sharafutdinova, G. (2010). What explains corruption perceptions? The dark side of political competition in Russia's regions. *Comparative Politics*, 42(2), 147–166. <https://doi.org/10.5129/001041510X12911363509431>
- Shkabatur, J. (2012). Transparency with(out) accountability: open government in the United States. *Yale Law and Policy Review*, 31(1), 79–140.
- Shkabatur, J., & Peled, A. (2016). Sustaining the open government data movement worldwide: Insights from developing countries. In *Proceedings of the 6th International Conference for E-Democracy and Open Government, CeDEM 2016* (pp. 172–179). <https://doi.org/10.1109/CeDEM.2016.10>
- Smulovitz, C., & Peruzzotti, E. (2000). Societal Accountability in Latin America. *Journal of Democracy*, 11(4), 147–158. <https://doi.org/10.1353/jod.2000.0087>
- Solis, J. A., & Antenangeli, L. (2017). Corruption is bad news for a free press: reassessing the relationship between media freedom and corruption. *Social Science Quarterly*, 98(3), 1112–1137. <https://doi.org/10.1111/ssqu.12438>
- Sovey, A. J., & Green, D. P. (2011). Instrumental variables estimation in political science: a readers' guide. *American Journal of Political Science*, 55(1), 188–200. <https://doi.org/10.1111/j.1540-5907.2010.00477.x>
- Stiglitz, J. E. (1999a). Knowledge as a global public good. In I. Kaul, I. Grunberg, & M. Stern (Eds.), *Global public goods: international cooperation in the 21st century* (1st ed., pp. 308–325). Oxford: Oxford University Press.
- Stiglitz, J. E. (1999b). On Liberty, the right to know, and public discourse: the role of transparency in public life. Retrieved July 25, 2019, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.594.93&rep=rep1&type=pdf>

- Surowiecki, J. (2005). *The Wisdom of Crowds* (1st ed.). New York: Anchor Books.
- Svensson, J. (2005). Eight Questions about Corruption. *Journal of Economic Perspectives*, 19(3), 19–42. <https://doi.org/10.1257/089533005774357860>
- The Conservative Party. (2010). 2010 Conservative Party Manifesto. Retrieved from <https://www.conservatives.com/~media/Files/Manifesto2010>
- The Labour Party. (1997). 1997 Labour Party Manifesto: New Labour because Britain deserves better. Retrieved from <http://www.labour-party.org.uk/manifestos/1997/1997-labour-manifesto.shtml>
- Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication*, 62(2), 363–379. <https://doi.org/10.1111/j.1460-2466.2012.01629.x>
- Vadlamamati, K. C., & Cooray, A. (2017). Transparency pays? Evaluating the effects of the freedom of information laws on perceived government corruption. *Journal of Development Studies*, 53(1), 116–137. <https://doi.org/10.1080/00220388.2016.1178385>
- Voltmer, K. (2013). *The media in transitional democracies*. Cambridge, UK: Polity Press.
- Vrushni, J., & Hodess, R. (2017). *Connecting the dots: Building the case for open data to fight corruption*. Berlin. Retrieved from [https://www.transparency.org/whatwedo/publication/connecting\\_the\\_dots\\_building\\_the\\_case\\_for\\_open\\_data\\_to\\_fight\\_corruption](https://www.transparency.org/whatwedo/publication/connecting_the_dots_building_the_case_for_open_data_to_fight_corruption)
- Waisbord, S. (2002). Antipress violence and the crisis of the state. *Harvard International Journal of Press/Politics*, 7, 90–109. <https://doi.org/10.1177/1081180X02007003006>
- Wheeler, B. (2012, November 9). Government online data ignored by “armchair auditors.” BBC News.
- Winters, M. S., & Weitz-Shapiro, R. (2013). Lacking Information or Condoning Corruption: When Do Voters Support Corrupt Politicians? *Comparative Politics*, 45(4), 418–436. <https://doi.org/10.5129/001041513X13815259182857>
- World Bank. (2018). World Bank Statement on Amendments to Tanzania’s 2015 Statistics Act. Retrieved February 10, 2018, from <https://www.worldbank.org/en/news/statement/2018/10/02/world-bank-statement-on-amendments-to-tanzanias-2015-statistics-act>
- World Wide Web Foundation. (2016). Open Data Barometer - 4th edition - ODB Methodology. Open Data Barometer. Retrieved from <https://opendatabarometer.org/4thedition/methodology/>
- Worthy, B. (2015a). The impact of open data in the UK: Complex, unpredictable and political. *Public Administration*, 93(3), 788–805. <https://doi.org/10.1111/padm.12166>
- Worthy, B. (2015b, June 10). Open data on council spending is largely unread by voters. *The Guardian*. Retrieved from <https://www.theguardian.com/public-leaders-network/2015/jun/10/council-finance-open-data-unread-by-voters>
- Yazaki, Y. (2017). Newspapers and political accountability: evidence from Japan. *Public Choice*, 172(3–4), 311–331. <https://doi.org/10.1007/s11127-017-0444-x>
- Yu, H., & Robinson, D. G. (2012). The new ambiguity of “open government.” *UCLA Law Review*, 178, 180–208.

- Zuiderwijk, A., & Janssen, M. (2015). Participation and data quality in open data use: Open data infrastructures evaluated. In Proceedings of the European Conference on e-Government, ECEG.
- Zuiderwijk, A., Janssen, M., & Dwivedi, Y. K. (2015). Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly*, 32(4), 429–440.  
<https://doi.org/10.1016/j.giq.2015.09.005>



# 1 Appendix

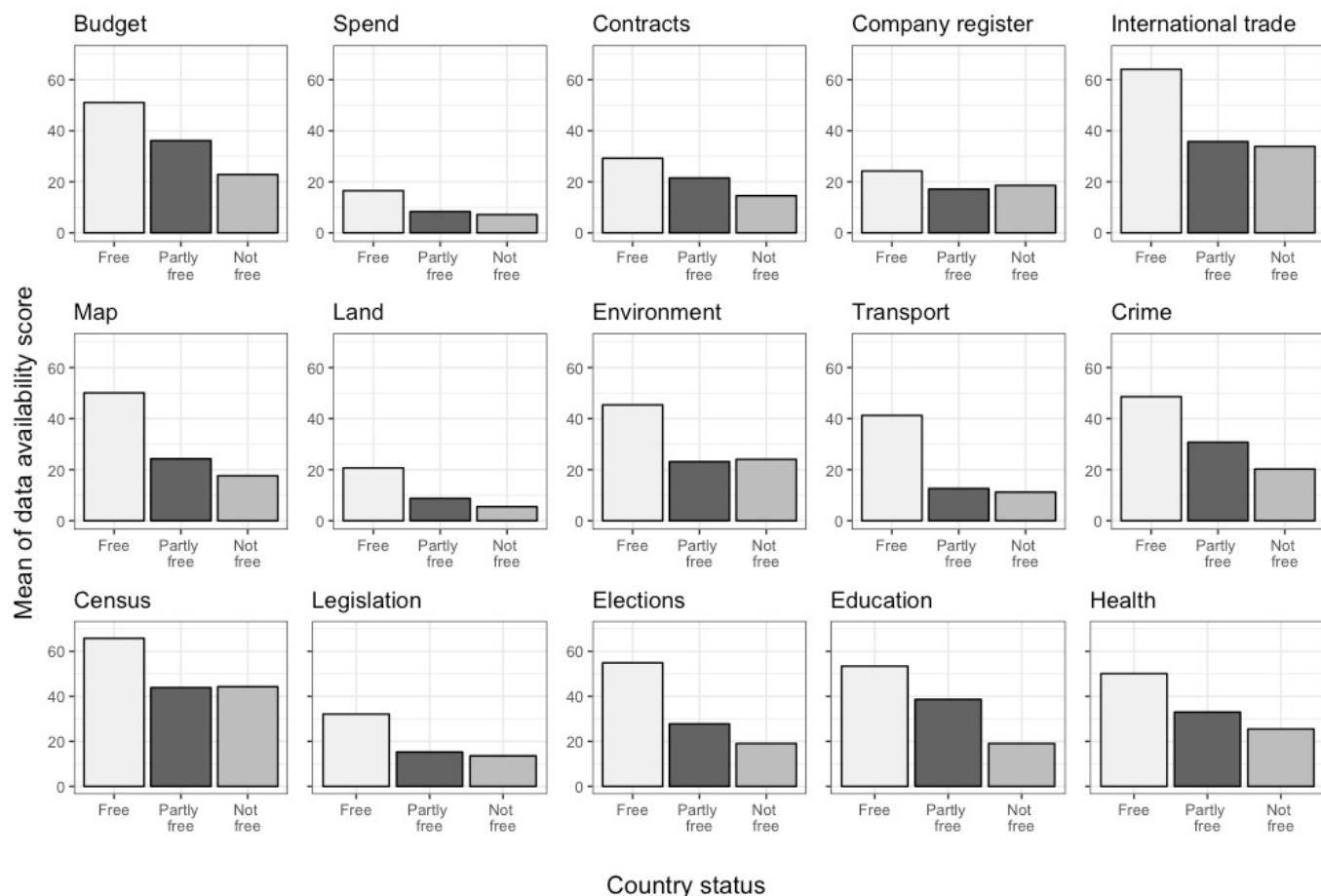
## 1.1 The number and list of countries included in the main model

Dependent variable: CoC (Control of corruption)

Number of observations: 110

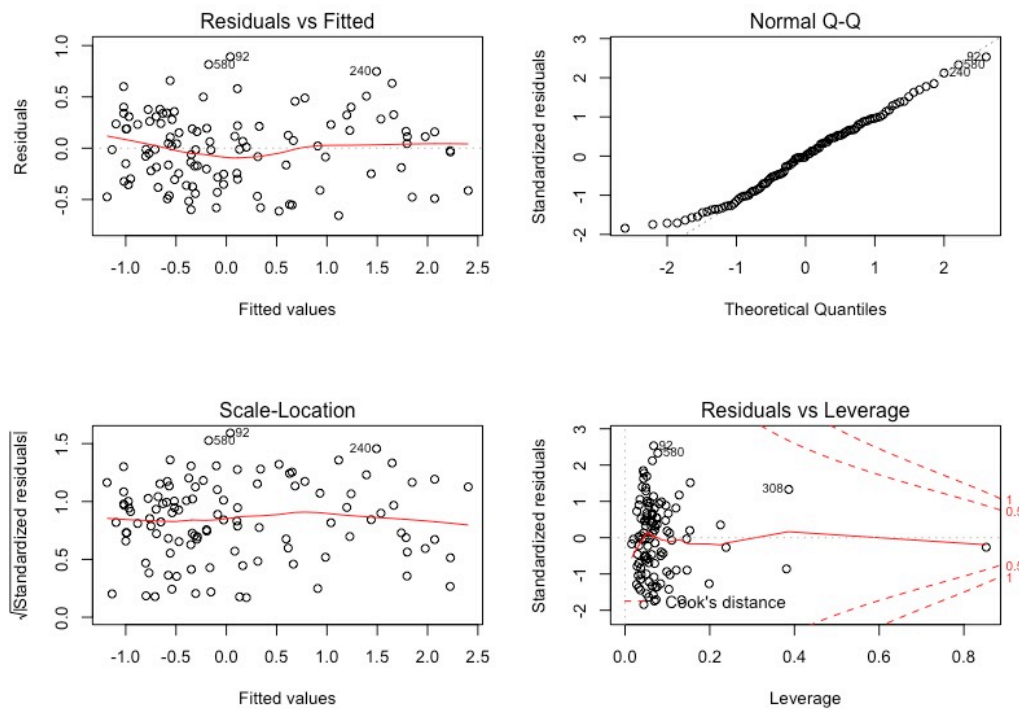
Albania, Argentina, Australia, Austria, Bangladesh, Belarus, Belgium, Benin, Bolivia, Bosnia and Hercegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Cameroon, Canada, Chile, China, Colombia, Congo DR, Costa Rica, Cote d'Ivoire, Croatia, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Greece, Guatemala, Haiti, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kosovo, Kyrgyz Republic, Latvia, Lebanon, Macedonia, Malawi, Malaysia, Mali, Mauritius, Mexico, Moldova, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Russian Federation, Rwanda, Saudi Arabia, Senegal, Serbia, Sierra Leone, Singapore, Slovakia, South Africa, South Korea, Spain, Sweden, Switzerland, Tajikistan, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Uganda, Ukraine, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe

Figure 1-1: Mean of data availability score by Freedom House country status



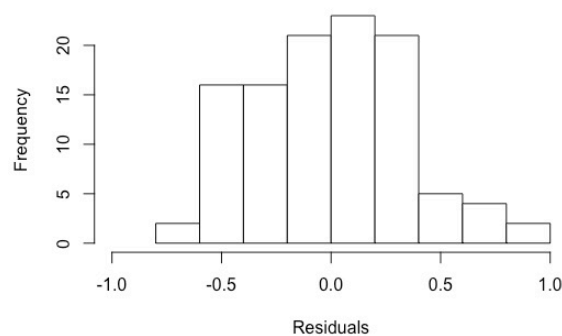
## 1.2 Regression diagnostic plots for all models

Figure 1-2: Diagnostic plots of residuals for model 1 (no interactions)



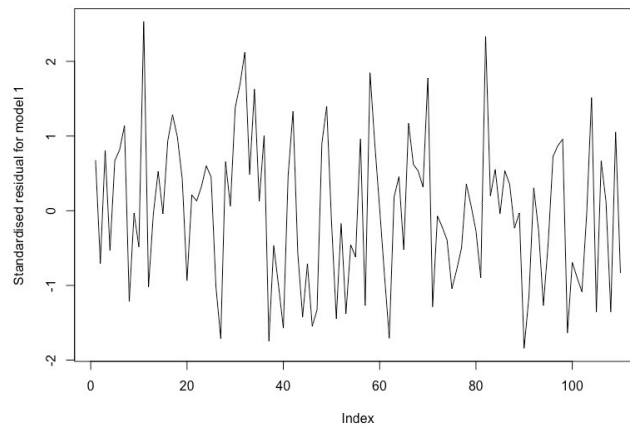
Note: Residuals vs. Fitted plot shows no distinct pattern; the line is relatively straight, horizontal at zero; the plot indicates that there is a linear relationship between the dependent and the independent variables, i.e. that the linear relationship assumption is met. Normal Q-Q plot shows that most residuals fall along the diagonal reference line; the plot indicates that the residuals are normally distributed, i.e. that the normality assumption is met. Scale-Location plot shows no distinct pattern; the line is relatively straight, horizontal, with residuals spread approximately evenly across the range of fitted values. Residuals vs. Leverage plot shows that none of the points falls beyond Cook's distance of 1, which was suggested as one of the cut-off points to identify influential observations (e.g. outliers and/or high-leverage points).

Figure 1-3: Histogram of residuals for model 1 (no interactions)



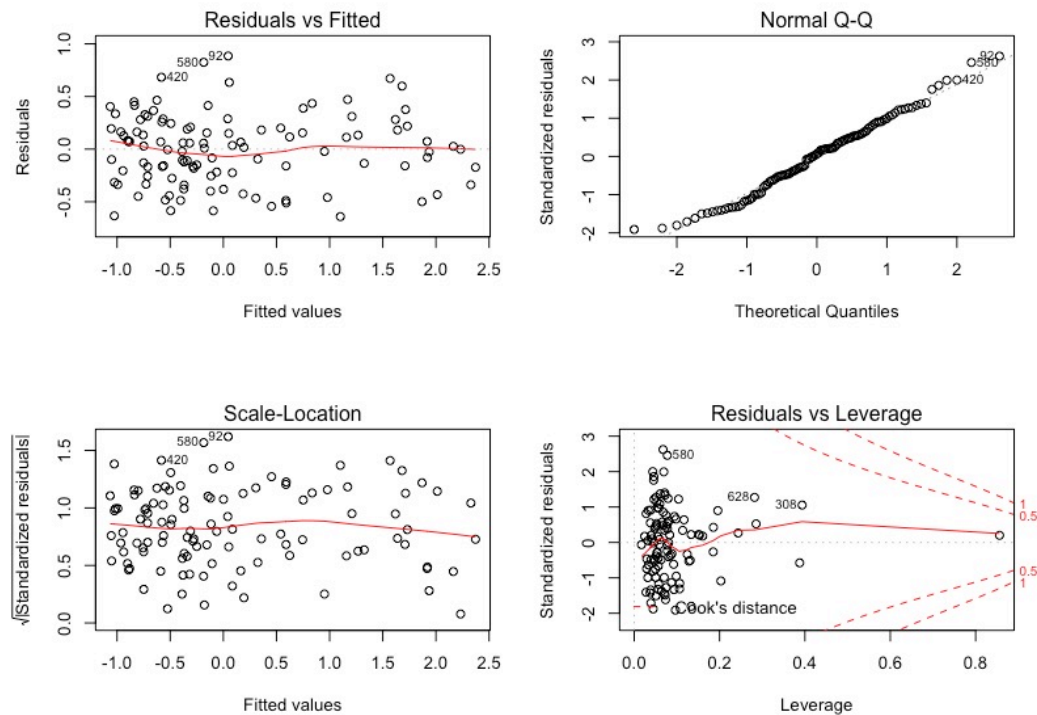
Note: The histogram shows that residuals are normally distributed, which was also confirmed by the results of a formal test. Shapiro-Wilk normality test results ( $p = 0.188$ ) suggest that the null hypothesis that the data (in this case, residuals) come from a normally distributed population cannot be rejected. Breusch-Pagan test results ( $p = 0.485$ ) suggest that the null hypothesis of homoskedasticity cannot be rejected (i.e., there may not be a heteroskedasticity problem).

Figure 1-4: Independence of residuals plot for model 1 (no interactions)



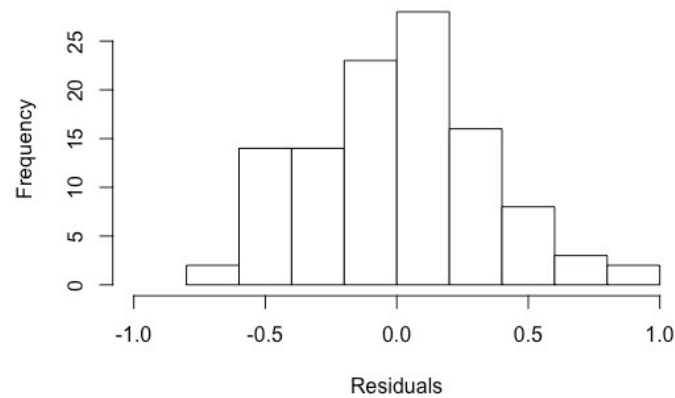
Note: The plot shows no serial pattern in the way the residuals of the model appear. It indicates that the residuals are independent of each other.

Figure 1-5: Diagnostic plots of residuals for model 2 (open data-press freedom interactions)



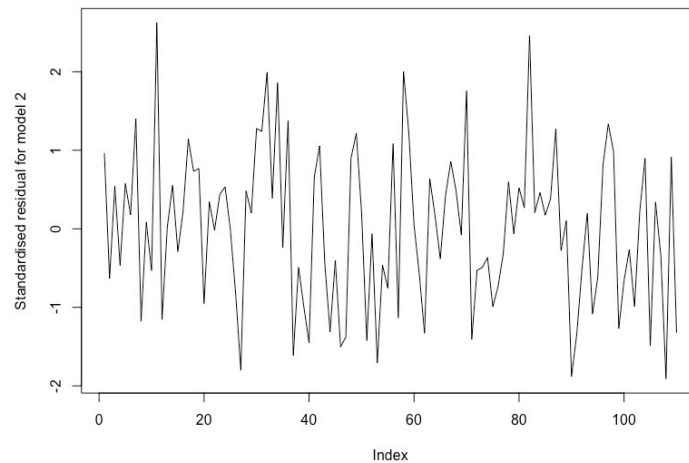
Note: Residuals vs. Fitted plot shows no distinct pattern; the line is relatively straight, horizontal at zero; the plot indicates that there is a linear relationship between the dependent and the independent variables, i.e. that the linear relationship assumption is met. Normal Q-Q plot shows that most residuals fall along the diagonal reference line; the plot indicates that the residuals are normally distributed, i.e. that the normality assumption is met. Scale-Location plot shows no distinct pattern; the line is relatively straight, horizontal, with residuals spread approximately evenly across the range of fitted values. Residuals vs. Leverage plot shows that none of the points falls beyond Cook's distance of 1, which was suggested as one of the cut-off points to identify influential observations (e.g. outliers and/or high-leverage points).

Figure 1-6: Histogram of residuals for model 2 (open data-press freedom interactions)



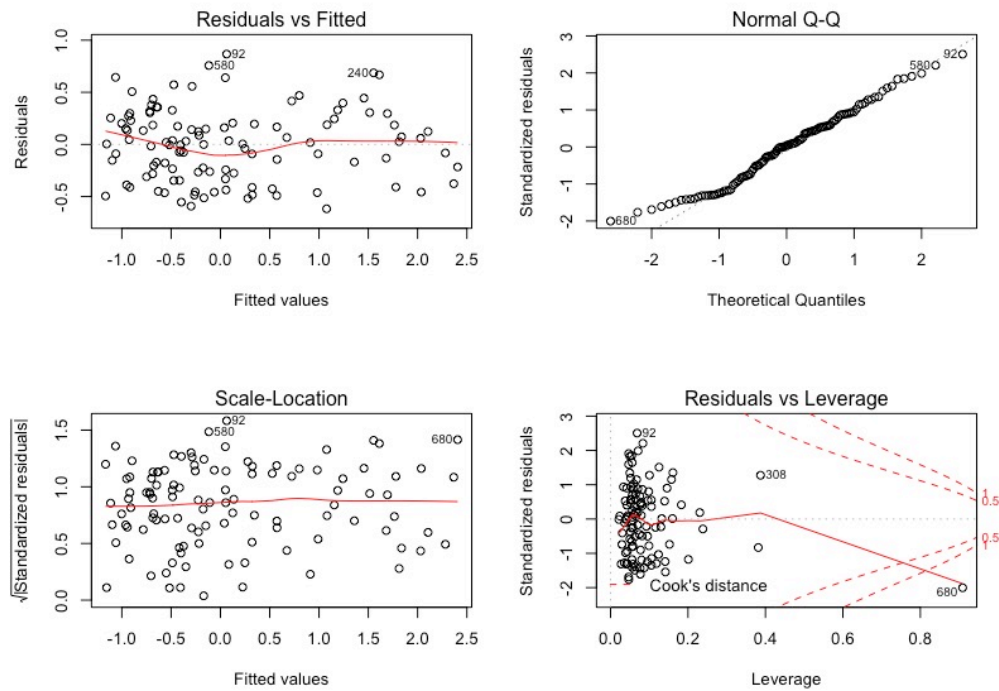
Note: The histogram shows that residuals are normally distributed, which was also confirmed by the results of a formal test. Shapiro-Wilk normality test results ( $p = 0.053$ ) suggest that the null hypothesis that the data (in this case, residuals) come from a normally distributed population cannot be rejected. Breusch-Pagan test results ( $p = 0.761$ ) suggest that the null hypothesis of homoskedasticity cannot be rejected (i.e., there may not be a heteroskedasticity problem).

Figure 1-7: Independence of residuals plot for model 2 (open data-press freedom interactions)

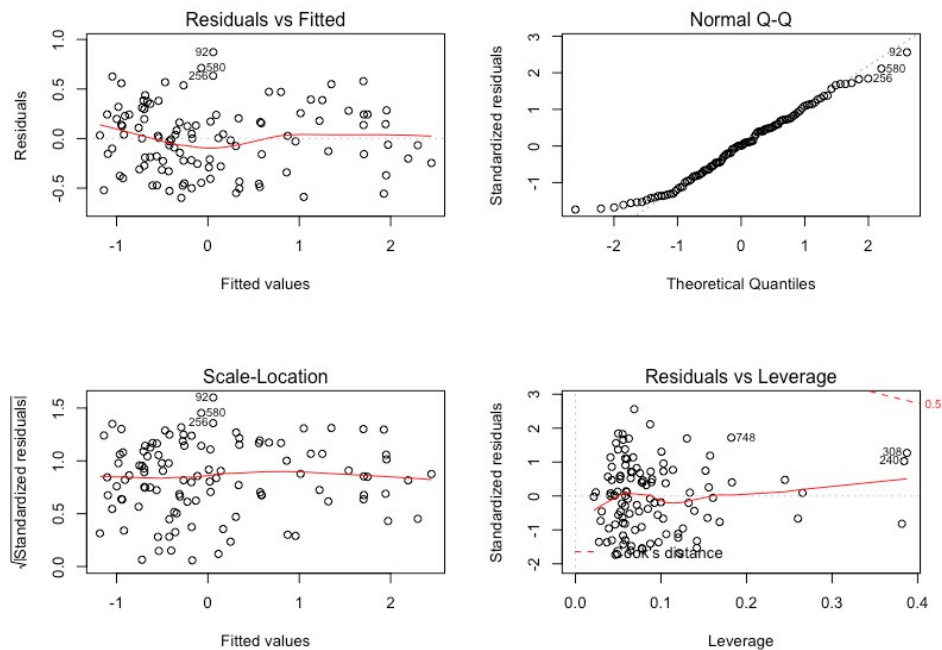


Note: The plot shows no serial pattern in the way the residuals of the model appear. It indicates that the residuals are independent of each other.

Figure 1-8: Diagnostic plots of residuals for model 4a and 4b (FOI years-press freedom interactions)

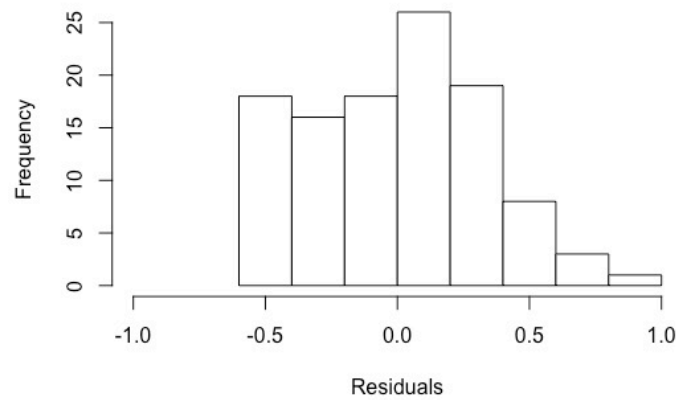


Note: Residuals vs. Fitted plot shows no distinct pattern; the line is relatively straight, horizontal at zero; the plot indicates that there is a linear relationship between the dependent and the independent variables, i.e. that the linear relationship assumption is met. Normal Q-Q plot shows that most residuals fall along the diagonal reference line; the plot indicates that the residuals are normally distributed, i.e. that the normality assumption is met. Scale-Location plot shows no distinct pattern; the line is relatively straight, horizontal, with residuals spread approximately evenly across the range of fitted values. However, as can be observed from the Residuals vs. Leverage plot, observation no. 680 (Sweden) is an influential observation because it falls beyond Cook's distance of 1, which was suggested as one of the cut-off points to identify influential observations. It might affect the slope of the regression line and have an effect on the parameter estimates. Therefore, it was removed from the data for the analysis.



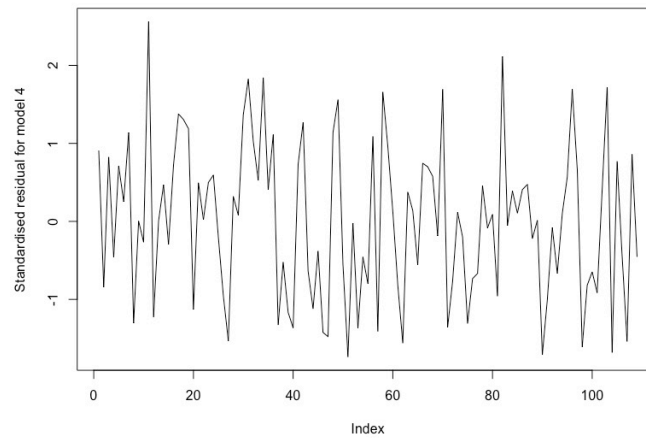
Note: Once the influential observation (no.680 - Sweden) was removed, the Residuals vs. Leverage plot shows that none of the points falls beyond Cook's distance of 1.

Figure 1-9: Histogram of residuals for model 4b (FOI years-press freedom interactions)



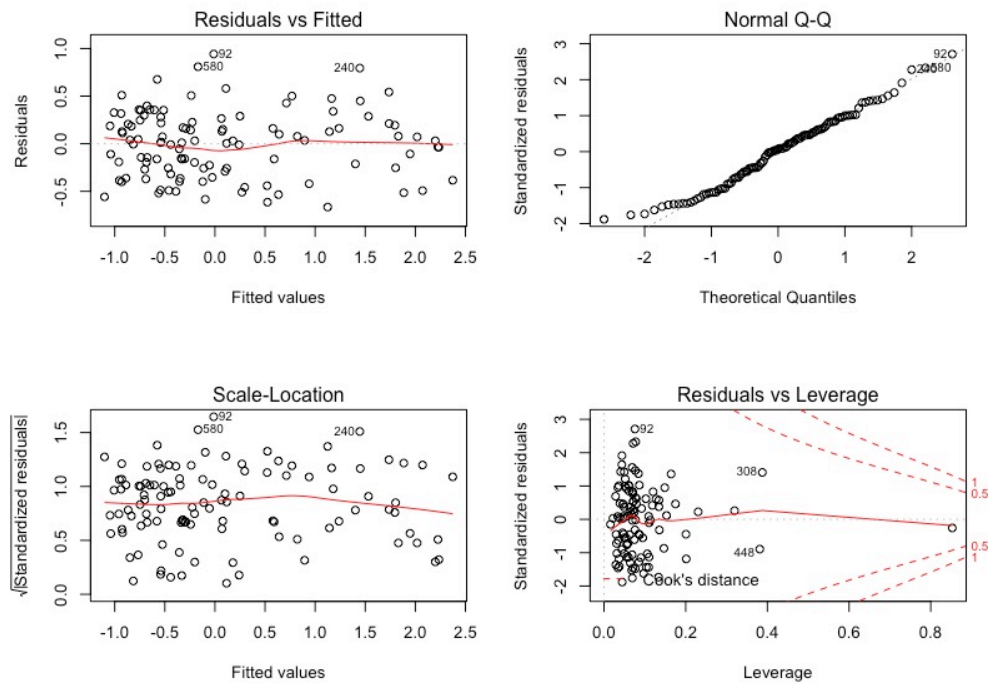
Note: The histogram shows that residuals are normally distributed, which was also confirmed by the results of a formal test. Shapiro-Wilk normality test results ( $p = 0.106$ ) suggest that the null hypothesis that the data (in this case, residuals) come from a normally distributed population cannot be rejected. Breusch-Pagan test results ( $p = 0.972$ ) suggest that the null hypothesis of homoskedasticity cannot be rejected (i.e., there may not be a heteroskedasticity problem).

Figure 1-10: Independence of residuals plot for model 4b (FOI years-press freedom interactions)



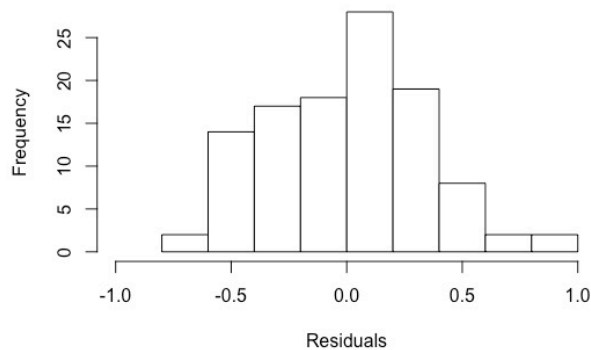
Note: The plot shows no serial pattern in the way the residuals of the model appear. It indicates that the residuals are independent of each other.

Figure 1-11: Diagnostic plots of residuals for model 3 (open data-judiciary interactions)



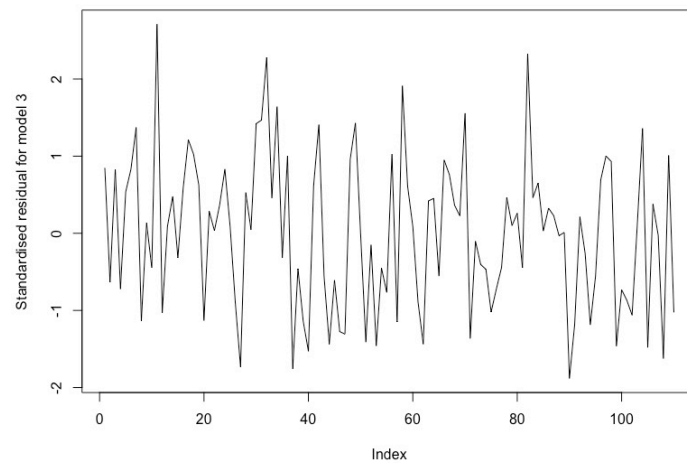
Note: Residuals vs. Fitted plot shows no distinct pattern; the line is relatively straight, horizontal at zero; the plot indicates that there is a linear relationship between the dependent and the independent variables, i.e. that the linear relationship assumption is met. Normal Q-Q plot shows that most residuals fall along the diagonal reference line; the plot indicates that the residuals are normally distributed, i.e. that the normality assumption is met. Scale-Location plot shows no distinct pattern; the line is relatively straight, horizontal, with residuals spread approximately evenly across the range of fitted values. Residuals vs. Leverage plot shows that none of the points falls beyond Cook's distance of 1, which was suggested as one of the cut-off points to identify influential observations (e.g. outliers and/or high-leverage points).

Figure 1-12: Histogram of residuals for model 3 (open data-judiciary interactions)



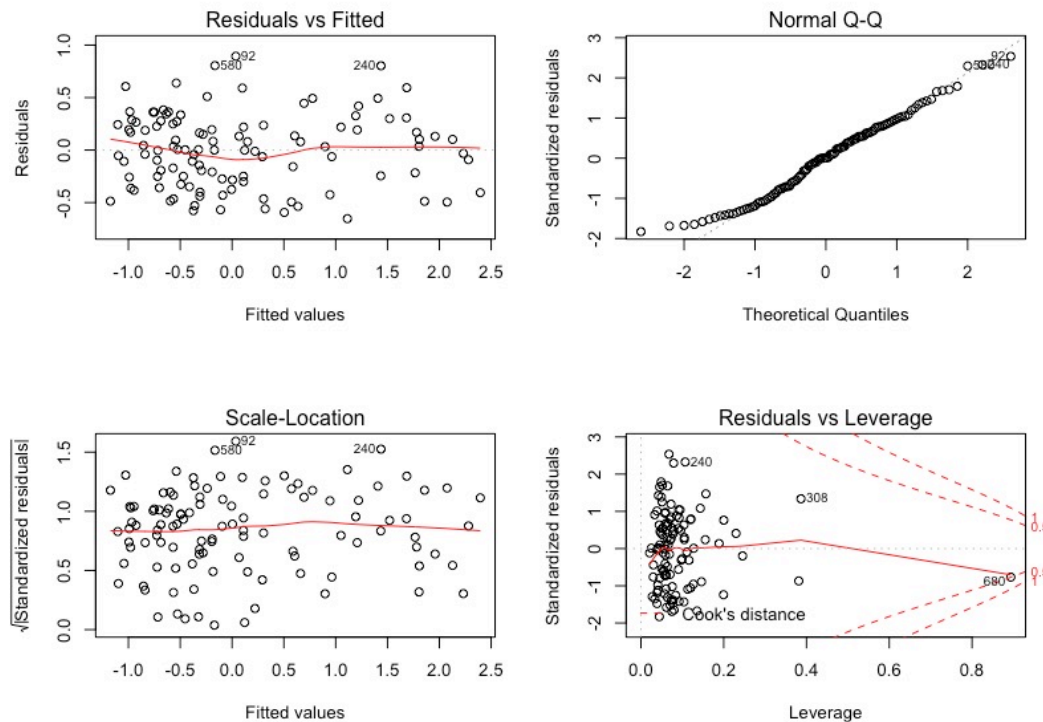
Note: The histogram shows that residuals are normally distributed, which was also confirmed by the results of a formal test. Shapiro-Wilk normality test results ( $p = 0.160$ ) suggest that the null hypothesis that the data (in this case, residuals) come from a normally distributed population cannot be rejected. Breusch-Pagan test results ( $p = 0.566$ ) suggest that the null hypothesis of homoskedasticity cannot be rejected (i.e., there may not be a heteroskedasticity problem).

Figure 1-13: Independence of residuals plot for model 3 (open data-judiciary interactions)



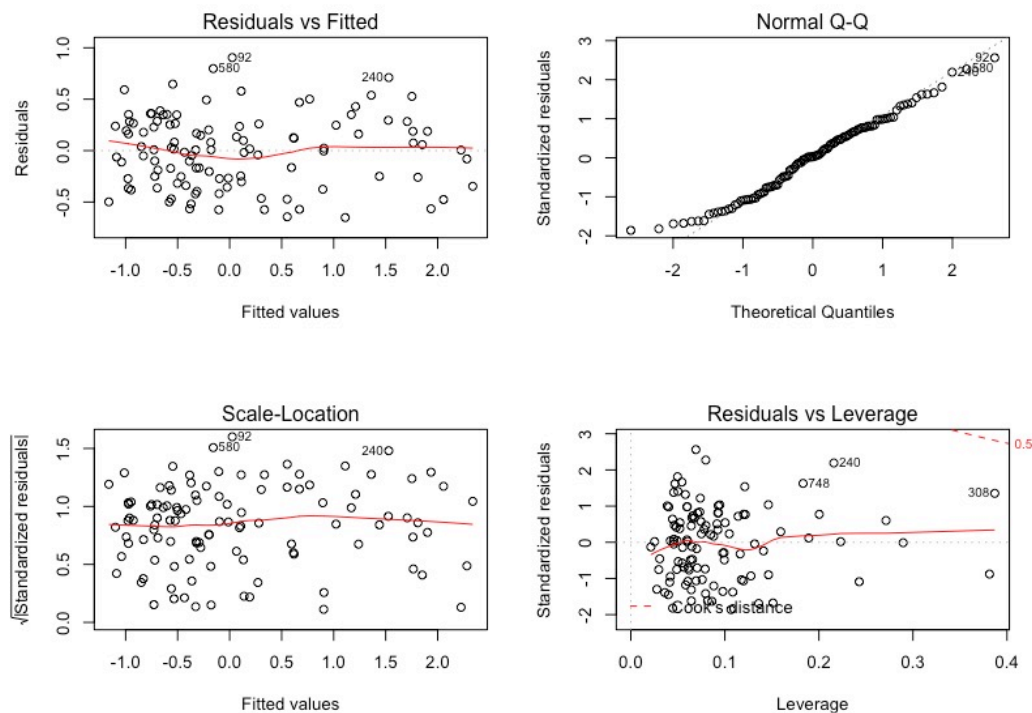
Note: The plot shows no serial pattern in the way the residuals of the model appear. It indicates that the residuals are independent of each other.

Figure 1-14: Diagnostic plots of residuals for model 5 (FOI years-judiciary interactions)



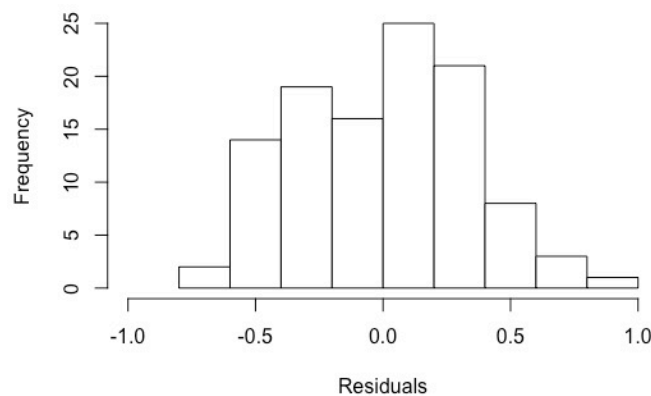
Note: Residuals vs. Fitted plot shows no distinct pattern; the line is relatively straight, horizontal at zero; the plot indicates that there is a linear relationship between the dependent and the independent variables, i.e. that the linear relationship assumption is met. Normal Q-Q plot shows that most residuals fall along the diagonal reference line; the plot indicates that the residuals are normally distributed, i.e. that the normality assumption is met. Scale-Location plot shows no distinct pattern; the line is relatively straight, horizontal, with residuals spread approximately evenly across the range of fitted values. However, as can be observed from Residuals vs. Leverage plot, observation no. 680 (Sweden) might be an influential observation. Although it does not fall beyond Cook's distance of 1, which was suggested as one of the cut-off points to identify influential observations, according to Cook (1977) any observations at values above 0.5 should be investigated as they might affect the slope of the regression line and affect the parameter estimates. Therefore, observation no. 680 (Sweden) was removed from the data for the analysis, as it was within Cook's distance of 0.5 to 1.





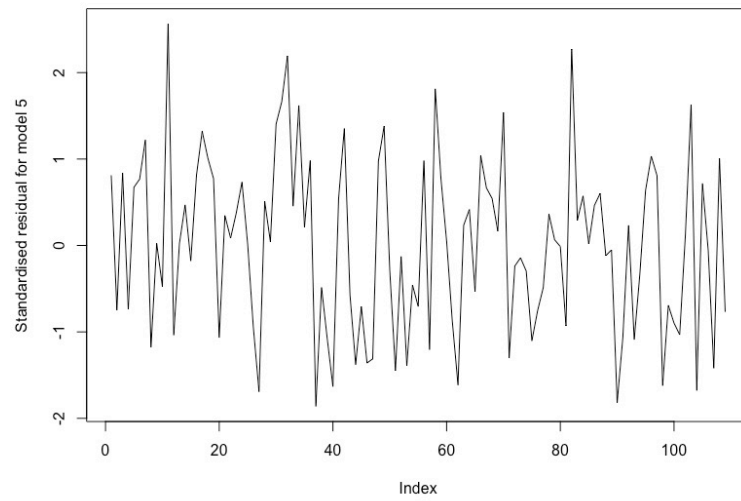
Note: Once the influential observation (no.680 - Sweden) was removed, the Residuals vs. Leverage plot shows that none of the points falls within Cook's distance of 0.5 to 1.

Figure 1-15: Histogram of residuals for model 5 (FOI years-judiciary interactions)



Note: The histogram shows that residuals are normally distributed, which was also confirmed by the results of a formal test. Shapiro-Wilk normality test results ( $p = 0.168$ ) suggest that the null hypothesis that the data (in this case, residuals) come from a normally distributed population cannot be rejected. Breusch-Pagan test results ( $p = 0.398$ ) suggest that the null hypothesis of homoskedasticity cannot be rejected (i.e., there may not be a heteroskedasticity problem).

Figure 1-16: (FOI years-judiciary interactions)



Note: The plot shows no serial pattern in the way the residuals of the model appear. It indicates that the residuals are independent of each other.

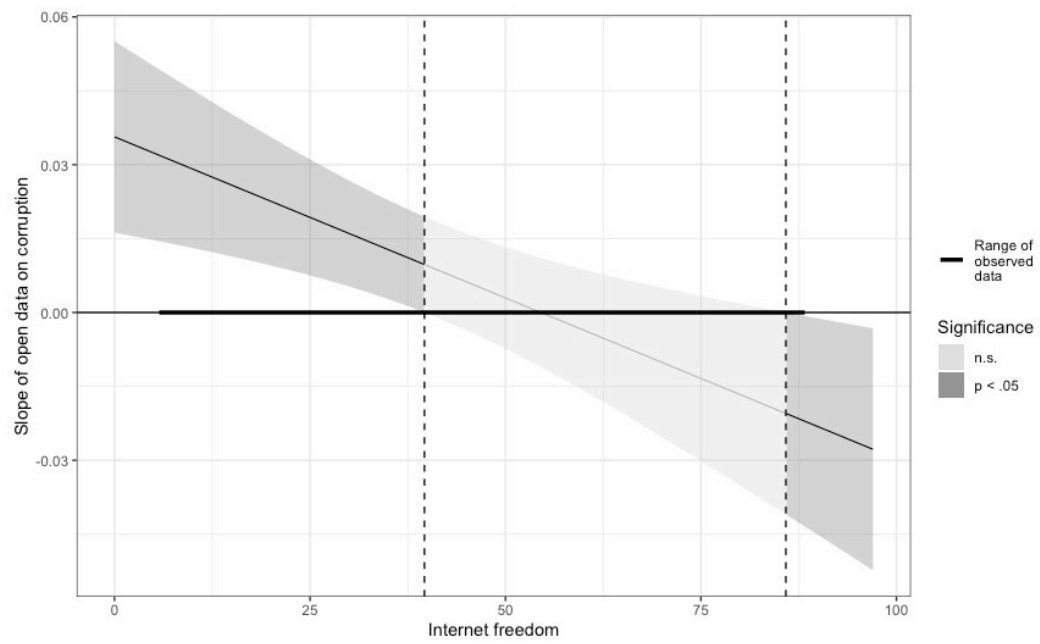
### 1.3 The moderating effects of internet freedom

Table 1-1: The associations between FOI laws, open data and corruption moderated by internet freedom

Control of corruption	Model 1	Model 2	Model 3
Open data	0.009 (0.005)	0.036*** (0.010)	0.009 (0.005)
Years of FOIA	-0.001 (0.007)	-0.005 (0.006)	0.012 (0.015)
Internet freedom	0.004 (0.007)	0.018* (0.007)	0.008 (0.008)
Murdered journalists	-0.252 (0.135)	-0.210 (0.122)	-0.246 (0.135)
Civil society	-0.032 (0.040)	-0.072 (0.038)	-0.032 (0.040)
Judiciary	0.221*** (0.060)	0.220*** (0.054)	0.224*** (0.060)
Free and fair elections	0.006 (0.004)	0.006 (0.003)	0.005 (0.004)
GDP	0.026*** (0.006)	0.018** (0.006)	0.023** (0.007)
Open data*Internet freedom		-0.001** (0.0002)	
Years of FOIA*Internet freedom			-0.0004 (0.004)
Constant	-1.878** (0.672)	-2.141** (0.611)	-2.020** (0.687)
N	50	50	50
R <sup>2</sup>	0.853	0.884	0.857
Adjusted R <sup>2</sup>	0.825	0.858	0.825
Residual Std. Error	0.404 (df = 41)	0.364 (df = 40)	0.404 (df = 40)
F Statistic	29.847*** (df = 8; 41)	33.911*** (df = 9; 40)	26.638*** (df = 9; 40)

\* p < .05; \*\* p < .01; \*\*\* p < .001

Figure 1-17: Conditional effect of open data on corruption as function of internet freedom



## 1.4 The OGP membership as an instrument

Table 1-2: The association between FOI laws and open data, and corruption (IV)

Control of corruption			
	Model 1	Model 2	Model 3
Open data	-0.013 (0.011)	0.023 (0.014)	-0.031 (0.028)
Years of FOIA	0.004 (0.002)	0.001 (0.002)	0.003 (0.002)
Press freedom	-0.003 (0.004)	0.023 (0.012)	-0.001 (0.005)
Murdered journalists	-0.110 (0.157)	-0.012 (0.167)	-0.116 (0.153)
Civil society	-0.073* (0.029)	-0.066* (0.028)	-0.071* (0.029)
Judiciary	0.201*** (0.039)	0.198*** (0.037)	0.089 (0.133)
Free and fair elections	0.014** (0.004)	0.017*** (0.005)	0.015** (0.005)
GDP	0.030*** (0.007)	0.022*** (0.006)	0.026*** (0.006)
Open data*press freedom		-0.001* (0.0003)	
Open data*judiciary			0.004 (0.004)
Constant	-1.074* (0.471)	-2.564** (0.780)	-0.721 (0.617)
N	110	110	110
R <sup>2</sup>	0.833	0.850	0.840
Adjusted R <sup>2</sup>	0.820	0.837	0.826
Residual Std. Error	0.428 (df = 101)	0.408 (df = 100)	0.422 (df = 100)
Weak instruments: open data	12.993***	8.634***	10.467***
Weak instruments: open data*press freedom		6.284**	
Weak instruments: open data*judiciary			15.761***
Wu-Hausman	5.162*	2.339	3.175*
Sargan	NA	NA	NA

\* p < .05; \*\* p < .01; \*\*\* p < .001

Table 1-3: The results of the first stage regression

	Open data
OGP non-member	-9.233*** (2.562)
Years of FOIA	0.068 (0.047)
Press freedom	0.075 (0.110)
Murdered journalists	7.782* (3.330)
Civil society	-0.571 (0.768)
Judiciary	-0.884 (0.976)
Free and fair elections	0.273*** (0.068)
GDP (in 1000 USD)	0.559*** (0.093)
Constant	14.328 (12.462)
N	110
R <sup>2</sup>	0.685
Adjusted R <sup>2</sup>	0.660
Residual Std. Error	10.986 (df = 101)
F Statistic	27.456*** (df = 8; 101)
* p < .05; ** p < .01; *** p < .001	

## 1.5 Specific government datasets and corruption

Table 1-4: The association between the availability of different government datasets and corruption (OLS)

	Control of corruption
Data on government budget	0.002 (0.001)
Data on government spending	0.0002 (0.002)
Government contracting data	-0.003 (0.002)
Company register data	0.002 (0.001)
Land ownership data	0.005* (0.002)
Years of FOIA	0.001 (0.002)
Press freedom	-0.004 (0.004)
Murdered journalists	-0.235* (0.115)
Civil society	-0.063* (0.026)
Judiciary	0.209*** (0.033)
Free and fair elections	0.008** (0.002)
GDP (in 1000 USD)	0.023*** (0.003)
Constant	-1.141*** (0.397)
N	110
R <sup>2</sup>	0.888
Adjusted R <sup>2</sup>	0.874
Residual Std. Error	0.359 (df = 97)
F Statistic	63.974*** (df = 12; 97)

\* p < .05; \*\* p < .01; \*\*\* p < .001

## 2 Notes

<sup>1</sup> For examples of the transparency rhetoric in the UK, see the following commentaries by the representatives of the Conservative Party: Francis Maude, the then Minister for the Cabinet Office and Paymaster General's editorial in the Guardian [here](#) or David Cameron, the then Prime Minister's opinion in the Daily Telegraph [here](#). For examples of the transparency discourse in the US, see, for example, the Memorandum for the Heads of Executive Departments and Agencies on Transparency and Open Government [here](#).

<sup>2</sup> Under the UK presidency, the G8's summit (Group of Eight consisting of leaders from Canada, France, Germany, Italy, Japan, Russia, the UK and US) focused on government transparency.

<sup>3</sup> See judicial decisions by ECtHR cited in References.

<sup>4</sup> See <https://www.economist.com/special-report/2010/02/27/data-data-everywhere>.

<sup>5</sup> Bentham was also an early proponent of free speech and press. He believed that that responsibility and accountability of political leaders could be secured only when the public has means to discuss their attitudes and actions freely and openly, which meant that the government should refrain from silencing criticism despite potential reputation harm (Schofield, 2006). In his view, publicity would not only ensure accountability, but it would also reinforce trust. Bentham was well aware that individual critical voices may be easily suppressed or too weak to hold political leaders accountable and recognised early on the role of a free press in ensuring government's answerability and responsibility. He argued that the invention of the printing press amplified the power of information. If this information was published in the newspapers, thanks to its "regularity and constancy of attention" its influence increased (Schofield, 2006). For a more detailed account, see Schofield (2006) and original Bentham's work (1999, 2001).

<sup>6</sup> The term armchair auditor was frequently used by the former UK Prime Minister David Cameron and Francis Maude, then Minister for the Cabinet Office and Paymaster General. For more contextual information, see <https://www.theguardian.com/commentisfree/2009/may/25/david-cameron-a-new-politics1>.

<sup>7</sup> Based on the data from September 2019, Facebook has 2.26 billion monthly active users, Youtube has 1.90 billion users, Instagram and Chinese multi-purpose messaging, social media and mobile payment app WeChat 1.00 billion users each, Tumblr 624 million and Chinese video-sharing social media TikTok has 500.00 million users.

<sup>8</sup> See Open Data Barometer at [https://opendatabarometer.org/?\\_year=2017&indicator=ODB](https://opendatabarometer.org/?_year=2017&indicator=ODB) and Global Open Data Index at <https://index.okfn.org/>

<sup>9</sup> See <https://info.worldbank.org/governance/wgi/pdf/cc.pdf>.

<sup>10</sup> The article on the rise of deforestation in Brazil is available [here](#). The government platform, monitoring deforestation and publishing satellite data, is available [here](#).

<sup>11</sup> Birchall (2015) and O'Neill (2006) argued that publishing more data will bring more revelations on mismanagement of public resources to the light, and if these are left unpunished, the public will become even more sceptic and mistrustful of the government, perceiving it as corrupt. In particular, data on contracts may reveal cases of suspicious use of public resources, overspending, illegitimate expenses etc. Such revelations might turn the public to become disillusioned of politics in a broader sense and see corruption in a country as pervasive. Bac (2001) suggested that access to contracting data might also improve outsiders' ability to detect the patterns of corrupt behaviour and increase their incentives to replicate these behaviours for personal gain. He proposed that this might eventually lead to an increase in corruption, which contradicts prevalent theories of transparency. A practical example of this would be if increased access to data on contracts would allow detecting which public agencies award procurement contracts repeatedly to the same companies, and how their value differs from comparable contracts in other public agencies. Recognising such agencies may encourage establishing connections at these agencies to offer bribes in exchange for awarded procurement bids. However, I observe nothing to confirm these speculations.