STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
BE551 Degree Project in Economics
Spring 2023

Does Corruption affect Fatalities from Natural Disasters? The Case of Floods

Felix Ling (24972) and Lalo Saleh (25189)

Abstract. Previous research has identified a positive effect of corruption on the fatalities of earthquakes and natural disasters in general. To test whether this relationship holds for other types of natural disasters, this paper aims to investigate the effect of corruption on flood fatalities using a panel dataset of 1 080 major flood events in 89 countries over a 33-year period. Seeking to provide additional explanatory value of the findings, the paper further investigates the possible transmission mechanism of flood magnitude. The research specification utilises a preliminary regression to account for the endogeneity of corruption before employing its fitted values in the Fixed Effects Poisson model. In contrast to previous findings, the results indicate that corruption has no statistically significant impact on flood fatalities, either directly or indirectly via flood magnitude. The results highlight the need of studying each type of natural disaster individually to identify the factors determining the severity of the resulting consequences for each particular type.

Keywords: Flood fatalities, Natural disasters, Corruption, Institutional arrangements

JEL: D73, O17, Q54

Supervisor:	Anders Olofsgård
Date submitted:	May 14, 2023
Date examined:	May 24, 2023
Discussants:	William Stålhök & Albin Nilsson
Examiner:	Johanna Wallenius

Acknowledgments

We would like to express our sincerest gratitude to our supervisor at the Stockholm School of Economics Anders Olofsgård (Stockholm Institute of Transition Economics (SITE)) for his invaluable feedback, guidance and constructive criticism during the undertaking of this study In addition, we are grateful for our friends and families that supported us throughout the whole process. Any remaining imperfections in the study are entirely the authors own.

TABLE OF CONTENTS

1 Introduction

Natural disasters have consistently resulted in significant human, ecological and economic loss throughout history. Major recent events – such as the 2020 South Asian, 2021 European, 2022 Pakistani floods, and the devastating earthquake in Turkey and Syria causing more than 58 000 casualties – have lifted the material and human cost of disasters to the front line of public attention. During recent years, disasters have frequently been brought up during various policy debates regarding national security, preparedness and ex ante disaster mitigations, especially in relation to the current issue of climate change (IPCC, 2021a) and the consequent pattern of climatic events that academia predicts will accompany it[1]. The predicted escalation in frequency and magnitude of natural disasters will certainly increase the pressure and responsibility of a country's institutions to adapt national and local disaster risk reduction strategies fit for their purpose (UN, 2023).

Institutions have a vital role in shielding citizens from deaths caused by different types of natural disasters through for instance preventative and mitigating measures such as adequate infrastructure quality, good planning and execution of disaster management, as well as through proper land usage. Multiple studies have reported that less democratic countries and countries with higher income inequality have a higher national death count by natural disasters, with further evidence indicating countries with weaker institutions suffer more fatalities through vulnerabilities in their institutional arrangements (Kahn, 2005; Kellenberg & Mobarak, 2008). Previous research has given different suggestions of mechanisms for why this is the case with the most preeminent being corruption. Fatalities of natural disasters have been linked to corruption and in particular, corruption within the built and infrastructure industry (Anbarci et al., 2005; Anbarci et al., 2007; Gleason et al., 2022), and government corruption (Kahn, 2005; Ambraseys & Bilham, 2011). Lacking enforcement of building codes, the development of non-resilient settler developments in vulnerable locations, and insufficient zoning are among the suggested mechanisms through which corruption could raise the number of fatalities. Several previous empirical studies have reported a significant, positive effect of corruption on natural disaster fatalities. However, an absolute majority of those studies have solely focused on earthquakes or natural disasters in general. The question is whether this significant relationship is isolated to a particular disaster type or not.

Even though it is the most frequent type of natural disaster, no previous work has focused solely on floods. Therefore, this study examines whether an aggregated relationship between corruption and disaster fatalities also holds for the case of floods to evaluate if corruption has a general effect on disaster fatalities, or is specific to certain disaster types. This study contributes to closing the research gap in this unexplored area by broadening the understanding on how vulnerabilities in institutions affect consequences of major flood events. Specifically, we analyse if a country with a lower level of corruption suffers fewer fatalities from major flood events, as well as suffers floods of lower magnitude than a country

[1] Changing sea, land and air temperatures, rising sea levels, changing patterns of rain, snow, extreme heatwaves and droughts, with an unstable climate are all likely catalysts of future weather-related events. However, for now, there is very little evidence that geo-physical disasters may also be affected and altered by global warming.

with a higher level of corruption to test if there is an indirect effect. To address these questions, we utilise data from the Dartmouth Flood Observatory (DFO), collecting a sample of 1 080 major flood events in 89 countries over the period 1985 to 2017. Unlike previous researchers in this context who typically use the Negative Binomial model, we employ the Fixed Effects Poisson estimator to account for all unobserved heterogeneity as it is concluded to be the more robust count data model amongst the alternatives. Additionally, the endogenous nature of corruption is accounted for by employing an instrumental-like approach where we predict corruption using exogenous variation, obtaining fitted values that are used as the dependent variable in lieu of the observed values in the final analysis.

Our results indicate, in contrast to most other previous studies on the effects of corruption on natural disasters, that there is no statistically significant direct effect of corruption on fatalities from flood events when relying on the three most commonly used corruption indices[2]. Furthermore, we find no robust evidence that corruption has an indirect effect on flood fatalities via the possible transmission channel of having an impact on the flood's magnitude. The results give a first indication that corruption has differing effects on different types of natural disasters. Hence, the determinants of the induced consequences of each disaster type need to be examined separately in order to fully understand how their impact can be mitigated.

The remainder of the paper is organised as follows. Section 2 provides a review of relevant previous research regarding flood vulnerability and corruption, as well as anecdotal evidence suggesting a relationship between corruption and flood fatalities. Section 3 presents our research specification and hypotheses. Section 4 and 5 provides further information regarding the employed data and the empirical methodology respectively. Section 6 presents the estimated results which are further discussed in section 7 alongside suggestions for further research. Lastly, section 8 concludes with the main findings of the paper.

2 Literature Review

This section presents strands of natural disaster related literature connected to the prevailing analysis. First, we present relevant research on the subject of what variables are important in terms of assessing flood vulnerability. What follows is a general definition of corruption and an explanation of its different forms. Thereafter we present recent cases where corruption likely had an impact on fatalities from flood events as anecdotal evidence and discuss the potential transmission channels. Lastly, we present previous literature on how institutional arrangements can affect the socioeconomic impact of disasters in general and lastly narrowing it down to corruption and floods specifically.

[2] The corruption indices used are from the International Country Risk Group, Transparency International and the World Bank.

2.1 The Case of Floods

Flood events are defined as the overflowing of water on land that is normally dry which can be induced by for instance heavy rainfall, insufficient drainage channels, large amounts of snow melting quickly, or when dams and levees break. Damaging floods may occur with only a few decimeters of water depth, or they could cover a house completely in extreme cases. Floods can occur within minutes and last for days or weeks hence showing great variability. Floods are viewed as the most common and widespread of all climate and weather-related natural disasters (NOAA, 2023b) with multiple factors determining their impact.

2.1.1 Determinants of Flood Vulnerability and Fatalities

Floods are a complex interplay of hydrology, climate, and human management. In addition to the extent and intensity of precipitation either independent or associated with other weather phenomena[3], other main factors for river, pluvial, urban and flash floods include the following. First, stream morphology i.e., the physical features and processes of flowing water that are limited by the waterbody's pathway (IPCC, 2021b). Second, river and catchment engineering, which refers to human intervention in the national course, characteristics, and behaviour or flow of rivers, with the intention of yielding protection against major flood events, but can cause the collapse of dams. An example is the 2009 Indonesian Situ Gintung dam collapse caused by poor maintenance and heavy monsoon rain, resulting in floods that cost at least 100 lives (ReliefWeb, 2009; BBC, 2009a; 2009b). Lastly, there is poor land-usage with unfavourable and hydrologically unsuitable land-cover characteristics such as the use of floodplains for developing new settlements, buildings and agriculture which explains why overpopulation is considered a contributing factor to flooding. Urban flooding often occurs as a result of infrastructure like drainage systems being overwhelmed by heavy and/or prolonged rains. Additionally, ground that has not been paved over will have better absorption abilities to mitigate the impact of floods. However in modern society, large areas with hard surfaces are becoming increasingly common and infrastructure is required for rainwater and meltwater to flow away efficiently as there is less unpaved ground to absorb it.

Disasters occur when different sorts of natural hazards meet vulnerability, typically due to lacking construction practices and infrastructure which is largely caused by man-made decisions or malpractice. Several different decisions that impact vulnerability including site-location of settlements, management and maintenance of flood prevention measures are clearly within the societal election. In addition, the specific design, development methods and materials employed to construct infrastructure and buildings are within this purview of choices (Green, 2005). A substantial factor, especially within construction, affecting vulnerability, risk and resilience is construction malpractice, commonly a result of corruption with large amounts of public funding for different infrastructural and building projects being channelled into the hands of private parties. In some countries it amounts to as much as 40

[3] Severe thunderstorms, hurricanes or tropical storms. There is not always a 1-to-1 correspondence between a flood event and an extreme precipitation (rainfall) event. Neither is this the case between change in extreme precipitation and changes in floods events, because as mentioned, flood events are affected by a plethora of factors.

percent of the industry (Gleason et al., 2022). The sole existence of private contractors does not necessarily mean that corruption or poor conduct is encouraged, but in certain circumstances with for instance certain institutional and societal arrangements, there exists a major risk if profits can be increased by cutting corners.

While there is a wide range of accepted relevant indicators and characteristics presented in previous literature, the true factors that determine flood vulnerability are to a large degree site-specific and hazard dependent. A case study by Müller et al. (2011) discussed various potential determinants of flood vulnerability, the most critical of which will be explained below. Firstly, a building's positioning in relation to the street level determines the exposure of the building to any flood event, where a low position implies increased susceptibility to damage (Schneiderbauer, 2007; Clark et al., 1998). Secondly, the quality of flood protection measures on buildings can alleviate effects of, for instance, heavy raining to avoid damages by altering precipitation and the flood's stream morphology. In the case study, households and buildings that had suffered from the impacts of floods improved precautionary measures, but only after a flood event had occurred and not prior, showing that in certain areas preventative protection measures were not seen as important. Connecting to flood frequency, experiencing floods increases the inhabitants' sensitivity to the problem at hand which generates a willingness to undertake private flood mitigation measures and a positive attitude towards preparedness even if it requires large financial resources to implement. Therefore, an increased experience likely creates a 'learning by doing' impact where a higher availability and increased knowledge pool of both flood hazards as well as community and private protection measures would lower and diminish the vulnerability (Cardona, 2003; Tanhueco & Velazquez, 2005). Thirdly, the materials used to construct a building determines its physical frailness towards a given flood event and dictates its resistance to damages, depending on the sheer force of the flood. Lastly connecting back to poor land-usage and paving natural grounds, the higher the amount of green spaces per building block and in the area in general, the higher the retention potential is and thus yielding higher protection (Bronstert et al., 2002).

With regard to flood fatalities, they are mainly inflicted by society's attitude, behaviour, decision making and measures in the long-term which of course differs from one society to another (Kelman, 2004) . For example, a significant portion of fatalities in Europe is due to risk-taking behaviour (Bierens & Brons, 2006). This was highlighted during the Côte d'Azur, France 2015 floods where eight individuals died as they tried to bring their cars out of underground parks with Vinet (2017) reporting that most likely these individuals did this after the issued flood warnings, and presumably managed to save their cars during several previous flood events. In the U.S., the majority of fatalities is due to obstacles and vehicle-related deaths (Jonkman & Kelman, 2005; Priest, 2009) when people attempt to drive across flooded bridges, roads or streams, being caught in sites or attempting to rescue others, with Coates (1999) showcasing similar behaviour in Australia. Thus, a large portion of flood fatalities are connected to victims' own unskilled, hazardous behaviours and unnecessary actions (Petrucci, 2022).

As explained, there are many widely-accepted indicators that determine the vulnerability towards flood events but a potential substantial factor influencing vulnerability, risk and resilience, that has not yet been widely investigated for the case of floods, is corruption.

2.1.2 Corruption and its Different Forms

There are different definitions of what corruption is. The European Commission defines it as 'any abuse of entrusted power for private gain' (2023) whilst the World Bank (2020; 2021) narrows it down to 'the abuse of public office for private gain', where the wide range of private gains include accepting bribes, fulfilling political and economical favours to peddling influence. Another common definition is given by Klitgaard (1988), with corruption materialising in institutions with assets that can be exploited by employees for their own private benefit to the cost of the body politic and individuals at large. Hence corruption is viewed as a principal-agent problem, where the basic elements allowing such behaviour to thrive can be summarised by the prominent Klitgaard formula:

$$Corruption = Monopoly + Discretion - Accountability \tag{1}$$

The formula explains the basic determinants of corruption and essentially states that corruption increases with the potential private benefits available to the corrupt and decreases with the risk of being caught alongside the severity of the potential consequences. This has been empirically proven by a number of papers (Banerjee & Duflo, 2020; Reinikka & Svensson, 2011; Olken, 2007). Moreover, corruption is often associated with political instability and authoritarianism (Alexander, 2017) and its concept is complex, nuanced and deeply embedded in society (Gleason et al., 2022). In some developing countries corruption amounts to a large fraction of GDP and it is also common in developed countries in the form of, for instance, defence officials selling contracts for personal gain, and local zoning officials being bribed to rezone (Becker & Stigler, 1974; Shleifer & Vishny, 1993).

Furthermore, there are different kinds of corruption. In corruption-without-theft, government representatives demand a bribe above any required fees where the government still receives its fees as state income. For example, construction companies pay the state to hire and send official building inspectors for the companies' constructions but bribe the inspectors on the side in order for their insufficient buildings to be approved. In many cases, a construction firm might need several permits and inspections from different agencies e.g., fire, water, and police which, depending on country-specific norms, could result in multiple bribes. In corruption-with-theft, the government representative demands a bribe but does not transfer the government its fees (Shleifer & Vishny, 1993). For example, customs agents who collect a bribe instead of collecting the actual customs fees robs the state of its income. Within this kind of corruption, both the buyer and seller benefit since it yields an opportunity for these government services to be produced more cheaply. Thus, it is seen as the sale of government property and services for personal gain. Lastly, state capture involves the endeavours of enterprises to shape and influence the underlying legislation and norms through private and/or corporate payments to public officials, thereby using political influence to deform legal arrangements and policy-making frameworks to acquire private gains with potentially damaging consequences to society at large (Hellman et al., 2000).

2.1.3 The Impact of Corruption on Flood Fatalities

To illustrate how flood fatalities can be affected by corruption, we present some anecdotal evidence, starting with the 2022 Pakistan floods. Leaving approximately one-fifth of Pakistan under water, the floods resulted in 1 739 casualties (World Bank, 2022; ReliefWeb, 2022). It was the deadliest flood event in the country since the 2010 floods which killed nearly 2 000 people. Even though international climate scientists agreed that global warming caused the flood event to be up to 50 percent worse, some of the contributors to the severity of the flood event were local to Pakistan itself (Harvey, 2022; Fickling, 2022; Mir, 2022).

Factors such as unauthorised constructions and illegal land usage in flood-prone areas have exacerbated the risk of flooding in Pakistan. This in combination with an uncontrolled population growth has caused an increasing number of Pakistanis to live in flood-prone areas with local governments and provincial irrigation agencies misusing and misdirecting public funds directed for flood protection management. Therefore, these parties are causing a lack of maintenance of canals and waterways. All issues have at some level been caused by corruption and a lack of accountability of those in charge (Glencorse & Yaseen, 2022). Many of the most important lines of defence against floods in Pakistan are projects of the British colonial-era such as the large Sukkur Barrage which is a system of dams and canals that divert the waters of the Indus river to irrigate the arid southern Sindh Province. However, the majority of such infrastructure is in a poor state of repair due to years of underinvestment in maintenance and incomplete new undertakings, corruption and disputes between the country's provinces about the allocation of water and public funds.

Moreover, Pakistani policemen and government officials regularly accepted bribes in order for the ‘timber mafia’ to fell trees. It operated hand-in-glove with local forest officials and politicians exploiting opportunities of corrupt conduct for their own vested interests, which explains why deforestation had grown so strong over the years and how legal action for cutting down trees was escaped (Mukhtar, 2021; Shaik & Tunio, 2014). Approximately, 80 million trees had been illegally cut in protected areas during the three years prior to the 2022 floods. Connecting back to the importance of natural prevention measures, riverline forest used to line the banks of the Indus which naturally paved way for utilising their beneficial absorbing capabilities of the flood stream’s ferocity as a first line of defence (UPI, 2010). Altogether, the lack of sufficient flood protection measures, diversions of flood-related funding into private hands (corruption with theft), in combination with other corrupt practices of assembly methods, non-existing maintenance of flood barriers and inappropriate land usage via the subsequent unauthorised siting of settlements (corruption without theft) increased the flood vulnerability of the growing population in Pakistan’s flood-prone areas.

Another case is the 2022 Nigerian floods that caused more than 600 fatalities. While the Nigerian government blamed heavy precipitation and excess water release from the

Cameroonian Lagdo dam, experts yet again stated that global warming, poor planning, and mismanagement of funds worsened the flood event (Ojewale, 2022). Besides climate change, another major component was the vulnerability of the infrastructure and development approach, which commonly causes societies to end up with such disasters (Ukomadu, 2022). Furthermore, Nigerian dams were intended to be constructed to reinforce the Lagdo dam, but their construction was never completed. Incomplete projects and lack of zoning allowed the construction of settlements in flood prone areas with poor irrigation, explaining why Nigeria was and still is ranked among the bottom 20 nations in its readiness to adapt and ranked highest in vulnerability to an increased frequency of natural disaster (ND-GAIN, 2023). It is necessary to examine how corruption via unaccountable spending of public funds could also have contributed to the problems. The Nigerian government established the Ecological Fund in 1981 with the objective being '..to have a pool of fund that would be solely devoted to the funding of ecological projects to ameliorate serious ecological problems nationwide.' (Ojewale, 2022; OSGF, 2023; EPO, 2023). Given that Nigeria is considered a high risk nation for fraudulent payoffs (GRP, 2023), the N548 billion put aside for the ecological funding account between the years 2012 and 2021 were deemed to be highly vulnerable to fraud, corruption and exploitation. This created and still creates major consequences for Nigerian citizens for which the funding was intended to assist and protect.

Several judicial cases of corruption regarding the administration and management of the fund had caught the Nigerian parliament's attention where state and federal authorities were suspected of diverting resources with evidence that corrupt spending of the fund existed across all tiers of government (Ojewale, 2022). The 2019 NEI Transparency Initiative Audit illustrated how government agencies embezzled and diverted approximately N4.35 billion between 2013 and 2015 to unqualified expenditures that the fund was not created to support with the report concluding that corrupt conduct of the Ecological fund has had a long history in the country (NEITI, 2019). Several high ranking government officials were indicted and charged of corruption, money laundering and bribery as a result, with the recent occurring in 2021. Such corruption exposed the Nigerian population to several preventable risks in terms of the human-induced parts of flooding with poor urban planning practices, insufficient environmental infrastructure made with substandard materials, and with incomplete projects being major contributing factors. Furthermore, fraud and corruption has undercut climate change interventions in for example flood-prone areas and decreased the scale of adapting preventative infrastructure measures. Even though the location of some Nigerian States in rainforest areas makes them susceptible to flood events, the consequences of the floods could have been significantly reduced with a proper and well-managed Ecological fund.

The aforementioned countries are not alone in this matter. China, Malaysia, India, and Peru are examples of other nations struggling with similar issues. The 2023 Peru floods exposed the country's corruption associated with past reconstruction efforts. Specifically after the devastating floods in 2017, large financial resources were not being spent as intended on necessary projects to prevent such disasters occurring again such as the maintenance of river defences and drainage channels, which could have mitigated or prevented flooding and casualties. The flood event also brought attention to the local and regional authorities' inability with regard to disaster prevention. Similar to Pakistan and Nigeria, over 100 incomplete and delayed projects related to disaster prevention have been identified in eleven Peruvian regions. Had these been completed on time, the region's resilience to the latest flood events would likely have been more robust. The major player in explaining this poor performance is yet again corruption, with the Comptroller General's office identifying almost 13 000 instances of ill-management, embezzlements, irregularities of project undertakings and procurement contracts involved (Ampuero, 2023; PSG UK, 2023). One must remember that hazards are natural, disasters are not.

To conclude, there is anecdotal evidence suggesting that corruption indeed has had an effect on flood fatalities. The question that remains is whether this effect is isolated to the case of specific countries or if there exists a general and aggregated causal inference on a larger scale amongst different countries across the globe, and if this effect is direct or through different indirect causal pathways.

2.1.4 Potential Transmission Mechanisms

Below, a few potential transmission mechanisms are discussed to highlight how the causal pathway behind an effect of corruption on flood fatalities might look. Thereby, pathways for an indirect effect may be identified alongside identifying the direct effect of corruption.

Magnitude: As the cases showcased, fraudulent and corrupt conduct were mainly concentrated within the built environment in terms of incomplete and underfinanced protective infrastructure projects with the already constructed preventative measures being ill-managed and not up to standard. Therefore, this kind of behaviour has and should undermine the resistance to major floods, amplifying the impact of the flood event. As will be explained in section 4.1, we defined flood magnitude as the logarithm of the product of the factors severity class, area affected and duration in days. All three factors are created naturally but they differ on the degree of impact by human interventions. The flood's severity class is unlikely to be humanly induced and impacted since a country's government cannot control the amount of precipitation or the force of floods created as the aftermath of tropical and hurricane storms. However corruption could potentially have an impact on the flood's duration and the area affected. If authorities provide insufficient flood protection measures, the flood's stream morphology will be unaltered and not mitigated, allowing floods to cover larger areas for longer periods of time. As follows, the number of individuals exposed to the flood will likely increase which increases the risk of yielding higher flood fatalities.

Population: If the usage of flood zones for residential purposes has broad links to corruption as was the case in the Pakistan and Nigeria examples above, the result would likely be that corruption causes a higher number of people to be exposed to increased flood risk. With more people exposed to such risk, the disaster potential is enhanced and an increased number of fatalities may follow. For example, in the Chinese city of Wuhan, the excessive urban development and overbuilding in areas previously being lakes and rivers was partly to blame for the over 180 casualties during the floods of 2016. With the number of lakes dropping significantly from 127 to 40 as they were being filled up to build cheap, and sometimes illegal, constructions, the city's ability to effectively absorb flood water had diminished dramatically (Huang, 2016). Moreover, the implementation of Wuhan's flood protective measures, e.g., dykes to keep floodwaters away, were being held back by corruption. One local official responsible for a 1 billion yuan dyke construction from 2005 to 2015 was charged with accepting bribes of 1.6 million yuan. One of the dykes that should have received maintenance and upgrades by the project was breached by the flood event in 2016. Consequently, during developments corruption may cause an increased population to live in flood-prone areas with insufficient protective measures, which can have an indirect impact on flood fatalities.

Frequency: Although corruption has no effect on the predetermined natural hazards of for instance extreme rainfall, man-made decisions can impact whether a natural hazard turns into a disaster. Both the potential lack of maintenance or non-existence of flood protection measures such as flood barriers, and an increased population vulnerable to floods resulting from corruption will likely increase the number of reported floods. Thereby, corruption may increase the frequency of floods and hence risk increasing the number of flood fatalities in total. Note that this mechanism would not necessarily increase fatalities per flood, only the total number of flood fatalities.

2.2 Institutional Framework and Socioeconomic Impact

The next relevant research strand considers how a country's institutional arrangements could mitigate the negative effects of natural disasters. We first look at studies examining this effect on natural disasters in general, before narrowing down to corruption, and lastly considering studies specifically relating to floods.

2.2.1 Natural Disasters in General

Anbarci et al. (2005) developed a theoretical model explaining how collective societal action – a form of informal institutions – in various forms such as the enforcement of strict building codes can reduce the damages caused by major earthquakes. Additionally, the model states that the probability of efficient collective societal action increases with income per capita as well as the level of income equality. Empirically, they found that income has a negative relationship with earthquake fatalities, and that income inequality has a positive effect, being in line with the forecast of the model. Similarly, Kahn (2005) found that countries with higher income inequality suffer more fatalities resulting from natural disasters in general, alongside evidence that high-income and democratic countries with higher quality and transparent institutions endure significantly less natural disaster related deaths. Kahn (2005) further proposed that one underlying mechanism behind the institutional effect could be government corruption, with results from Anbarci et al. (2007) supporting this proposal. Moreover, Skidmore and Toya (2007) reported that countries with smaller governments, more open economies, more complete financial systems, and higher levels of education suffer less deaths and lower economic damage from natural disasters. Povitkina and Sjöstedt (2017) narrowed their focus and studied small island developing states (SIDS), typically viewed as particularly vulnerable to natural disasters, and found that government effectiveness strongly reduces the number of people killed or affected by natural disasters in SIDS. However in contrast to Kahn (2005), they found that democracy has no such effect when government effectiveness is included in the model. Moreover, fiscal decentralisation is commonly suggested as a method to improve government effectiveness through a more efficient provision of local public goods. Escaleras and Register (2012) displayed that fiscal decentralisation is indeed associated with lower fatalities following natural disasters in developing countries thus in line with Povitkina and Sjöstedt's conclusions in SIDS.

2.2.2 Natural Disasters and Corruption

In previous literature, a commonly suggested connector between the consequences of natural disasters and corruption is corruption specifically within the construction and infrastructure industries, which are further commonly found to be the most corrupt in the world (Gleason et al., 2022). Ambraseys and Bilham (2011) reported that 83 percent of all casualties from building collapses caused by earthquakes over the past 30 years occurred in countries that are anomalously classified as corrupt. An additional notable contribution, closely related to the topic examined by this study, is made by Anbarci et al. (2007) who found that public sector corruption has a positive effect on fatalities of earthquakes, further supporting previous findings. The paper argues that the main mechanism behind the effect is substandard building practices as resulted from public sector corruption in the construction industry. By bribing inspectors, developers are able to use substandard practices to gain abnormal profits by reducing construction costs. Moreover, the physical process and layering of materials involved in construction covers up any potential malpractice, thus making it difficult to discover. Following substandard construction practices, buildings are at greater risk of collapsing once an earthquake hits and hence more people risk losing their lives as a result. Furthermore, the paper's theoretical framework argues that the decision of a corrupt inspector to accept a bribe or not depends on monetary incentives that can be negated with good institutions However, it is less obvious how corruption in the construction sector would impact the consequences of floods, especially if considering fatalities as most floods are seemingly unlikely to cause even substandard residential buildings to collapse. In fact, a literature review by Gleason et al. (2022) concerning the extent to which natural hazards are worsened by corruption in the built environment found that most research is indeed focused on earthquakes with more research being needed in the case of floods and storms. The review further suggests that corrupt practices may lead to the development of residential areas in substandard land in flood zones, thus increasing the flood risk exposure to these settlements.

Others have attempted to estimate the reverse relationship specifically relating to corruption, i.e. if natural disasters induce corruption (Escaleras & Register, 2016; Yamamura, 2015). The results point towards that natural disasters yield opportunities for corruption, with Yamamura reporting a stronger effect in more developed countries. Similarly Wenzel (2021), showcased that more severe drought exposure is followed by more corruption with the effect holding for subsamples of developing and developed countries. As a natural disaster strikes, public resources flow in to aid the affected areas. Governments will further assist with multiple urgent tasks such as search and rescue, the need of emergency medical treatment and to rebuild damaged infrastructure. As follows, the aftermath of natural disasters can create an environment of chaos, distrust[4] and moral hazard with huge sums of capital in the centre. For example, Latip et al. (2018) showcased several shortcomings in the Malaysian flood disaster management system with opportunities of corruption occurring at every phase. Riskier phases include the response and recovery phase, especially activities involving funds, donations and reconstruction, where a lack of transparency is highlighted as one probable cause amongst several others. A discussion of the potential reverse causality that may follow from this relating to our study is held under section 4.2.

2.2.3 Flood specific

Although institutional arrangements appear to mitigate the impacts of natural disasters in general, such conclusions may be oversimplified in the case of floods, albeit research on how institutions can mitigate the effects of floods is limited. Ferreira et al. (2013) found that governance, represented as an aggregated index consisting of seven governance indicators from the ICRG including bureaucratic quality and democratic accountability[5], has no significant impact on flood fatalities. To hypothesise the causal pathway of an effect, their study refers to White et al. (1975) who pointed out that there is an evident risk that increased flood protection and forecasting capabilities promotes permissiveness, causing poor land usage of floodplains which may increase the disaster potential. This would imply that while improving flood management systems may reduce the frequency of floods, the permissiveness that comes with it may increase the magnitude of floods when they occur, hence meaning that such systems may not have an unambiguously positive effect. The prediction by White et al. (1975) appears to hold according to Ferreira et al. (2013) if considering income, as it was found to have a linear negative effect on flood frequency and a U-shaped quadratic effect on flood magnitude. However, their governance index was not found to have a similar significant effect, making it unclear whether the prediction by White et al. (1975) is purely income driven, or whether good institutions also promote permissiveness. With regard to corruption, it could also be the case that the opposite relationship holds (Gleason et al., 2022), i.e. that corruption may exacerbate the permissiveness of public officials to use floodplains.

3 Research Specification

The different strands of previous literature presented in the prior section investigated the relationship between natural disasters, institutions and corruption. However, there still exists several gaps for further research to fill. One gap is that no previous study has attempted to estimate the effects of corruption on flood fatalities, despite floods being the most frequent type of natural disaster (Jonkman, 2005). Although flood events in general kill fewer individuals than earthquakes in single events, they affect a large number of people with the

[4] Carlin et al. (2013) find that people who suffer damages from natural disasters have lower faith in their country's institutions, which may undermine their influence in the long run.
[5] The other governance indicators included in the aggregated index are control of corruption, law and order, government stability, lack of ethnic tensions and lack of religious tensions.

WHO (2023) illustrating that floods affected more than 2 billion people worldwide between 1998-2017. Moreover, Jonkman et al. (2008) estimated that floods caused 175 000 casualties and affected 2.2 billion people from 1975 to 2001. Therefore, analysing the relationship between corruption and flood fatalities is a useful and important line of inquiry. For example, floods cause more casualties each year than tornadoes and hurricanes in the U.S. (NOAA, 2023b). This section intends to further define the research gap we intend to fill and the effect of corruption the study aims to examine, before concluding the section by presenting our research questions with its accompanying hypotheses.

3.1 Delimitations

Most of the research undertaken relates to earthquakes and the role of corrupt practices in the area of construction that induce building collapses. The large focus on earthquakes is of no surprise due to the severe, striking and visceral outcomes of poor construction following seismic disasters. However to bridge existing gaps, research that estimates the impacts of corruption with regard to other disaster types besides earthquakes is needed. Even though floods do not offer the same plain and apparent casual passage as ruptured buildings following earthquakes, these estimates are of interest nonetheless. Consequently, our purpose is to establish if there exists a causal effect of corruption on fatalities of flood events, and to compare the results with relevant previous studies in other disaster contexts. The main objective is focused around identifying a potential direct effect and we intentionally focus on a decomposed indicator of governance to better be able to analyse the disaggregated effects of corruption on fatalities from flood events. Secondly, a vital delimitation of this study is that of the aggregate level of corruption of the countries during the time period of our sample. It would perhaps have been better to use measurements of corruption of the province affected in order to receive more accurate estimates at the local level as there may exist regional differences within countries. Even if there might not be a significant impact on the aggregate national level, there might exist a causal impact of corruption on flood fatalities within countries i.e., at the regional or county level as previous studies have indicated (Calgaro & Lloyd, 2008; Marks, 2015). However, due to lack of data when it comes to local level corruption within countries, such an empirical investigation was not possible at this time.

Thirdly, the study only focuses on the potential transmission mechanism of magnitude rather than in detail identifying all structural mechanisms behind a potential effect, hence leaving the other two mentioned mechanisms for future research. When examining the indirect effects of corruption on flood fatalities through a larger affected population, we argue that it would be better to employ more precise measures of the population in the affected areas than the measures collected by us for reasons that will be explained in section 4.6. With regard to frequency, an entire new data set would have to be created with each observation being the number of floods in a particular country in a particular year for all entities and years, rather than each observation being one flood event as in our data. Our data is simply not adapted for estimating corruption's indirect effect through this mechanism.[6] Hence, we leave this transmission mechanism as a future consideration. A final important delimitation concerns

[6] See section 4 for a comprehensive description of the data employed.

the specific sample of flood observations employed for the analysis. We believe that the most severe floods are of highest interest as mitigating its consequences yields the greatest benefit and hence the sample used is intended to include the most severe floods. As most researchers, our ambition was to include as many flood observations as possible but due to the merging of multiple datasets, only flood observations for countries that had complete data from all datasets were used. The countries included have a broad variation of flood fatalities and level of corruption and thus we still obtain a representative sample since the number of observations lost were only approximately 65. The results presented are based on the sample of countries during the specific time period and chosen range of magnitude as will be explained in section 4. For a full list of the sample countries, see Appendix 1.

3.2 Academic Contributions

We believe the contributions of the performed study to be multiple. Most notably, this study will aid in closing a large gap in the current literature and create a more broad understanding of how corruption affects fatalities of different types of natural disasters while contributing to the growing research specifically relating to flood events. Some empirical work examined the impact of variables viewed as transmission channels but these did not connect it to corruption. We add to the literature by including a potential transmission mechanism, magnitude, which could provide additional explanatory value of any findings on how corruption impacts fatalities from floods and thus support findings in previous literature. Furthermore, this study contributes to the growing set of model approaches employed in disaster research by undertaking a quantitative approach based on observational data in which the study will be the first to employ the chosen model estimator, Fixed Effects Poisson, to analyse the effect of corruption on natural disasters. This makes the modelling approach in our study unique from an empirical view and further adds to the general discussion of modelling approaches within this research context. Thirdly, the study assists in policy making with regard to the eleventh Sustainable Development Goal: Make cities and human settlements inclusive, safe, resilient and sustainable (UN, 2023). An expected consequence of climate change is an increased number of natural disasters and rising sea levels according to climate research, which highlights the importance of flood and risk mitigating measures (Logan, 2023).

3.3 Research Question and Hypotheses

With the academic contributions presented, we restate our research focus in more precise terms, leaving us with the main research questions this empirical study attempts to answer:

Does a country with a lower level of corruption suffer fewer fatalities than a country with a higher level of corruption when affected by a major flood event?

Does a country with a higher level of corruption suffer floods of greater magnitude compared to a country with a lower level of corruption?

Given the previous literature and arguments presented below, our hypothesis is that corruption will have no statistically significant effect on fatalities caused by major flood events. We further hypothesise that a country's corruption level will have no statistically significant effect on the magnitude of floods suffered by the country. Below, we reiterate the main arguments for the presence of a significant intuitive effect and respond with arguments against an effect existing before arriving at our conclusion.

The anecdotal evidence from section 2.1.3 illustrates certain cases where corruption is believed to have exacerbated the impact of floods to the point where fatalities may have increased as a result of it. Amongst others, the aforementioned Pakistani 'timber mafia', widespread diversion of funds in Nigeria and insufficient or poorly maintained flood prevention measures in Peru are all examples of this. Additionally, building codes may be overly lax following state capture, potentially increasing the flood vulnerability of buildings which may lead to increased flood damages should a flood occur. Such is also the case if corrupt contractors economise on materials such as cement or iron bars to achieve abnormal profits at the cost of increased building vulnerability. Another case could be the adoption and monitoring of regulations concerning land-usage since poor land usage could increase flood risk and its disaster potential. Making flood plains available for commercial usage may further distort the natural flood protection such as green spaces and could cause more severe floods. In addition, it may increase the number of people living in areas vulnerable to floods, thus exposing more people to a higher flood risk. While several pieces of reasoning have been presented suggesting the existence of an effect of corruption on flood fatalities, either direct or indirect via magnitude, we now move on to discuss why we hypothesise that such an intuitive, significant effect does not exist.

With the determinants of flood vulnerability in mind, the impact corruption may have on land usage in flood-prone areas and in terms of yielding substandard flood protection measures is likely to increase the damages and economic loss created by major flood events. While this is the case, it appears less likely that flood vulnerabilities caused by corruption on an aggregated level would yield a higher number of fatalities from flood events since it is hard to see how it could cause an increase in deaths. Approximately 75 percent of flood fatalities are estimated to be the result of drownings (WHO, 2023). This makes it difficult to imagine a general, significant causal pathway from malpractice while constructing buildings or ill-maintenanced flood infrastructure to flood fatalities even if the possibility in specific regional cases could exist. For instance, while substandard material usage in a residential building may imply that a flood causes greater damage due to the building's decreased resilience, it will likely not cause it to collapse, making it difficult to imagine a pathway for it to cause death. The main realistic scenario would be if the flood event would be classified as a flash flood with a strong stream morphology and force, that induces landslides in flood-prone areas which in turn causes buildings to collapse or be dragged along the current, which is deemed as a very rare occasion. As follows, the most likely broad link between corruption and flood fatalities appears to be decisions regarding the development of substandard land in floodplains. With drownings being generally accepted as direct flood disaster casualties, if a house located in a floodplain becomes flooded and its residents drown, the questions that arise are why they were living there, and why seemingly few precautions were taken to prevent death and the flood itself. Examples of possible solutions would be proper disaster planning and land usage (Kelman, 2004). On a broader level however, it is less probable that the main causal reason for why the solutions were not being enacted to be the result of corruption as suggested by Gleason et al. (2022). It appears more likely that it is the result of permissiveness and overoptimism due to the perceived high-quality flood protection measures as suggested by White et al. (1975), as well as unnecessary risk-taking behaviour by decision-makers and victims that contributes significantly to flood fatalities (Jonkman & Kelman, 2005).

As mentioned, two of three factors constituting the flood magnitude are seen to be able to be impacted by poor decision-making and corrupt conduct, being the affected area and duration of a flood. The driver behind such an effect would likely be lacking flood prevention infrastructure or other permissive behaviour such as the deforestation alongside the Indus river as explained above in the Pakistan example. However, if the collapse or malfunction of major flood prevention infrastructure such as dams or flood barriers would pose a major risk to settlements nearby, they would presumably have heavier monitoring even in more corrupt societies. Additionally, major collapses of such structures are highly uncommon. Thereby, while lacking practices of flood disaster management following corruption could increase the magnitude of floods, it is deemed unlikely to cause or exacerbate major floods such as flash floods[7] on a larger scale. Thus, it is deemed that if corruption impacts flood magnitude, it would probably mainly affect the economic damages caused by the flood rather than fatalities. There could exist inference in some specific cases, such as in the Pakistan case, but since there exist a plethora of other determinants to the magnitude of flood events that are site-specific and local, it is difficult to see that corruption would have a significant, and widespread effect on magnitude. For example, precipitation is seen as the most important factor in creating floods and corruption clearly has no effect on how much it rains directly or as a consequence from storm surges.

4 Data

To empirically analyse the effect corruption has on flood fatalities, an unbalanced panel dataset[8] was constructed with observations on the number of individuals killed in flood events, as well as variables capturing the physical magnitude of the floods, affected

---

[7] According to NOAA (2023a), 'Flash floods are very dangerous floods that can happen with little or no warning. When there is more rain than the soil can absorb, the excess water quickly runs into rivers and creeks, overwhelming storm drains and ditches and causing a flash flood. Flash floods can cause water to rise significantly in a short amount of time.'

[8] The mechanics of fixed effects estimation with an unbalanced panel data are overly different than with a balanced panel. If T_i is our number of time periods for a cross-sectional unit (here a country), one simply uses the T_i for implementing the time-demeaning with one degree of freedom lost for every cross-sectional observation because of the time-demeaning. It is simple to realise that countries for which the sample only has one observation play no role in a fixed effects analysis since the time-demeaning for such observations yields all zeros and therefore not used in the estimation. Moreover, dropping observations with only one time period will – contrary to a criticism some makes against fixed effects – not cause bias or inconsistency since $T_i = 1$ observations contributes nothing to our learning about β_j in the FE environment.

population's exposure and vulnerability for floods during 1985-2017. Anbarci et al. (2005) and Anbarci et al. (2007) focuses solely on 269 and 344 'significant' earthquakes, respectively, with a 6+ magnitude on the Richter scale for their analysis of the deadly interaction between corruption and earthquakes, with the reasoning that even the most poorly constructed structures rarely fail following earthquakes of lower magnitude. Since there is no universally accepted magnitude measure of magnitude for floods as the Richter Scale (earthquake) or wind speed (storms), we adhere to definitions of previous studies (Cunado & Ferreira, 2014; Ferreira et al., 2013), which appears as the most reasonable to use. We use floods above the third quartile (6.08) within the range of magnitudes of the full DFO archive which gives an unbalanced panel dataset[9] of 1 080 floods in 89 countries with complete data between 1985-2017. Table 1 presents descriptive statistics on the data, which are more thoroughly defined, along with their sources, in Appendix 1.

4.1 Flood Fatalities and Characteristics

For the given purpose, a flood-specific disaster dataset was used to investigate the questions at hand. Most previous studies on disasters have drawn and used data from the Emergency Events Database (EM-DAT) maintained by the School of Public Health at Université catholique de Louvain, Belgium. However, with the context at hand, data were drawn from the Dartmouth Flood Observatory (DFO). Hydrologists use it more often than the EM-DAT, in part because it provides more detail on flood events, including the flood's magnitude, which enables investigating the impact of different indicators of flood events as controls. The DFO is made possible by data acquired from NASA, the Japanese and European Space Agencies, with funding support from NASA and the European Commission, through the Global Disaster Alert and Coordination System (GDACS) project of the Joint Research Centre in Ispra, Italy (DFO, 2023).

The DFO records data on large flood events, defined as events inducing significant damage to structures or agriculture, long reported intervals (measured in decades) since the last similar event, and/or the number of fatalities. Flood fatalities recorded in disaster databases such as the DFO and EM-DAT, are commonly from severe injuries and drowning during the time period of which the flood event is active (Combs et al., 1998; Jonkman & Kelman, 2005). The fact that data on fatalities are solely available for countries that indeed experienced flood events could construct a sample selection problem, namely incidental truncation. Moreover, the fact that the DFO only includes data on what they define as large floods further supports that data might be truncated. However, several studies (Ferreira et al., 2013; Cunado & Ferreira, 2014) have tested the DFO data by applying the truncation test by Wooldridge (2010) where the studies failed to reject the null that there is no truncation bias.

[9] A more difficult issue is determining the reason for the panel being unbalanced. With countries, for example, data on key variables are sometimes missing for certain years. Provided the reasoning for missing data for some country i is not correlated with the idiosyncratic errors, u_{it}, the unbalanced panel incurs no problem. In this case, the reason for missing data is due to the number of floods a country has had during the specified time period with a magnitude of 6.08. It would not matter if this cutoff existed or not since different countries have more floods than others hence the full DFO archive is unbalanced by itself.

TABLE 1. Descriptive Statistics

Variable	N	Mean	Median	St. Dev.	Min	Max
FATALITIES	1 080	231.3	14.0	3 139.8	0	100 000
COR-ICRG	1 080	2.78	2.50	1.18	0	6
COR-TI	800	4.14	3.50	2.06	0.4	9.4
COR-WB	878	2.41	2.10	0.985	0.8	4.6
MAGNITUDE	1 080	6.71	6.60	0.497	6.08	8.49
FREQUENCY	1 080	6.10	4	6.07	1	32
RAINFALL (m)	1 080	1.04	0.79	0.67	0.028	3.50
GDPPC (Constant 2015$)	1 080	12 078	3 852	17 799	190.012	57 356
POPDENISTY	1 080	145.85	56.76	222.84	0.076	2 721.3
POPULATION	1 080	17 669 797	3 898 000	29 695 126	11 811	223 807 649

Note: The sample period covers 1985-2017 for 89 countries for floods of magnitude 6.08 or greater.
Sources: See Table A2 in Appendix 1

Such a test is used when the dependent variable is unobserved in parts of the sample and there is a reason for why it is not unobserved. An example in this case would be if the DFO would detect flood events with the affected country not giving any information of the number of people affected or killed. Therefore, in this case, the inclusion of a country in the sample depends on the country's decision, not the surveyor's decision. However, given the parties handling the DFO and the sources it uses to collect information, it is concluded that it is highly unlikely for it to cause significant sample selection bias. Moreover, in general, it seems unlikely for a country to withhold information of the number of fatalities caused by a flood event since the value it could gain from it with e.g., increased foreign monetary aid and assistance towards the affected areas. Another case might be due to technological and reporting limitations, especially in more poorer and less advanced countries, however, in this case the country does not purposefully choose not to share information due to internal constraints. An important note to make is that the dependent variable is observed for all flood observations in our sample data. Therefore, it is concluded that the main reason for why fatalities from other countries are not included is simply because those countries have not been affected by flood events according to the DFO standards[10]. Furthermore, this does not

[10] The same argumentation can be made for the EM-DAT and for the earthquake-related NGDC's Significant Earthquake Database, which contains information on more than 6,500 destructive quakes occurring worldwide since 2150 BC. As the title indicates, this is a catalogue of 'significant earthquakes'. To be included, a quake must meet one of the following criteria: cause approximately $1 million or more in property damages; have a Richter Scale value of 7.5 or greater; or cause 10 or more deaths. Similarly, EM-DAT excludes smaller destructive natural events. To be included in it, an event needs to fulfil at least one of the following criteria: *(1)* 10 or more people killed; *(2)* 100 or more people reported affected (typically, displaced); *(3)* a declaration of a

necessarily imply that there does not exist any correlation between selection in a certain period and the employed independent variables or specific country-effects, but it does mean we can exclude an existence of correlation between the idiosyncratic errors and selection.

FATALITIES: The main dependent variable is flood fatalities which is connected to the collected flood observations from the DFO. In compliance with previous literature on disaster fatalities, we choose to focus on major flood events as defined by the variable of magnitude. One could argue that economic costs of floods would be a better dependent variable, but determining the direct economic cost and damages of flood events, and other natural disasters for that matter, involves estimation uncertainties and is hence less accurate. Therefore, we adhere to methods employed by previous research and employ the number of fatalities. The number of fatalities within the sample are relatively small for most observations, with 22 per cent of events causing no fatalities, approximately a third killing no more than 3 people and 52 percent of them causing no more than 15 casualties. As Table 1 illustrates, the mean number of flood fatalities is 231, with a great level of variation since its standard errors were approximately 13.5 times larger. By investigating our dependent variable's frequency distribution, we can see that the high standard deviation and mean are induced by the large right tail, with 91 percent of observed values being below the mean.

MAGNITUDE: The database reports the specified magnitude of a certain flood as the logarithm of the product of three factors: the area (in km^2) affected by the flood × duration of the flood (in number of days active) × severity of the flood[11]. The expectation of this variable should be quite obvious since the number of fatalities of a given flood event should be greater for floods of a higher magnitude and power. Moreover, magnitude varies much less than flood fatalities and flood frequency since its standard deviation is smaller than its mean.

FREQUENCY: The variable of flood frequency is constructed as the count of floods in a country in a year. The expectation here is that there might be a 'learning by doing' reduction in the number of fatalities in a specific country that is regularly hit by large flood events, which is improving the country's ability to adapt to similar events in the future. The mechanism can be through a more robust flood management system and other channels societies can better prepare themselves for the onset of disasters and its determinants of causing direct damage. Table 1 showcases that, on average, there were just over six floods per country-year observation within our sample, however the statistic has large variation with a standard deviation nearly as large as the mean.

RAINFALL: We include a variable that controls for the annual mean precipitation in a country during a year. The CRU CY dataset consists of country averages at a monthly, seasonal and annual frequency, for ten variables, where precipitation is one, covering over 120 countries

---

state of emergency; or *(4)* a call for international assistance. No other paper among the many that has been read as preparation for this study has tested for incidental truncation except the studies mentioned above.

[11] Floods are divided by the DFO into three different classes of severity depending on the flood's estimated recurrence interval. Class 1 floods have a 10 to 20 year long reported interval, class 1.5 have a 20 to 100 year recurrence interval, and class 2 flood events have a recurrence interval greater than 100 years.

and land areas through the period 1901-2021 (Harris et al., 2020)[12]. An increased area of impenetrable land that is connected with larger urbanisation and population growth is thought to aggravate flood events through increasing the portion of precipitation that keeps running off instead of being absorbed into the ground. Since one of the causes of floods is heavy precipitation, we find it natural to control for when regressing magnitude on corruption as in Equation (6) presented below. Additionally, Ferreira et al. (2013) showcases that a country's yearly mean precipitation (rainfall) has no effect on flood fatalities and only has a positive, significant effect on the frequency and magnitude of flood events, albeit only a minor effect for the latter. The most relevant informational contribution of rainfall when discussing fatalities is its degree of force that causes major problems for residents, and since our magnitude is constructed to account and capture such a potential effect, we chose not to include it when regressing fatalities on corruption in Equation (3) and (4) below. Looking at Table 1, we see that, on average, a country had a mean national precipitation of 1 metre during the sample period, with its standard deviation being approximately two thirds the mean' size, indicating a high degree of variation.

4.2 Corruption

COR-ICRG: The primary indicator of the level of corruption used is from the International Country Risk Guide (ICRG), published by the Political Risk Services Group, as assembled by the Institutional Reform and the informal Sector Center at the University of Maryland, USA. The ICRG is a popular source of governance indicators, not solely restricted to the level of corruption, used in several previous studies (Anbarci et al., 2005; 2007; Ferreira et al., 2013; Yamamura, 2014). Beginning in 1984, the ICRG reports data on a broad range of over 100 countries, reducing the risk of selection bias (Kaufmann et al., 1999; Johnston, 2001), and its availability covers the entire time period of the DFO data, unlike corruption indicators from other sources. The variable ranges from zero to six, with higher values denoting less corruption. The indicator evaluates the overall level of corruption in government by evaluating whether higher level government officials are likely to require special payments and whether such inducements are anticipated throughout all levels of government (PRS Group, 2014).

COR-TI: As a robustness check, the corruption indicator produced annually by Transparency International since 1996 was also used. This measure runs from zero to ten with larger values indicating less corruption and is reported as a 'poll of polls' summary[13] of the level of corruption perceived within an economy (Transparency International, 2022).

[12] The dataset was released on May 26 2022 but the method and data are based on the CRY TS data from 2020 with Spatial averages calculated using area-weighted means. The dataset was developed and has been subsequently updated, improved and maintained with support from a number of funders, principally the UK's Natural Environment Research Council (NERC) and the US Department of Energy. Long-term support is currently provided by the UK National Centre for Atmospheric Science (NCAS), a NERC collaborative centre.

[13] According to Transparency International (2022), each country's score is a combination of at least 3 data sources drawn from 13 different corruption surveys and assessments. These data sources are collected by a variety of reputable institutions, including the World Bank and the World Economic Forum.

COR-WB: As a further sensitivity check, the corruption indicator produced by the World Governance Indicator of the World Bank is employed (World Bank, 2023c). This measure runs from zero to five with larger values indicating less corruption[14]. In all cases, indicator scales run nearly their entire ranges respectively in the sample data but only the primary indicator of ICRG contains observations with the cornervalues zero and six. We see on all three corruption indices that the mean level of corruption is under the mid-level of their respective scales, with their standard deviations being less than half the value of their respective means, indicating less variation within the sample and data points being more clustered around the mean.

4.2.1 Reverse Causality

One consideration that should be discussed is the potential of reverse causality, i.e., perhaps flood fatalities may cause higher levels of corruption. The research mentioned in section 2.2.2 is focused on how the number of natural disasters in a country impacts the level of corruption and opportunities for corrupt conduct in terms of resource allocation and subversion of financial measures to enterprises and the more well-off part of communities. For example, in the aftermath of Hurricane Katrina, 907 charges were made against entities, construction firms included, and individuals over a large range of crimes including emergency-benefit fraud, identity theft, procurement procedures fraud and public corruption to an estimate of US$ 2 billion (FBI, 2023; Goldenberg, 2006). However, it is less clear how the number of flood fatalities would have a causal pathway to increase corruption in a country since it should be the severity of the flood event itself and not flood fatalities that may lead to higher corruption. A more severe flood in terms of damages will probably require more resources and financial aid within the larger recovery phase that can be diverted for private benefits. While the severity of a flood and the number of fatalities are likely correlated, it is more likely that fatalities themselves have no causal impact on corruption. Additionally, we control for flood magnitude and frequency in our main model thus it is deemed to not be an issue.

4.3 Income

GDPPC: The indicator for income is the country's real GDP per capita converted to constant 2015 U.S. dollars as it is reported in the World Development Indicators (World Bank, 2023b). The indicator is expressed in natural logarithmic form, which is a common transformation in the literature, not restricted to that of natural disaster economics. Models were estimated with the logarithmic transformation and without it, and the produced results were not very sensitive to it and therefore adherence was taken to convention with the natural logarithm of GDP per capita. Even though it has been mentioned that GNI per capita is an indicator that more accurately measures the income of a specific nation's population, it was available to slightly fewer countries and years than GDP per capita, thus dictated the choice of income indicator. Moreover, the WDI did not have data on real GDP per capita in constant 2015 U.S.

[14] The World Bank's Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests, collected from multiple sources which are stated in the database. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5, which we have re-scaled to 0-5.

dollars for one Canadian and four Venezuelan observations in the sample data. However, it did offer values of real GDP per capita in current U.S. dollars. To include these observations, the five values of GDP per capita in current U.S. dollars were converted to constant 2015 U.S. Dollars according to the methodology applied by the United States Census Bureau (2022) and the U.S. Bureau of Labor Statistics (2023), which is thoroughly explained in Appendix 2.

Overall, GDP per capita is a relatively good proxy for all dimensions of development that may matter on a yearly basis. It is reasonable to expect that greater levels of economic development should allow for increased adherence to high-level building codes and other relevant regulations (Anbarci et al., 2007; Lewis, 2005; Gleason et al., 2022; Ambraseys & Bilham, 2011), better health-care infrastructure and yielding a greater level of self-insurance from disaster risk on the single individual's part. It has been shown that as poorer households are only able to utilise limited ex-ante risk-mitigating strategies, the households' ex-post risk-coping strategies are critical (Sawada & Takasaki, 2017). However, formal insurance mechanisms against natural disasters are relatively more limited in poorer nations in general. Another viewpoint is that since richer countries are deemed to be able to have top-standard prevention measures, an increased knowledge pool and higher technological adoption, it can create an overoptimism in its resilience toward flood events. In that way, one can argue that richer countries may allow for more housing in flood-vulnerable areas to a higher degree as suggested by White et al. (1975), exposing a higher share of the population to a greater disaster potential when a flood strikes. Moreover, heavier investment in infrastructure solutions in flood controls can reduce magnitude and frequency e.g., dams, whilst others such as levees and channelisations could increase river levels during heavy precipitation or events causing it (Ferreira et al., 2013; Criss & Shock, 2001; Pinter, 2005). Therefore, there are possible arguments in both directions with regard to the effect of income on flood fatalities.

4.4 Population in Flood-Affected Areas

POPULATION: Each observation entry in the DFO's data archive of large flood events has its own associated coordinates representing the regional area affected by that specific flood event. The DFO uses news and governmental sources to conclude the area affected. A measure of the population affected by each flood event was estimated by using the accompanying coordinates in connection to sources of subnational population. The primary source used was the Subnational Population Database constructed by the World Bank (2023a). Since the World Bank only has data for the period 2000-2016 and only covered a majority of sample countries, other sources were employed. For observations before 2000, and in countries that were not included, such as Australia, Canada and France, the population data were collected from sources such as Eurostat (2023), European Commission's Joint Research Center (2023), the respective country's national bureau of statistics, and national census bureaus. For example, for observations of the United States, the United States Census Bureau (2023) was employed.

The intended thought was to have population and population density of the most affected province as control variables, but an early realisation was that some major floods affected areas larger than some provinces (often in smaller countries). Therefore a discussion was enacted on whether one should take the most affected province (or other first administrative level) in any case or if one should apply the national population instead. It was concluded that the most affected province would be used because there can be wide variation in population density between provinces within countries. Moreover, the ensuing estimates of population affected and population density to a large flood event should be more proper than estimates based on country-level population measures attempted by previous empirical work.

POPDENSITY: The variable population density was created by dividing the population estimate with the land area in square kilometres of the most affected province where the data on land area stems from the same sources as previously mentioned above (United States Census Bureau, 2010). As mentioned by the Intergovernmental Panel on Climate Change (IPCC) (2007; 2021b) and Freeman et al. (2003), an increasing population in floodplains is seen an explanation for the observed growth in the number of reported flood events, because as floods enacts a larger impact on the residing population, the higher the probability it becomes that floods are reasonable reported[15].

4.5 Determinants of Corruption

As will be described in section 5.4, one of the complications that is to be addressed before reliable estimates can be produced is the potential endogeneity of the independent variable. Previous literature on the subject of corruption suggests a great amount of determinants to be used as instruments common within the empirical analysis. The idea is that these should be closely related to corruption, while being relatively uncorrelated with flood fatalities. The literature proposes four widely accepted categories of determinants, namely economic, political and environmental, social, and historical factors. These categories are well in line with the findings of Serra (2006) who tested the robustness of 28 common determinants of corruption across 60 different countries. Serra concluded that income per capita, extent of democratic institutions, political stability, share of Protestants, and a country being of English colonial heritage are the most robust determinants of a country's level of corruption with determinants exercising a negative influence. We utilise these contributions in combination with the approach of Treisman (2000) to define a number of institutional variables closely correlated with corruption but relatively uncorrelated with flood fatalities to be used.

GDPPC: The income per capita indicator is defined exactly as the one under section 4.3. As a country gains higher wealth, corruption will likely fall as the increased abundance of wealth allows for a reasonable increase in the chance of uncovering and penalising any illicit deals such as inspector-bribes. A concern worth mentioning regards the possibility that GDP per capita is endogenous. While economic development influences corruption, causation may also run in the opposite direction with corruption at least modestly impacting GDP measures

[15] See Appendix 2 for an explanation of why we do not include the share of forest (which has been used by other authors) in our regressions.

i.e., yielding reverse causality. However, since it is used as control in the main regression it should not be a major problem.

PROTESTANT: This variable indicates the percentage of a country's population that is of the Protestant branch within the religion of Christianity as of 1980 with La Porta et al. (1999) being the source of the measures. Since Protestantism is relatively more egalitarian and less hierarchical than its fellow branch Catholicism or other major religions such as Judaism or Islam, it is more probable that an increasing extent of Protestantism within a country should be negatively related with corruption. Acemoglu et al. (2001) concluded that settler mortality influenced the development of democracy which yielded more robust institutions, and in turn greater economic development. In contrast however, Woodberry (2012) finds, when controlling for Protestantism in the form of exposure to protestant missionaries, that settler mortality is statistically insignificant and further concludes that more protestant missions led to more stable democratic institutions which lead to more economic freedom that finally lead to higher average income, i.e. greater economic development.

ENGLISH: To account for differences in historical origins, a dummy variable is constructed that takes on the value of one for a country with English legal (common law) origin, again the study on quality of governments by La Porta et al. (1999) being the source for the variable. With the conclusion of previous literature in mind, having an English legal origin, which is based on judicial precedents, judiciaries typically positioned themselves with landowners counter to the sovereign, which was on the contrary to the common position in countries based on Roman, Germanic or Nordic civil law (Glaeser & Shleifer, 2002). Therefore, a negative relationship is to be expected with corruption.

DEMOCRACY: In order to control for the degree of governmental openness and democratic institutions, we use the democracy data offered by PolityV (2022). The variable scale runs from zero to 10 with lower values reflecting lesser degrees of openness, and ranks countries annually as to the general openness of their public institutions. It is typically expected that a higher general openness is associated with a lower level of corruption since enhanced democracy should yield a higher likelihood that the greater population will monitor public sector corruption with a better ability and thus the relationship is expected to be negative to corruption.

INTCONFL: A variable is made to reflect a country’s internal conflict by indicating the degree of political stability within a country and therefore control for differing political structures and environments. The variable takes on values from zero to twelve, with higher values indicating less political violence within a country and stems from the ICRG. This should take into account other differences e.g., ethnolinguistic since a country with higher ethnic division will enact decisions favouring the government officials' own ethnic group, but at the same time could play a role in balancing constituency blocs that disrupts corruption based on ethnicity from other internal blocs.

FEDERAL: For capturing differences in governmental structures, a dummy variable is included that takes on the value one when a country is relatively decentralised, and does not follow a unitary system of government i.e., governed as a single entity in which the central government is the supreme authority. A centralization could incur a type of oligopoly where various ministries collude to ignore each other's corrupt conduct. Additionally, different decentralised authorities empowered with a degree of supervision could potentially lessen corruption. Therefore, the nature of relationship relative corruption is expected to be negative.

DUMMIES: To account for differences between countries, we create five dummy variables indicating where a specific country is located. We adhere to the procedure of including a group of dummy variables, and include four of the five total in our specification, with their interpretation being in respect to the omitted dummy variable. We do this since country fixed effects are not possible to use for the estimation since some variables do not vary over time, i.e, they have no within country variation.

In conclusion, the chosen instruments are expected to exhibit a negative relationship with corruption, i.e., a positive correlation with the corruption indices since higher value indicates less corruption.

4.6 Limitations of Data

The data used are subject to limitations. First, since the measure of public sector corruption is vital to the performed empirical analysis, a remissness would occur if it was not expressly noted that as with all available measures of the level of corruption, data offered by the ICRG, TI and WB is subjective in that it is based on surveys. Nonetheless, Alesina and Weder (2002) and Anbarci et al. (2007) show that there is a very high correlation between each of the available measures of corruption and that each measure tends to be highly correlated over time. From a practical angle, previous empirical studies have illustrated that these subjective assessments of public sector corruption are vigorous in describing economic and social phenomenon connected to varying subfields, ranging from foreign direct investments decisions and economic growth (Habib & Zurawicki, 2002; Smarzynska & Wei, 2000; Mauro, 1995) to fatalities of traffic accidents (Anbarci et al., 2006).

Second, due to a lack of alternatives of subpopulation data and time constraints, the data used for subnational population is at the first-administrative level, which still is arguably a more reasonable estimate than the usage of country-level population statistics. Instead, one could use the coordinates in combination with overlays of gridded population maps (Anbarci et al., 2007; Ferreira et al., 2013) to more rigorously determine a more accurate estimate of the population affected by a particular flood[16]. However, such population grid data is either only available in intervals of five years or in inconsistent intervals depending on the source, and is not available for the entire sample period. Therefore, a calibration of an extrapolation with an

[16] For examples of how this can be done, see Appendix 2.2

appropriate curve must be constructed to receive values for the remaining years.[17] Additionally, employing the population grid method would not be feasible due to time constraints. Therefore, to ensure a reasonably valid estimate, we use subnational population at the first-administrative level as our population control with a further advantage being that data is available for each individual year for the majority of observations. For the few observations where data on subnational populations were not available for a particular year, it was estimated using the compound annual growth rate between a year before and after the observation where data were available. Although subnational population is argued to work as a proxy for the population affected by a flood event as a control variable, we argue that it is not suitable for investigating population as a potential transmission mechanism. Predicting the population and population density of a province or region within a country using the national level of corruption as a predictor would likely yield nonsense predictions as they should have no true relationship. Hence, such an estimation would require more reliable data of the true affected population of a flood, which is what was previously argued to be the potential transmission channel.

Third, the full sample includes floods dating back to 1985 and the alternative corruption indices of TI and WB does not have consistent individual measures for the years before 1998 and 1996 respectively[18]. Previous literature has, for example, assigned the 1998 value of the corruption index to all disaster observations between their samples starting period and 1997 (Anbarci et al., 2007). This adjustment will reduce the variability in the corruption measure and implies that there is no within-country variation in the independent variable before 1998 which is not ideal in a fixed-effects analysis. Therefore, we do not adhere to that method and we instead drop all observations before 1998 and 1996 for the respective alternative corruption indices when performing our sensitivity and robustness analysis which should still provide a good test of the results given by *COR-ICRG*. While these measures are not as broad as the primary measure, it does allow us to evaluate 800 and 878 of 1 080 floods respectively.

Fourth, if a flood event affects several countries, there is a possibility that the number of fatalities caused by that event could be linked to a single country that represented the majority of casualties, even though some fatalities occurred elsewhere. This could potentially bias the estimate. A potential remedy could be to drop all observations of flood events that occurred in multiple countries. However, an issue with such an approach could be that certain areas hit by, for instance, seasonal monsoon rain or hurricanes affecting multiple countries may become systematically excluded from the data which may bias the results. As it stands, the vast majority of all fatalities should be linked to the correct country in the data and thus this is not believed to have a major impact. Lastly, we intend to estimate the impact of corruption on the magnitude of a flood event where the results could be affected by the choice to only include flood events that have a magnitude equal or greater to the third

[17] For a further explanation of why this may pose a problem, see Appendix 2.2

[18] As pointed out by Alesina and Weder (2002) and Anbarci et al. (2007), a high degree of correlation is present across the different measures and sources of control of corruption indexes. In this case, the ICRG and TI indices have a correlation of 0.79 which is similar to previous work. Further, while the measure from Transparency international is available for the 1995-1997 period, it has been demonstrated that it has been relatively inconsistent prior to 1998 and therefore data is only used from this source starting from 1998.

quantile. To yield a more accurate overall estimate, one should include the full amount of observations from the DFO. However, the study is purposely focused on the most severe floods as these are deemed to be of highest interest with regard to a potential impact of corruption, but one should indeed note that the results are thereby mainly applicable to the most severe floods.

5 Empirical Methodology

This section motivates and presents the chosen methodology adopted to estimate the effect of corruption on flood fatalities and magnitude. We first discuss the issue of the regular choice of count data distribution and thereafter we present the reasonings for the selected econometric approach: Fixed Effects Poisson (FEP) estimator, followed by the model specifications. The section concludes by presenting model assumptions and limitations.

5.1 Flood Fatalities – A Form of Count Data

Flood fatalities are nonnegative integer count data i.e. the dependent variable is a counting number with the unit of our analysis being the number of individuals killed in a specific flood event. Linear regression models, estimated by Ordinary Least Squares can be used for count data with the estimated coefficient interpreted as the expected value of the dependent variable, conditional on the used regressors. However as with the Binary regression model, OLS regressions do not take into account the special structure of count data and can therefore yield nonsense predictions e.g., negative fitted values, which is not possible since one can not have a negative amount of fatalities (Stock & Watson, 2020). The two most widely used models to analyse count data are the Poisson and Negative Binomial regression models. Furthermore, since these models were used by most similar previous studies on fatalities from natural disasters, we adhere to them.

5.1.1 The Issue of Overdispersion

The Poisson distribution is dictated by the condition that the mean is equal to the variance but this restrictive condition has been demonstrated to be violated in a number of applications. Fortunately for us however, the Poisson distribution has a very beneficial robustness property that we take advantage of. Whether or not the condition is fulfilled, one will nonetheless receive consistent, asymptotically normal estimators of the sought coefficients. The variance is assumed to be proportional to the mean

$$Var(Y|x_j) = E(Y|x_j)\sigma^2 \qquad (2)$$

with $\sigma^2 > 0$ being an unknown parameter. If $\sigma^2 = 1$, then we have the regular 'Poisson case' that the conditional variance equals the conditional mean and thus no issues. However, if $\sigma^2 > 1$, it reflects that the variance is greater than the mean for all independent variables x_j and this is defined as overdispersion, where the variance is larger than the regular 'Poisson case'. Overdispersion is a phenomenon observed in a number of applications of count regressions[19]. A first hint that overdispersion exists is if when estimating a regular Poisson regression, the

[19] A more rare case is that of underdispersion which means that the variance is lower than the mean.

residual deviance is much larger than the degrees of freedom. Furthermore, one can calculate the dispersion parameter, and when performing this calculation, we received a parameter significantly greater than 1, thus confirming a first existence of overdispersion. Secondly, standard goodness of fit tests strongly reject the null hypothesis that there is no overdispersion when estimating an ordinary Poisson regression without fixed effects[20].

The issue with overdispersion is not that one receives inaccurate estimates but it does mean that the accompanying standard errors will be underestimated and in many cases strongly underestimated if the data are highly overdispersed. If one underestimates the standard errors, it yields an inflated type 1 error i.e., the p-value is too small and thus more significant than warranted by the data, which means that one's decision is more likely to be a false positive. Moreover, there can be different causes of overdispersion[21] of which the presence of unobserved heterogeneity is one of the most common i.e, there may be some underlying clustering or heterogeneity in the sampled population that causes this underestimation of standard errors. (Cameron & Trivedi, 2009; Dunn & Smyth, 2018)[22]. When estimating the regular Poisson regression and formally testing for overdispersion[23], we reject the null that there is no overdispersion at a five per cent significance level, being a third and final piece of evidence for the existence of overdispersion.

In conclusion, the existing overdispersion in our sample implies that one may need to fit an alternative model to the Poisson that accounts for overdispersion. A commonly employed alternative is the Negative Binomial model following the Negative Binomial distribution. However, this model comes with several shortcomings of its own and we argue thoroughly in Appendix 3.1 for why it is not ideal in our case.

5.2 Econometric Model

We now turn to the introduction of the robust Fixed Effects Poisson (FEP) model employed in order to accurately assess the impact of corruption on fatalities from major flood events.

5.2.1 The Fixed Effects Poisson Estimator

For models of a nonnegative dependent variable, one would prefer an estimator requiring minimal distributional assumptions with further relaxation of moment assumptions present in previous literature[24]. Wooldridge (1999) directly proved that the multinomial Quasi

[20] See Table A7 in Appendix 3 for results and a more thorough explanation.

[21] Others could be that there are predictor variables that have not been included in the model and thus yielding a mixture of different Poisson distributions i.e., a misspecification or there might be an excess of expected zero counts in the sample population which is called Zero-Inflation, both of which our chosen model approach can account for (Wooldridge, 2010; Ferrerira et al., 2013).

[22] For a more empirical explanation the reader is referred to Cameron & Trivedi (2009) CH 17.2. When events being counted arise in clusters or are mutually supporting in some way, it yields a positive correlation between underlying events hence resulting in overdispersion of the counts.

[23] See Table A6 Appendix 3

[24] Most previous studies regarding unobserved effects models were focused on standard linear unobserved effects under correct specification of the conditional mean, and strict exogeneity of the independent variables conditional on the latent individual effect. Less focus was on alternatives within the field of distribution-free estimation of nonlinear unobserved effects models since most methods for nonlinear panel data relied on the

Conditional Maximum Likelihood Estimator – better known as the Fixed Effects Poisson estimator – consistently estimates the conditional mean parameters, leading naturally to method of moments estimators that could improve the FEP estimator's efficiency. Most practical is that the FEP estimator is fully robust within the sense that the conditional mean assumption is the only assumption needed for consistency and asymptotic normality. Thereby, Wooldridge (1999) showcased that the FEP estimator is completely robust to every failure of the Poisson distributional assumptions except, of course, for having the correct conditional mean. The Fixed Effect Negative Binomial model, which is not considered a true fixed effects model (Allison & Waterman, 2002; Greene, 2007; Guimarães, 2008), does not come close to nesting the Poisson distributional assumptions unless in the case when the heterogeneity is zero in a panel data setting. Moreover, the FEP estimator suffers from none of the shortcomings mentioned in Appendix 3.1 with the additional advantage that testing for overdispersion is technically of no need. For example, in the case of having a short *T* panel dataset, testing for overdispersion is not an easy task to complete but because the FEP estimator is completely robust, there is no need to do anything even if overdispersion is found to be present (Wooldridge, 2010).

Additionally, the estimator is more robust and less sensitive to extreme outliers due to the employed regressional link which should take care of such an issue. One should tread carefully when removing potential outliers since removing legitimate outliers that are not results of error may cause the output of coefficients to differ significantly from the 'true' ones[25]. It will further cause loss of valuable information concerning the nature of the dependent variable data and would potentially disturb the exponential functional form following the conditional mean assumption. Furthermore, the FEP allows any kind of variance-mean relationship, thus allowing some units being overdispersed, some underdispersed, and some exhibiting both depending on the covariate values. One of the advantages relevant in this case is that the FEP allows for any kind of serial correlation with the only adjustment needed being to cluster the standard errors at the appropriate level. This is beneficial in our case for controlling for unobserved heterogeneity at the country level.

Moreover, the proof presented by Hahn (1997) of the FEP estimator's achievement of semiparametric efficiency, bound under the FEP model assumed by Hausman et al. (1984), further justifies its employment. A concern is that when including variables that are not of substantive interest – meaning they would not individually improve how the likelihood function is modelled – one could rapidly bias the rest of the sought estimated coefficients and standard errors i.e., the incidental parameter problem. Given that we do not observe every country for the 33-year period, and others are observed for a limited year until they 'leave' the sample, incidental parameters need to be taken seriously. However, due to its high level of

---

CMLE where the sum of the dependent variable over time is conditioned to remove the unobserved effect, for example the count data model of Hausman et al. (1984).

[25] Justification for such removal should come with information that is external to the sample data since the removal will otherwise invalidate standard theory. For example, is the value of fatalities impossible? Is the value collected in a special situation that might have caused measurement errors or mistakes? The only reason for removing such observations is in fact if they are incorrect or strongly unrealistic, taking into account the meaning of the data.

robustness under very general conditions, the FEP estimator does not suffer from this problem (Wooldrige, 1999; Silva & Tenreyro, 2010). Incidental parameters are not an issue if one does not have fixed effects, but even in a cross-sectional context the Poisson is more robust than the Negative Binomial. Using an estimator that addresses the consequences of overdispersion is a more robust approach than modelling the overdispersion with the assumption of a particular distribution e.g., the Negative Binomial (Greene, 2007).

To summarise, there are many reasons why one should employ the FEP estimator, with the main one being that it is fully robust to any distributional misspecification when estimating the conditional mean. Furthermore, its estimation is straightforward and the estimator controls for the full amount of unobserved country fixed effects. Therefore, we chose to employ the FEP estimator as our main regression model.

5.2.2 Model Specifications

We can now define the model used for the analysis. We estimated two versions of the main regression, with the first one not controlling for flood magnitude and the second controlling for it. The choice of whether or not to include magnitude as a control depends on if magnitude is viewed as exogenous or not. While a considerable amount of the variation in magnitude is likely weather driven, we are further investigating it as a potential transmission mechanism for an indirect effect of corruption. Therefore, both versions of the model are included in the analysis. The former, unconditional model, has the following specification:

$$F_{ijt} = \beta_0 + \beta_1 C_{jt} + \beta_2 G_{jt-1} + \beta_3 Q_{jt} + \beta_4 P_{ijt} + z_j + \theta_t + u_{ijt}\,, \tag{3}$$

where F is number of fatalities from flood event i in country j in year t, C is the level of corruption, G is a vector of income per capita, Q is flood frequency, P is a vector containing the population and the population density in the most affected province, with z_j and θ_t being unobserved effects for country and time, and lastly u_{ij} being the error term. Moreover, we lagged the income per capita variable by one year to allay potential endogeneity within the variable since the current income per capita could have been affected by the flood disaster itself. Several econometricians argue that employing conventional standard errors within panel data models could result in extreme underestimation if the error term is not independently distributed within entities with the result of overstating the significance of independent variables (Angrist & Pischke, 2009; Wooldridge, 2020). Therefore, we cluster the standard errors in our model by country to diminish this bias and its risk of independent variables being falsely accepted as significant. In accordance with standard goodness of fit tests and previous undertaken studies (Ferreira et al., 2013; Kellenberg & Mobarak, 2008) we estimated a specification with income taking a quadratic form. The country fixed effects, z_j , are treated as fixed, not random, and estimated using Equation (3). Fixed effects models are preferred to random effects due to the nature of our sample which is the complete set of countries with observed floods instead of a set of countries drawn at random (Wooldridge, 2010). Furthermore, we employed time fixed effects, θ_t, for any unobserved year effects in order to control for time-varying factors that affect all countries i.e., changes in reporting procedures and quality of flood information in the flood database which the DFO themselves have noted explicitly. If reporting has improved across all countries, the country fixed effects will not control for this effect, hence defending the usage of time fixed effects. Another reason is due to different weather phenomena that affect certain parts of the world during certain periods such as monsoon, hurricane and storm seasons in South East Asia, the Americas and Africa, or the climate effects of El Niño (Readfearn, 2023). Employing two-way fixed effects reduces the risk that the differences across countries and over time will perplex the effect of corruption e.g., more corrupt countries systematically underreports flood fatalities (Kahn, 2005).

The conditional model employed enabled an estimation of a model that isolated the direct effect of corruption, in other words, its effect for a flood of a certain magnitude. The specification for the conditional regression model has the same independent variables as Equation (3) with the difference being that we added the variable controlling for flood magnitude:

$$F_{ijt} = \beta_0 + \beta_1 C_{jt} + \beta_2 G_{jt-1} + \beta_3 M_{ijt} + \beta_4 Q_{jt} + \beta_5 P_{ijt} + z_j + \theta_t + v_{ijt}\,, \tag{4}$$

where M is the flood magnitude and v_{ijt} is the error term with all other notations being the same as the unconditional model previously presented. By entering the chosen variables into for example the conditional model we receive the following:

$$\begin{aligned} FATALITIES_{ijt} = \beta_0 &+ \beta_1 COR\text{-}INDEX_{jt} + \beta_2 GDPPC_{jt-1} + \beta_2 GDPPC^2_{jt-1} \\ &+ \beta_3 MAGNITUDE_{ijt} + \beta_4 FREQUENCY_{jt} \\ &+ \beta_5 POPULATION_{ijt} + \beta_6 POPDENSITY_{ijt} \\ &+ z_j + \theta_t + v_{ijt} \end{aligned} \tag{5}$$

5.3 Flood Magnitude

In addition to including a magnitude variable as a control in the main regression, we extended the analysis by investigating the effects of corruption on flood magnitude by estimating the following regression model:

$$M_{ijt} = \gamma_0 + \gamma_1 C_{jt} + \gamma_2 G_{jt-1} + \gamma_3 R_{jt} + \gamma_4 Q_{jt} + \gamma_5 P_{ijt} + z_j + \theta_t + v_{ijt}\,, \tag{6}$$

where v_{ijt} is the error term, R is the mean national precipitation, with remaining being the same as in Equation (3). A major difference, however, is that flood magnitude is a continuous variable and not a count as flood fatalities. Therefore, we estimate Equation (6) employing an OLS two-way fixed effects model with clustered standard errors at the country level. In addition to potentially having a direct effect on flood fatalities, corruption may have an indirect effect through the transmission channel of magnitude which makes this model of interest. The final model can be stated as:

$$\begin{aligned} MAGNITUDE_{ijt} = \gamma_0 &+ \gamma_1 COR\text{-}INDEX_{jt} + \gamma_2 GDPPC_{jt-1} + \gamma_2 GDPPC^2_{jt-1} \\ &+ \gamma_3 RAINFALL_{jt} + \gamma_4 FREQUENCY_{jt} + \gamma_5 POPULATION_{ijt} \\ &+ \gamma_6 POPDENSITY_{ijt} + z_j + \theta_t + v_{ijt} \end{aligned} \tag{7}$$

5.4 Endogeneity of Corruption

It is highly reasonable to proclaim that the chosen main independent variable is endogenous since corruption within the public sector is known to be highly correlated with a plethora of other, omitted institutional factors that would be explained by the error term (Kahn, 2005; Anbarci et al., 2007; Skidmore & Toya, 2007). The dependent variable could be correlated with this disturbance, leading to a violation of the zero conditional mean assumption, $E(u|x_1,x_2,...,x_K) = E(u|x_j) = 0$, as each independent variable would need to be necessarily exogenous for it to hold. Different approaches have been taken within the relevant previous literature, where some studies take this into account by performing a preliminary regression using exogenous variation with the corruption variable as the dependent variable (Anbarci et al., 2005; Anbarci et al., 2007), whilst other studies do not (Ferreira et al., 2013; Kellenberg & Mobarak, 2008**;** Yamamura, 2014).

5.4.1 Model to Determine Corruption

To account for this, we re-perform the main regression. The difference is that instead of the observed corruption value as the dependent variable, a predicted value, using an instrumental-like approach with exogenous variation, is added in lieu of the observed value. The preliminary regression is estimated to offer a channel to correct for the potential endogeneity in the used corruption measures. Since the primary corruption index, ICRG, is bounded by its corner values, zero and six, and our sample includes each of these extreme values, the preliminary regression for the estimation utilises the Two-bound Tobit Regression Model (Tobin, 1958). The Tobit methodology is used when there is a limited dependent variable that is 'censored'[26] between the left side, right side or both boundaries, (Maddala, 1983; Wooldridge, 2010). Applying this to our case, the underlying model is

$$Y_i^* = x_{ij}\alpha_j + u_i, \;\; u|x \sim Normal(0,\sigma^2), \qquad (8)$$

where Y_i^* is our continuous latent dependent corruption variable, x_{ijt} the set of independent variables defined in section 4.5 and showcased in Equation (9), α_j the vector matrix of coefficients to be estimated, and lastly u_i is the vector containing normally distributed error terms with the variance σ^2. As with our primary corruption index, the observed values of our alternative corruption indices are bounded by corner values of zero and ten, and zero and five respectively. However, since the sample does not take any of these extreme values for the TI and WB indices, Ordinary Least Squares can be used rather than a Tobit methodology. The results of the preliminary regression estimations are presented in the result section with an explanation of its interpretation.[27] Thus, entering our determinants of corruption into the specification used in our Two-bound Tobit and OLS models, we receive:

[26] For corner solution outcomes, it makes more sense to call the resulting model a corner solution model. Unfortunately, the name "censored regression model" appears to be firmly entrenched according to Wooldridge (2010). One could apply a Poisson regression to a dependent variable that is a Tobit-like outcome, provided that the underlying model setup of the latent dependent variable is not violated. In this case however, Tobit is deemed more appropriate.

[27] For a more detailed walkthrough of the econometric estimation process, the reader is referred to Wooldridge (2010) p. 703-705

$$COR\text{-}INDEX_{it}^* = \alpha_0 + \alpha_1 GDPPC_{it} + \alpha_2 DEMOCRACY_{it} + \alpha_3 INTCONFL_{it} + \alpha_4 PROTESTANT_i + \alpha_5 ENGLISH_i + \alpha_6 FEDERAL_i + \alpha_7 AFRICA_i + \alpha_8 ASIA_i + \alpha_9 EUROPE_i + \alpha_{10} OCEANIA_i + u_i \quad (9)$$ [28]

5.5 Model Assumptions and Limitations

To employ our choice of econometric model and yield fully robust estimates, no other assumptions are needed for the FEP regressions except for the structural replacement of the linear functional form for the mean with the exponential functional form i.e., the conditional mean assumption[29]. Therefore, the model is valid under very general conditions. The only assumption employed for valid inference is that the functional form of the FEP is correctly specified. The leading case is a FEP model with an exponential mean function as the most reliable and reasonable approximation estimator of the mean (Wooldridge, 1999; Cameron & Trivedi, 2013). The next question is then how one ensures that the functional form is indeed properly specified. The truth of the matter is that one can never be sure of its fulfilment. Given the structure and collection of our dependent variable with the chosen controls employed by previous research using the model in a similar context, it is deemed highly likely that the conditional mean is correctly specified. In addition, the appropriate pseudo measures would give a first indication of a good specification fit. Furthermore, the exponential conditional mean has been logically consistent for nonnegative variables and has the feature that coefficients can be interpreted as semi-elasticities. Should there be weak evidence for misspecification, the FEP estimator is still valid due to its robustness properties and still seen as the best estimation of the conditional mean. Even though the FEP is valid with the existence of overdispersion, a limitation could be that we cannot know if we have overdispersion after conditioning on fixed effects. To our knowledge there are no practical ways of actually testing for overdispersion in that context. However, since we do not intend to compute the probability of certain events, we are not too worried about it (Wooldridge, 2010) as even if overdispersion existed in the FEP model, the coefficient estimates would still remain consistent. Additionally with the appropriate clustered standard errors employed, an inflated type 1 error appears unlikely.

6 Empirical Results

Tables 2-5 introduces the estimated results. First, we present estimates of the performed preliminary regression predicting the main corruption index and the two other indices for sensitivity analysis, followed by the unconditional flood fatalities regression that does not include magnitude as a control. Next, the conditional flood fatalities regression that includes magnitude is introduced, before the section ends with presenting the flood magnitude regression to analyse the particular transmission channel.

[28] The Americas is the dummy variable that has been omitted and the results given for the remaining dummy variables are to be interpreted in comparison to the Americas, all else equal.
[29] For a more thorough definition of the assumption, see Appendix 3.2.

TABLE 2. Estimation of Preliminary Regression

	Dependent variable:		
	COR-ICRG	COR-TI	COR-WB
	Tobit (Censored)	*OLS*	*OLS*
log(GDPPC)	0.010	0.811^{***}	0.333^{***}
	(0.035)	(0.063)	(0.029)
DEMOCRACY	0.044^{***}	0.114^{***}	0.064^{***}
	(0.001)	(0.015)	(0.007)
INTCONFL	0.142^{***}	0.132^{***}	0.099^{***}
	(0.016)	(0.023)	(0.010)
PROTESTANT	0.028^{***}	0.021^{***}	0.012^{***}
	(0.003)	(0.005)	(0.002)
ENGLISH	−0.028	0.320^{*}	0.122^{*}
	(0.091)	(0.128)	(0.056)
FEDERAL	−0.040	-0.209^{*}	-0.071^{*}
	(0.067)	(0.102)	(0.047)
AFRICA	-0.749^{***}	0.028	0.089
	(0.136)	(0.238)	(0.107)
ASIA	−0.050	0.557^{***}	0.218^{**}
	(0.110)	(0.155)	(0.069)
EUROPE	−0.100	-0.325^{*}	-0.191^{**}
	(0.100)	(0.136)	(0.066)
OCEANIA	1.008^{***}	1.832^{***}	0.754^{***}
	(0.127)	(0.157)	(0.069)
logSigma	-0.182^{***}		
	(0.0022)		
Constant	1.037^{***}	-5.254^{***}	-5.952^{***}
	(0.333)	(0.667)	(0.298)
Observations	1 080	800	878
R^2		0.831	0.835
Adjusted R^2		0.829	0.833
Pseudo R^2	0.223		
Log Likelihood	−1 339.533		
LR Chi-Square FM	770.99^{***}		
Akaike Inf. Crit.	2 703.066		
Bayesian Inf. Crti.	2 762.883		
Residual Std. Error		0.850	0.403
F Statistic		388.812^{***}	437.381^{***}
Wald test	$1\,123.432^{***}$		

*Note: Robust Heteroeskedsic standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Sources: Authors' rendering of data from sources listed in Tables A2 and A3 in Appendix 1

6.1 Preliminary Regression

From the regressions reported in Table 2, we find robust and solid goodness of fit measures as is proven by for example the reported strong LR Chi-Square FM value, Wald-test and F-statistics indicating that the included explanatory variables as a group are highly significant in determining the level of corruption. Additionally the adjusted R^2 indicates a strong explanation of the variation in our sample overall and the pseudo R^2 indicates a good model fit[30]. The model specification yields individual estimates that are, based on previous literature and the discussion in section 4.6, of the expected signs, with exception to *FEDERAL*, and are mostly statistically significant. Specifically, *DEMOCRACY*, *INTCONFL, GDPPC* and *PROTESTANT* are all positively and significantly related to the majority of the corruption indices where each is found to reduce the level of corruption (remember that higher levels of *COR-INDEX* reflect less corruption). Moreover, *EUROPE* is negatively related to the corruption indices, suggesting a higher level of corruption in the region, relative to the Americas, all else equal, whilst both *OCEANIA* and *ASIA* show a positive relation to dependent variables, relative to *AMERICAS*. There are a few minor differences between the results of the main index, ICRG, and the other two indices. For instance, *GDPPC* and *ENGLISH* are not significant when using the ICRG index, but are significant at a 1% and 10% level respectively for the other two indices. However, overall the results are fairly similar for all three indices. The high level of significance of most of the independent variables and the overall high fit of the models as suggested by pseudo R^2 and adjusted R^2 respectively indicate a close relationship between the determinants and corruption, suggesting that the instruments are valid.

6.2 Estimates of Flood Fatalities

With the fitted values collected, we now turn to the estimation of the remaining regressions described by Equations (3) to (7), and presented in Table 3-5. We start by regressing flood fatalities on corruption and the results of these regressions are summarised in Table 3 and 4. The only difference between these regressions is that magnitude is only included in Table 4. Each model displays a pseudo R^2 value over 0.8, indicating a strong fit of the model and good evidence that the functional form could be correctly specified i.e., an indication of upholding the conditional mean assumption. Overall, the results displayed in the two tables are very similar in terms of significance levels and signs of the coefficients with the size of the coefficients differing marginally. Hence, the inclusion of magnitude as a control variable does not seem to have a major impact on the results and thereby we focus on interpreting Table 4 to avoid repetition. Additionally, given the endogenous nature of corruption as explained in section 5.4, we focus on the results using the fitted values. However, it is interesting to compare these results with the results using the observed values to see how they differ.

---

[30] The pseudo R^2 employed are the McFadden R^2, i.e., the ratio of log-likelihood. For his definition of pseudo R^2, McFadden (1979) further recommended that values ranging from 0.2 to 0.4 indicate a good model fit and values beyond 0.4 indicate an excellent model fit.

TABLE 3. Flood Fatalities (Unconditional)

	Dependent variable:					
	FATALITIES					
	ICRG		TI		WB	
	Observed	Fitted	Observed	Fitted	Observed	Fitted
Model:	(1)	(2)	(3)	(4)	(5)	(6)
COR-ICRG	0.475^{*} (0.273)					
COR-ICRG*		−0.033 (0.541)				
COR-TI			−0.456 (0.519)			
COR-TI*				−0.542 (0.742)		
COR-WB					-2.779^{***} (1.067)	
COR-WB*						−0.822 (0.871)
FREQUENCY	-0.110^{***} (0.029)	-0.130^{***} (0.024)	−0.090 (0.057)	−0.089 (0.057)	-0.130^{***} (0.044)	-0.138^{***} (0.036)
log(POPULATION)	-0.393^{***} (0.137)	-0.408^{***} (0.128)	-0.440^{**} (0.194)	-0.424^{**} (0.181)	-0.340^{***} (0.130)	-0.406^{***} (0.153)
POPDENSITY	0.0030^{***} (0.0008)	0.0032^{***} (0.0008)	0.0031^{***} (0.0009)	0.0031^{***} (0.0009)	0.0025^{***} (0.0005)	0.0031^{***} (0.0008)
log(GDPPC)	4.214 (3.261)	5.522^{**} (2.756)	−4.963 (7.505)	−6.977 (8.526)	3.365 (5.417)	−1.400 (6.272)
$\log(GDPPC)^2$	−0.272 (0.221)	-0.395^{**} (0.174)	0.237 (0.449)	0.377 (0.510)	−0.278 (0.341)	0.0092 (0.377)
Fixed effects						
COUNTRY	Yes	Yes	Yes	Yes	Yes	Yes
YEAROFOBS	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations[31]	1 074	1 074	792	792	870	870
Squared Correlation	0.97545	0.97512	0.98200	0.98180	0.98378	0.97928
Pseudo R^2	0.81823	0.81531	0.85400	0.85314	0.85235	0.84431
BIC	317056.7	322132.6	236705.0	238105.9	244150.7	257420.3

*Note: Clustered (COUNTRY) standard-errors in parentheses   Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sources: Authors' rendering of data from sources listed in Tables A2 in Appendix 1

[31] As observed, the number of observations does not equal 1 080 as in Table 1. When the resulting outcome for one particular group is always equal to zero, those observations are not rather informative about the slope parameters. Hence, these are dropped by the model to facilitate a better estimation. It is not to be worried about since it is beneficial and does not cause any bias or inconsistency, it is simply a matter of estimation efficiency.

TABLE 4. Flood Fatalities (Conditional)

	Dependent variable:					
	FATALITIES					
	ICRG		TI		WB	
	Observed	Fitted	Observed	Fitted	Observed	Fitted
Model:	(1)	(2)	(3)	(4)	(5)	(6)
COR-ICRG	0.526* (0.296)					
COR-ICRG*		−0.164 (0.645)				
COR-TI			−0.5340 (0.524)			
COR-TI*				−0.882 (0.788)		
COR-WB					−3.292*** (1.103)	
COR-WB*						−1.482 (0.930)
MAGNITUDE	0.960*** (0.117)	0.943*** (0.122)	0.858*** (0.215)	0.882*** (0.211)	1.136*** (0.200)	0.993*** (0.223)
FREQUENCY	−0.096*** (0.028)	−0.119*** (0.023)	−0.100* (0.059)	−0.097 (0.059)	−0.126*** (0.048)	−0.141*** (0.045)
log(POPULATION)	−0.394*** (0.143)	−0.404*** (0.135)	−0.445** (0.212)	−0.403** (0.197)	−0.3105** (0.124)	−0.366** (0.162)
POPDENSITY	0.0032*** (0.0009)	0.0033*** (0.0009)	0.0031*** (0.0010)	0.0031*** (0.0009)	0.0024*** (0.0006)	0.0032*** (0.0008)
log(GDPPC)	3.490 (3.164)	5.244** (2.507)	−4.184 (8.021)	−6.994 (9.185)	4.623 (5.307)	−1.409 (6.833)
$\log(\text{GDPPC})^2$	−0.235 (0.215)	−0.395** (0.154)	0.156 (0.481)	0.348 (0.546)	−0.385 (0.327)	0.0199 (0.406)
Fixed effects						
COUNTRY	Yes	Yes	Yes	Yes	Yes	Yes
YEAROFOBS	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1 074	1 074	792	792	870	870
Squared Correlation	0.97611	0.97573	0.98188	0.98180	0.98489	0.97980
Pseudo R^2	0.83085	0.82747	0.85968	0.86902	0.86267	0.85212
BIC	295121.4	300994.7	227532.5	228594.3	227146.6	244536.0

*Note: Clustered (COUNTRY) standard-errors in parentheses   Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sources: Authors' rendering of data from sources listed in Tables A2 in Appendix 1

If considering our main variable of interest, corruption, the results differ quite a bit depending on which index is chosen, and whether one considers the observed values or the fitted values. Regression 1 in Table 4 shows that using observed values for the ICRG index, corruption has a significant positive effect on flood fatalities, implying that less corrupt countries suffer more flood fatalities. Although this could be considered a counter-intuitive result, regression 3 using *COR-TI* shows an insignificant result of the opposite sign and regression 5 using *COR-WB* shows a significant result of the opposite sign, suggesting that this is not a robust result as it is highly dependent on the choice of corruption index. Additionally if considering regressions 2, 4 and 6 using the fitted values rather than the observed ones, they are all insignificant regardless of which index is used. As evident, the usage of fitted values from the preliminary regression to take into account the endogenous nature of corruption has a major impact on the results. Therefore, failing to take this endogeneity into account, as well as only using one corruption index, could cause the wrong conclusions to be made. Although the coefficients for the different indices vary in size even when using fitted values, they all share a negative sign and are all statistically insignificant, suggesting that corruption has no significant impact on flood fatalities. This conclusion holds regardless of whether or not magnitude is included as a control.

Briefly examining the control variables, *MAGNITUDE, POPULATION, FREQUENCY* and *POPDENSITY* all mainly have statistically significant results, mostly with the expected signs, perhaps with the exception of *POPULATION* as it indicates that areas with higher populations suffer fewer fatalities. However, this interpretation is made holding the population density fixed, meaning that the result essentially suggests that larger provinces in terms of land area suffer fewer fatalities. Considering the four mentioned variables, only *FREQUENCY* is not significant when using the TI-index in particular, but is highly significant for the other indices. Note that *MAGNITUDE* has a significant positive effect which will be relevant below under section 6.3 where we analyse it as a potential transmission mechanism. Interestingly, *GDPPC* is only significant when using the ICRG-index where it has an inverse U-shaped effect. It could be the case that the quadratic income specification for some reason has a good fit with this index in particular while a linear specification would have been better for the other indices.

6.3 Estimates of Flood Magnitude

Because flood magnitude is a continuous dependent variable, not count data, the estimates in Table 5 are made using the standard FE OLS model with clustered standard errors. A striking finding is the poor fit of this model, with low adjusted R^2 values in the range of 0.063-0.079 for each regression in spite of all the included fixed effects dummies. This suggests that the independent variables and fixed effects explain only a small amount of the variation in *MAGNITUDE.* However, the significant F-statistics of each model suggest that the variables are jointly significant. The fitted version of the main corruption index ICRG does not indicate a significant effect of corruption on flood magnitude, which is also the case when employing the WB index.

TABLE 5. Flood Magnitude

	Dependent variable:					
	MAGNITUDE					
	ICRG		TI		WB	
	Observed	Fitted	Observed	Fitted	Observed	Fitted
Model:	(1)	(2)	(3)	(4)	(5)	(6)
COR-ICRG	−0.026 (0.029)					
COR-ICRG*		0.106 (0.073)				
COR-TI			−0.011 (0.029)			
COR-TI*				0.186^{**} (0.081)		
COR-WB					0.022 (0.137)	
COR-WB*						0.114 (0.132)
FREQUENCY	-0.018^{***} (0.004)	-0.018^{***} (0.004)	-0.015^{***} (0.004)	-0.015^{***} (0.004)	-0.015^{***} (0.003)	-0.015^{***} (0.003)
RAINFALL	0.318^{*} (0.165)	0.317^{**} (0.161)	0.309^{*} (0.164)	0.331^{**} (0.156)	0.277^{*} (0.160)	0.284^{*} (0.160)
log(POPULATION)	−0.009 (0.022)	−0.007 (0.023)	−0.004 (0.027)	−0.004 (0.026)	−0.009 (0.024)	−0.008 (0.024)
POPDENSITY	0.00001 (0.0002)	−0.00000 (0.0002)	0.00001 (0.0002)	0.00000 (0.0002)	0.00001 (0.0002)	0.00003 (0.0002)
log(GDPPC)	−0.556 (0.494)	−0.701 (0.528)	-1.645^{**} (0.660)	-1.223^{**} (0.129)	-1.006^{**} (0.414)	-0.839^{*} (0.432)
$\log(GDPPC)^2$	0.038 (0.031)	0.050 (0.034)	0.105^{***} (0.041)	0.082^{**} (0.037)	0.067^{**} (0.026)	0.057^{**} (0.027)
Observations	1 080	1 080	800	800	878	878
R^2	0.173	0.174	0.198	0.202	0.193	0.194
Adjusted R^2	0.063	0.064	0.076	0.079	0.076	0.077
F Statistic	1.811^{***}	1.825^{***}	1.922^{***}	2.041^{***}	1.841^{***}	1.864^{***}

Note: FE OLS with year dummies included in all models. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
Standard errors, shown in parentheses, are robust and clustered by country in all models.
Sources: Authors' rendering of data from sources listed in Tables A2 in Appendix 1

In contrast, the fitted version of *COR-TI* is positively significant at a 5 percent level, indicating that countries with less corruption encounter floods of higher magnitude. This finding appears to suggest that *COR-TI* has an indirect positive effect on flood fatalities since *MAGNITUDE* is positively significant in regression 4 of Table 4. However, the unconditional result of *COR-TI* in regression 4 of Table 3 is not significant and has a negative sign, suggesting that *COR-TI* has no total effect (including its effect on *MAGNITUDE*) on flood fatalities. Additionally, the negative sign would suggest a total effect of the opposite direction relative to the supposed indirect effect. The lack of a total effect for *COR-TI*, alongside the low adjusted R^2 values and the fact that no other corruption index indicates a significant effect on magnitude would suggest that corruption has no robust indirect effect on flood fatalities via the path of magnitude.

FREQUENCY has a highly significant negative effect indicating that countries with a larger number of floods suffer flood events of smaller magnitude as was expected. Furthermore, *RAINFALL* has a significant positive effect on magnitude for each index, albeit at a 10% level when using the WB index, which is an intuitive result. Curiously, *GDPPC* has a significant U-shaped effect when using the TI and WB indices, but is not significant when using the ICRG index. Comparing this with the results in Table 3 and 4, this suggests that income has a strictly indirect effect on flood fatalities via magnitude if considering the TI and WB indices, and a strictly direct effect if considering the ICRG index. Lastly, comparing the results using fitted values with the results employing the observed values, the indices of ICRG and TI obtain negative signs with WB retaining its positive sign and each index is insignificant.

In combination, the results in Tables 3-5 indicate that there is no robust evidence that a country's corruption level affects flood fatalities either directly or indirectly by having an effect on the flood's magnitude. Therefore, we conclude that there is no statistically significant result supporting that countries with lower levels of corruption suffer fewer fatalities of major floods or suffer floods of lower magnitude, than countries with higher levels of corruption.

7 Discussion

7.1 Result Discussion

7.1.1 Flood Fatalities

Unlike the finding of Anbarci et al. (2007) that corruption has a positive significant effect on earthquake fatalities, the same inference can not be drawn in the case of floods based on our sample, demonstrating differing effects of corruption on fatalities caused by different natural disasters. Previous literature stated that institutions play a meaningful role in protecting a population from death by natural disasters, with evidence showcasing that countries with stronger institutions suffer lower disaster death counts. However, given our results, we reject the notion of Kahn (2005) and Gleason et al. (2022) that one possible mechanism behind this effect is lower corruption for the case of floods. In contrast, our findings support the suggestion of White et al. (1975) that economic and governmental development likely do not unambiguously mitigate flood consequences since no such effect is found to be associated with reduced corruption. This finding reiterates one of the core conclusions of Ferreira et al. (2013), being that floods differ substantially from other natural disasters which has implications on the specific causes of its fatalities, further highlighting the importance of studying how to mitigate the consequences of each disaster type. One important note is that we have not attempted to re-investigate corruption's impact on earthquakes or previous studies focused on natural disasters in general using our employed methodology. Therefore, we cannot state whether or not previous studies have exaggerated the significance of corruption, as suggested by Ferriera et al. (2013), by not sufficiently controlling for unobserved heterogeneity and correlation of the errors due to the shortcomings of the commonly employed Negative Binomial model.

Furthermore, one has to remember that floods and flood fatalities are determined by multiple factors specific to the particular setting which goes for all types of natural disasters. This could create an issue of omitted variable bias since it is possible that an omitted variable is correlated with an included explanatory variable and create confounding results. However, given that we have employed an empirical approach with control variables heavily supported by previous literature, we find it unlikely to generate biassed results in our case or severely disrupt the conditional mean assumption by e.g., yielding a misspecification of the functional form. Moreover, we further re-estimated all regressions with additional explanatory variables used, such as a country's urban population and the resulting output did not yield any major difference in the estimated results and fit of the models.

7.1.2 Flood Magnitude

Similar to what Ferrira et al. (2013) reported for the case of governance in general, we find that corruption has no robust significant effect on flood magnitude, and no indirect effect on flood fatalities through this particular transmission channel. While a significant effect on magnitude is found when using the TI index, no total effect on flood fatalities is found. This effect could be an indicator that less corrupt countries have better flood prevention systems, likely causing a lower count of floods, yet floods of larger magnitude once they occur, similar to the findings of Ferreira et al. (2013) and Kahn (2005). An alternative explanation for the positive significant impact of the TI index in particular might be that it has fewer observations than the other two indices with a slightly altered sample. Another plausible reason is, as mentioned in the previous subsection, that the results are dependent on the choice of corruption index and their construction. A possible way to investigate this would be to construct identical samples for all three indices covering the same observations and compare their results. The observed difference of significance levels, signs and sizes of the coefficients when using different corruption indices highlight the difficulties in both measuring corruption, and finding robust inference of which is not dependent on the choice of index.

7.2 General Discussion

The vast majority of previous research has concluded that good institutions and a low level of corruption reduces casualties aggregated across different types of disasters, or solely for earthquakes. To our knowledge, only one study (Ferreira et al., 2013) focused on floods which, similar to our results, found no mitigating effects of institutional arrangements on flood fatalities, proving the need to analyse different individual types of natural disasters. Relative to the most heavily examined disaster type of earthquakes, it is not as clear how higher development and better governmental factors such as lower corruption would mitigate flood magnitude or fatalities. The true cause of such consequences is rather more likely to be that an overly optimistic approach is taken towards land usage, preventive measures and other aspects of flood management. Pinter (2005) showcases that the increased number of dams, channelisation and levees have narrowed the natural course of large rivers in the U.S. Narrower channels have a higher and more varying water flow than rivers that have not been manipulated. This mitigating approach is still being adapted in the country even after disastrous floods such as the Great Midwest flood of 1993, West Virginia floods of 2016 and the more recent nationwide floods in June-August 2022 killing 38 people. Connecting back to White et al. (1975), this type of approach may create a sort of false perception of safe settler development with growing populations in low-positioned areas that potentially overestimate the provided protection measures. As follows, the population may be exposing itself to additional risk and vulnerability, especially if individuals do not take private preventive and protection measures (Burby, 2006). Therefore, it might be the perception of the country's ability to protect its population via economic development that has an impact rather than its level of corruption.

Since corruption can be expressed via multiple different acts within different categories, a general shortcoming of research using corruption as an explanatory variable is that measuring it is highly difficult and each attempt to measure it uses a different method. This explains the observed different sizes of coefficients between the indices but more importantly, the difficulty to measure corruption raises the question of to which extent results using these indices can be trusted. While they are likely the best available estimates, one should note the possibility that the true effect of corruption may not be what our results indicate, especially if looking at individual events as the anecdotal evidence from section 2.1.3 highlighted. In addition, while most indices of corruption reflect perceived corruption at the national level, Calgaro and Lloyd (2008), points to the fact that corruption at a more local level might be most influential at yielding higher vulnerability to natural disasters. For example, this was evident in local level corruption within tsunami planning regulation and development approvals in Thailand, with Marks (2015) reporting similar findings when investigating the causes of local government decisions regarding the 2011 Thai floods. Their findings can be linked to well-known principles of disaster related outcomes being primarily a function of local process and conduct, hence might suggest that local level corruption, not national, is the most relevant factor contributing to the fatal outcome. Therefore, it appears more likely that a significant effect on both flood fatalities and flood magnitude could exist within country-specific cases at a local level, rather than on a national level.

Another factor to acknowledge and discuss deeper is the choice of sample selection as only floods with magnitudes above the third quartile are part of the used sample, with the idea being that the most severe floods are of most interest to examine. The question then becomes whether the magnitude measure is sufficient to determine which floods are indeed the most severe. The measure consists of 3 dimensions, being area affected in km^2, the duration in days and the severity class. While this seemingly covers the main characteristics determining the intensity of a flood, it may fail to capture the harshness of flash floods, which are arguably the most dangerous kinds of floods. Even though flash floods may have a short duration and affect a small land area, they can still cause vast amounts of damage and can be a major threat to human lives. An illustrative example of this is the 2011 tsunami in Japan which resulted in 10 000 flood related fatalities, but only had a magnitude of 3.98 according to the DFO data, being well below the 6.08 threshold set for observations to be included in our sample. This flood in particular had a high severity class, but affected a small area and had a duration of just one day, and yet is one of the deadliest floods recorded in the DFO database. Having this in mind, the specification of the magnitude variable may have caused some floods that were highly severe in reality to be excluded from the sample which could have an effect on the results. Therefore, we believe utilising a fuller set of observations without a limit on magnitude, and a more accurate estimate of population in affected areas would give more confidence in drawing a robust conclusion. However, our results provide a first glance of the aggregated relationship between corruption and flood fatalities.

7.3 Future Research

Given the results presented with the following undertaken discussion, we have a number of recommendations for further research. The first is to employ a similar methodology as ours but with a larger sample regardless of the flood event's magnitude. This could considerably increase the number of observations to over 4 700 with higher variation in the data employed, yielding a more robust estimate to evaluate the validity of our results. Secondly, we suggest analysing potential differences in the effect of corruption on flood fatalities between developed and developing countries since most flood fatalities in the past have occurred in less developed countries, suggesting that they are more vulnerable. The intention is to help decision makers in less developed countries by yielding more knowledge of how to mitigate excessive casualties and create more sustainable societies with infrastructure that is more resilient towards natural disasters, not restricted to flood events. Third, future studies could examine the other two mentioned potential transmission channels of frequency and population in affected areas to broaden the understanding of corruption's potential effect on flood events. Lastly, we encourage future research to investigate whether institutional arrangements in various forms, can mitigate the effects of other types of natural disasters where current research is lacking, such as windstorms and landslides.

8 Conclusion

Despite floods being the most common and widespread of all natural disasters, the relationship between a country's level of corruption and flood fatalities has so far remained unstudied. Therefore, the aim of this study has been to bridge this research gap by employing an attractive choice for modelling count data – the Fixed Effects Poisson estimator – to determine the effect of corruption on flood fatalities and investigate if an effect exists for natural disasters in general or depends on the specific disaster type. Using the Dartmouth Flood Observatory data set amongst others, we obtained a panel data sample of 1 080 major flood events in 89 countries over the period 1985-2017 that was utilised to answer these questions.

The evidence put forth by our study indicates that countries with less corruption do not suffer fewer flood fatalities following a major flood event than countries with higher corruption – directly or indirectly through the possible transmission channel of flood magnitude – when applying three different commonly used corruption indices. This is in line with the notion of previous flood studies' that development in governance, for example in the form of lower corruption as in our case, or higher rule of law, does not reduce flood fatalities. The lack of a positive significant relationship between corruption and fatalities, as was reported by previous studies in other disaster contexts, indicates that the effect of corruption depends on the specific type of natural disaster. This highlights the importance of investigating different types of natural disasters individually rather than only at an aggregate level to fully understand the determinants of the severity of the resulting aftermath for each type.

The findings and limitations of our study leave room for further research that is needed in order to receive a more solid understanding of how corruption and other vulnerabilities in a country's institutional arrangements affects the socioeconomic consequences of floods and other natural disasters. We therefore believe it to be important to evaluate our results with a larger unrestricted sample, and to complement our results by also studying the other remaining types of natural disasters. Moreover, it would be of great value to conduct an analysis for a casual relationship in specific regions prone to major flood events, and to test if there are significant differences between developed and developing countries. Given the likely increased frequency of different natural disasters following climate change and global warming, this inconclusive field is an important area for more research to be undertaken in and we look forward to following the advancement of it.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). *'The Colonial Origins of Comparative Development: An Empirical Investigation'*. The American Economic Review, 91(5), 1369–1401.

Alesina, A., & Weder, B. (2002). *'Do corrupt governments receive less foreign aid?'*. American Economic Review, 92, 1126-1137.

Alexander, D. (2017). '*Corruption and the Governance of Disaster Ris*k'. Oxford University Press, Oxford.

Allison, P. D., & Waterman, R. P. (2002). '*Fixed-Effects Negative Binomial Regression Models*'. Sociological Methodology, 32, 247–265.

Ambraseys, N., & Bilham, R. (2011). '*Corruption kills*'. Nature. 469(7329). pp. 153–155.

Ampuero, A. (Mars 15, 2023). '*Corrupción afectó ejecución de obras de reconstrucción*'., La República. https://larepublica.pe/politica/actualidad/2023/03/15/emergencia-en-peru-corrupcion-afecto-ejecucion-de-obras-de-reconstruccion-autoridad-para-la-reconstruccion-con-cambios-contraloria-ciclon-yaku-desastre-naturales-948300 (accessed April 5, 2023)

Anbarci, N., Escaleras, M., & Register, C. (2005). ''*Earthquake Fatalities: The Interaction of Nature and Political Economy.*' Journal of Public Economics 89:1907–33

Anbarci, N., Escaleras, M., & Register, C. (2006). '*Traffic fatalities and public sector corruption*'. Kyklos, 59, 327-344.

Anbarci, N., Escaleras, M., & Register, C. (2007). '*Public Sector Corruption and Major Earthquakes: A Potentially Deadly Interaction*'. Public Choice, 132(1/2), 209–230.

Angrist, J.D., & Pischke, J. S. (2009). *'Mostly Harmless Econometrics: An Empiricist's Companion'*. Princeton and Oxford: Princeton University Press

Banerjee, A.V., & Duflo, E. (Printed Edition 2020). *'Poor Economics'*. New York: PublicAffairs, 2011

BBC (News Archive). (March 27, 2009a). '*Indonesia dam burst kills dozens*'. The British Broadcasting Corporation, BBC. http://news.bbc.co.uk/2/hi/asia-pacific/7967205.stm (accessed April 4, 2023)

BBC (News Archive). (March 28, 2009b). '*Indonesian dam burst toll rises*'. The British Broadcasting Corporation, BBC. http://news.bbc.co.uk/1/hi/world/asia-pacific/7969397.stm (accessed April 4, 2023)

Becker, G.S., & G.J. Stigler. (1974). '*Law Enforcement, Malfeasance, and Compensation of Enforcers.*' The Journal of Legal Studies 3, no. 1: 1–18.

Bierens J.J., & Brons R.K. (2006). '*Water-Related, Disasters, in Handbook on drowning: Prevention, rescue, treatment*'. Berlin: Springer Science & Business Media; 2006. pp. 535–585

Bradshaw, C.J.A., Brook, B. W., Peh, K. S. H., & Sodhi, N. S. (2007). *'Global evidence that deforestation amplifies flood risk and severity in the developing world'*, Global Change Biology 13: 2379–2395

Bronstert, A., Fritsch, U., & Niehoff, D. (2002). '*Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany*', J. Hydrol., 267, 1–2, 80–93.

Bruijnzeel, S.L.A., Calder, I.R., Chappell, N.A., Schellekens, J., Van Dijk, A.I.J.M., & Van Noordwijk, M. (2009). *'Forest–flood relation still tenuous – comment on 'Global evidence that deforestation amplifies flood risk and severity in the developing world' by C.J.A. Bradshaw, N.S. Sodi, K.S.-H. Peh, and B.W. Brook'*. Global Change Biology 15: 110–115.

Burby, R.J. (2006). '*Hurricane Katrina and the paradoxes of government disaster policy: bringing about wise governmental decisions for hazardous areas*', Annals of the American Academy of Political and Social Science 604: 171–191.

Calgaro, E., & Lloyd, K., (2008). '*Sun, sea, sand and tsunami: examining disaster vulnerability in the tourism community of Khao Lak, Thailand'*. Singapore Journal of Tropical Geography. 29(3). pp. 288–306

Cameron, A.C., & Trivedi, P. K. (1986). *'Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests'*. Journal of Applied Econometrics 1, no. 1 : 29–53.

Cameron, A.C., & Trivedi P .K. (2009). *'Microeconometrics Using Stata'* (revised edn), College Station, TX: Stata Press.

Cameron, A.C., & Trivedi P .K. (2013). '*Regression Analysis of Count Data'*, 2nd edition, Econometric Society Monograph No.53, Cambridge University Press

Cardona, O. (2003). '*Indicators for disaster risk management',* in: First expert meeting on disaster risk conceptualization and indicator modelling, 37 pp., March 2003, Manizales.

Carlin, R. E., Love, G. J., & Zechmeister, E. J. (2014). *'Natural Disaster and Democratic Legitimacy: The Public Opinion Consequences of Chile's 2010 Earthquake and Tsunami.'* Political Research Quarterly, 67(1), 3–15.

CIFOR org., & FAO org. (2005). *'Forests and floods: drowning in fiction or thriving on facts?'*. RAP Publication 2005/03 Forest Perspectives 2, CIFOR and FAO Regional Office for Asia and the Pacific Bogor, Indonesia

Clark, G. E., Dow, K., Emani, S., Jin, W., Kasperson, J. X., Kasperson, R. E., Moser, S., Ratick, S., Meyer, W. B., & Schwarz, H. E. (1998). *'Assessing the vulnerability of coastal communities to* *extreme storms: the case of Revere, MA., USA'*, Mitigation and Adaptation Strategies for Global Change 3, 1, 59–82.

Coates, L. (1999). '*Flood Fatalities in Australia, 1788–1996*'. Australian Geographer. 30(3). pp. 391–408.

Combs, D.L.,Parrish, R.G., & Quenenmoen, L.E. (1998). '*Assessing disaster attributable mortality: development and application of definition and classification matrix*', International Journal of Epidemiology 28: 1124–1129

Criss, R.E., & E.L. Shock (2001). '*Flood enhancement through flood control*', Geology 29: 875–878.

Cunado, J., & Ferreira, S. (2014). *'The Macroeconomic Impacts of Natural Disasters: The Case of Floods'*. Land Economics, 90(1), 149–168.

Dartmouth Flood Observatory (DFO), University of Colorado, USA. (2023). 'Global Active Archive of Large Flood Events' [Database]. http://floodobservatory.colorado.edu/ (accessed January 10, 2023)

Dunn, P. K., & Smyth G.K., (2018) '*Generalized Linear Models With Examples in R*'. Springer Texts in Statistics (STS), Springer

Ecological Project Office (EPO). (2023).'*About us*'. Office of the Secretary of the Government of the Federation, the Presidency (OSGF). https://ecologicalproject.gov.ng/about-us/ (accessed April 5, 2023)

Escaleras, M., & Register, C. (2012). *'Fiscal decentralization and natural hazard risks'*. Public Choice, 151(1/2), 165–183.

Escaleras, M., & Register, C. (2016). *'Public Sector Corruption and Natural Hazards'*. Public Finance Review, 44(6), 746–768.

European Commission. (2023). '*Corruption*'. https://home-affairs.ec.europa.eu/policies/internal-security/corruption_en (accessed 3 April, 2023)

European Commission's Joint Research Center. (2023). *'Knowledge Centre on Migration and Demography (KCMD) Data Portal'*. [Database]. https://migration-demography-tools.jrc.ec.europa.eu/atlas-demography (accessed 18 March, 2023)

Eurostat. (2023). [Database]. https://ec.europa.eu/eurostat/data (accessed 18 March, 2023)

Federal Bureau of Investigation. (Mars, 2023). '*Famous Cases: Hurricane Katrina Fraud*' FBI.gov https://www.fbi.gov/history/famous-cases/hurricane-katrina-fraud (accessed April 6, 2023)

Ferreira, S., & Ghimire, R. (2012). *'Forest cover, socioeconomics, and reported flood frequency in developing countries'*, Water Resources Research 48, W08529

Ferreira, S., Hamilton, K., & Vincent, J.R. (2013). *'Does Development Reduce Fatalities from Natural Disasters? New Evidence for Floods.'* Environment and Development Economics 18, no. 6: 649–79.

Fickling, A. (August 31, 2022). '*Pakistan Could Have Averted Its Climate Catastrophe*'. The Washington Post & Bloomberg Joint Analysis. https://www.washingtonpost.com/business/energy/pakistan-could-have-averted-its-climate-catastrophe/2022/08/30/4c2d3d92-28b8-11ed-a90a-fce4015dfc8f_story.html (accessed April 4, 2023)

Freeman, P.K., Keen, M., & Mani, M. (2003). *'Dealing with increased risk of natural disasters: challenges and options'*, IMF Working Paper No. WP/03/197, Washington, DC: IMF.

Glaeser, E. L., & Shleifer, A. (2002). '*Legal Origins*'. The Quarterly Journal of Economics, 117(4), 1193–1229.

Gleason, K., Loosemore, M., Patel, S.S., Ronak, P., Sanderson, D., & Sharma, A. (2022). *'Corruption and disasters in the built environment: a literature review'*. Disasters, 46(4), 928-945

Glencorse, B., & Yaseen, F. (September 17, 2022). '*PAKISTAN IS FLOODED WITH CORRUPTION*'. Diplomatic Courier. https://www.diplomaticourier.com/posts/pakistan-is-flooded-with-corruption (accessed April 4, 2023)

Global Risk Profile (GRP). (2023). '*Global Corruption Index 2022*'. Global Risk Profile - True Diligence. https://risk-indexes.com/global-corruption-index/ (accessed April 5, 2023)

Goldenberg, S. (June 15, 2006). '*Fraudsters stole $1bn of Hurricane Katrina relief cash, Congress told* '. The Guardian. https://www.theguardian.com/world/2006/jun/15/hurricanekatrina.usa (accessed 6 April 2023)

Gourieroux, C., Monfort, A., & Trognon, A. (1984). *'Pseudo Maximum Likelihood Methods: Theory'*. Econometrica, 52(3), 681–700.

Greene, W. (2004). *'The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects'*. The Econometrics Journal 7 (1):98-119.

Greene, W. (2007). *'Functional form and heterogeneity in models for count data'*. Foundations and Trends in Econometrics, 1(2), pp.113-218

Green, P. (2005). *'DISASTER BY DESIGN: Corruption, Construction and Catastrophe'*. The British Journal of Criminology 45, no. 4: 528–46.

Guimarães, P. (2008). *'The fixed effects negative binomial model revisited'*. Economics Letters, vol. 99(1), pages 63-66.

Habib, M., & Zurawicki, L. (2002). '*Corruption and Foreign Direct Investment*'. Journal of International Business Studies, 33(2), 291–307.

Hahn, J. (1997). *'A Note on the Efficient Semiparametric Estimation of Some Exponential Panel Models'*. Econometric Theory, 13(4), 583–588.

Harris, I., Jones, P., Lister, D & Osborn, T.J. (2020) '*Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset*'. Sci Data 7, 109 (2020). https://doi.org/10.1038/s41597-020-0453-3 & Climate Research Unit, University of East Anglia [Data sets available at https://crudata.uea.ac.uk/cru/data/hrg/]

Harvey, F. (September 15 , 2022). '*Pakistan floods 'made up to 50% worse by global heating'*. The Guardian.https://www.theguardian.com/environment/2022/sep/15/pakistan-floods-made-up-to-50-worse-by-global-heating (accessed April 4, 2023)

Hausman, J., Hall, B., & Griliches, Z. (1984). *'Economic models for count data with an application to the patents—R&D relationship'*, Econometrica, 52, pp. 909-938

Hellman, J.S., Geraint, J., Kaufmann, D., & Schankerman, M. (2000). '*Measuring Governance, Corruption, and State Capture: How Firms and Bureaucrats Shape the Business Environment in Transition Economies*'. Policy Research Working paper 2312, (Joint product of Governance, Regulation, and Finance, World Bank Institute, and the Chief Economist's Office, European Bank for Reconstruction and Development, April 2000)

Hilbe, J.M. (2012). *'Negative Binomial Regression'*. 2nd edn, Cambridge: Cambridge University Press

Hilbe, J.M. (2014). '*Modeling Count Data'*. Cambridge: Cambridge University Press

Huang, Z. (July 7, 2016). '*China's devastating floods can be traced back to corruption and overbuilding*'. Quartz Publication. https://qz.com/725468/chinas-devastating-floods-can-be-traced-back-to-corruption-and-overbuilding (accessed April 7, 2023)

Intergovernmental Panel on Climate Change (IPCC). (2007). '*Impacts, Adaptation and Vulnerability*', Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). (August 9, 2021a). *'Climate Change widespread, rapid, and intensifying'*. https://www.ipcc.ch/2021/08/09/ar6-wg1-20210809-pr/ (accessed March 18, 2023)

Intergovernmental Panel on Climate Change (IPCC). (2021b). *'Sixth Assessment Report Working Group 1: The Physical Science Basis'*., Chapter 11: Weather and Climate Extreme Events in a Changing Climate.

Johnston, M. (2001). *'Measuring corruption: numbers versus knowledge versus understanding'*, in A.K. Jain (ed.), The Political Economy of Corruption, London and New York: Routledge, pp. 157–179.

Jonkman, S.N. (2005). *'Global perspectives on loss of human life caused by floods'*, Natural Hazards 34: 151–175

Jonkman, S.N., & Kelman, I., (2005). '*An analysis of the causes and circumstances of flood disaster deaths*', Disasters 29: 75–97

Jonkman S.N., Vrijling J.K., & Vrouwenvelder A.C. (2008). '*Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method*'. Nat Hazards. 2008;46(3):353–89.

Kahn, M.E. (2005). '*The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions"*. The Review of Economics and Statistics, 87(2), 271–284.

Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (1999). *'Aggregating governance indicators'*. Policy Research Working Paper No. 2195, World Bank, Washington, DC.

Kellenberg, D.K., & Mobarak, A.M. (2008). *'Does rising income increase or decrease damage risk from natural disasters?'*. Journal of Urban Economics 63: 788–802.

Kelman, I. (2004). '*Philosophy of Flood Fatalities*'. FloodRiskNet Newl, Issue 1, Winter 2004, pp. 3-4.

Klitgaard, R. (1988). '*Controlling Corruption*'. Berkeley: University of California Press.

Lancaster, T. (2000). *'The incidental parameter problem since 1948'.* Journal of Econometrics 95 (2):391-413.

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

Latip, E., Ahmad Zawawi, E.M., Ismail, Z., & Mohd Nordin, R. (2018). *'Opportunities for Corruption across Flood Disaster Management (FDM)'*. IOP Conference Series: Earth and Environmental Science, 117(1), 012011

Lewis, J. (2005). *'Earthquake destruction: Corruption on the fault line'*. Transparency International, Global corruption report 2005, pp. 23-30.

Logan, T. (April 9, 2023). '*El Niño forecasting a hard task during unpredictable autumn, climate scientists say*'. ABC News Australia. https://www.abc.net.au/news/2023-04-10/el-nino-forecasting-difficult-unpredictable-autumn-climate/102189516 (accessed April 12, 2023)

Maddala, G.S. (1983). *'Limited Dependent and Qualitative Variables in Econometrics'*. New York: Cambridge University Press

Marks, D. (2015). '*The urban political ecology of the 2011 floods in Bangkok: the creation of uneven vulnerabilities*'. Pacific Affairs. 88(3). pp. 623–651.

Mauro, P. (1995). '*Corruption and growth'.* Quarterly Journal of Economics, 110, 681-712.

McFadden, D. (1979). '*Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments.*' p. 279–318 in Behavioural Travel Modelling, edited by Hensher D. A., Stopher P. R. London, UK: Croom Helm.

Mir, H. (September 13, 2022). '*Opinion Pakistan didn't contribute to climate change — but it's paying the price*'. The Washington Post. https://www.washingtonpost.com/opinions/2022/09/13/pakistan-paying-price-climate-change/ (accessed April 4, 2023)

Mukhtar, I. (April 1, 2021). '*FEATURE-Pakistan sends in armed force to stop logging in northern forests*'. Reuters, Thomson Reuters Foundation. https://www.reuters.com/article/pakistan-forests-military-idINL8N2L831A (accessed April 4, 2023)

Müller, A., Reiter, J., & Weiland, U. (2011). '*Assessment of urban vulnerability towards floods using an indicator-based approach - a case study for Santiago de Chile'.* Natural Hazards and Earth System Sciences. 11. 2107-2123. 10.5194/nhess-11-2107-2011.

Nigeria Extractive Industries Transparency Initiative (NEITI). (2019). '*SHARE OF DERIVATION & ECOLOGY 2012 - 2016*'. The Presidency, NEITI. https://neiti.gov.ng/cms/wp-content/uploads/2021/10/FASD-2012-2016-Share-of-Derivation-and-Ecology-1.pdf (accessed April 5, 2023)

National Oceanic and Atmospheric Administration (NOAA). (February, 28, 2023a). '*What Causes a Flood?*' NOAA SciJinks. https://scijinks.gov/flood/ (accessed April 12, 2023)

NOAA NAtional Severe Storms Laboratory. (2023b).'*SEVERE WEATHER 101 - Flood Basics*' National Oceanic and Atmospheric Administration (NOAA). https://www.nssl.noaa.gov/education/svrwx101/floods/ (accessed April 14, 2023)

Notre Dame Global Adaptation Initiative (ND-GAIN). (2023). '*RANKINGS*'. Notre Dame Research, University of Notre Dame. https://gain.nd.edu/our-work/country-index/rankings/ (accessed April 4, 2023)

Office of the Secretary of the Government of the Federation, the Presidency (OSGF). (2023). '*Background on Ecological Fund*'. OSGF. https://www.osgf.gov.ng/storage/app/media/uploaded-files/BACKGROUND%20ON%20ECOLOGICAL%20FUND.pdf (accessed April 5, 2023)

Ojewale, O. (November 10, 2022). '*Climate change, flooding and Nigeria's tide of corruption*'. Institute for Security Studies (ISS). https://issafrica.org/iss-today/climate-change-flooding-and-nigerias-tide-of-corruption (accessed April 4, 2023)

Olken, A.B. (2007). *'Monitoring Corruption: Evidence from a Field Experiment in Indonesia'*. Journal of Political Economy, 2007, vol. 115, no. 2.

Peru Support Group United Kingdom (PSG UK). (Mars 18, 2023). '*FLOODS EXPOSE CORRUPTION ASSOCIATED WITH PAST RECONSTRUCTION EFFORTS*'. PSG UK. https://perusupportgroup.org.uk/2023/03/floods-expose-corruption-associated-with-past-reconstruction-efforts/ (accessed April 5, 2023)

Petrucci, O. (2022). ‘*Review article: Factors leading to the occurrence of flood fatalities: a systematic review of research papers published between 2010 and 2020*’, Nat. Hazards Earth Syst. Sci., 22, 71–83

Pinter, N. (2005). ‘*One step forward, two steps back on U.S. floodplains*’, Science 308: 207–208

PolityV. (2022). ‘*Polity5 Annual Time-Series, 1946-2018*’, Integrated Network for Societal Conflict Research (INSCR). Center for Systemic Peace. https://www.systemicpeace.org/inscrdata.html (accessed March 23, 2023)

Povitkina, M., & Sjöstedt, M. (2017). *‘Vulnerability of Small Island Developing States to Natural Disasters: How Much Difference Can Effective Governments Make?’* The Journal of Environment & Development, 26(1), 82–105.

Priest, S. (2009). ‘*Building a model to estimate Risk to Life for European flood events*’. T10-07-10. Collection of Hydraulic Engineering Reports, TU Delft, Netherlands. Publisher: Middlesex University, London, UK.

Readfearn, G. (April 11, 2023). ’*Climate models warn of possible ‘super El Niño’ before end of year*’. The Guardian. https://www.theguardian.com/australia-news/2023/apr/12/climate-models-warn-of-possible-super-el-nino-before-end-of-year (accessed April 12, 2023)

Reinikka, R., & Svensson, J. (2011).*‘The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture’*, Journal of Public Economics 95 (2011) 956–966.

ReliefWeb. (March, 30, 2009). ‘*WHO emergency situation report no. 2: Collapsed dam of Situ Gintung, Tangerang, Banten Province, Republic of Indonesia - 30 Mar 2009*’ United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Services. https://reliefweb.int/report/indonesia/who-emergency-situation-report-no-2-collapsed-dam-situ-gintung-tangerang-banten (accessed April 4, 2023)

ReliefWeb. (October 12, 2022). ‘*Pakistan Monsoon Floods 2022 Islamic Relief Pakistan (12 October 2022’*). United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Services https://reliefweb.int/report/pakistan/pakistan-monsoon-floods-2022-islamic-relief-pakistan-12-october-2022 (accessed April 4, 2023)

Sauer, T. (2012). ‘*Numerical Analysis*’. Second Edition, Pearson Education, Inc.

Sawada, Y., & Takasaki., Y. (2017). *‘Natural disaster, poverty, and development: an introduction’*. World Dev., 94, pp. 2-15.

Schneiderbauer, S. (2007). ‘*Risk and vulnerability to natural disasters – from broad view to focused perspective. Theoretical background and applied methods for the identification of the most endangered populations in two case studies at different scales*’, Phd, Freie Universitaẗ Berlin.

Shaik, S., & Tunio, S. (January 29, 2014). '*Pakistan's timber mafia threaten forest protection plans - environmentalists*'. Reuters, Thomson Reuters Foundation News Archived. https://news.trust.org/item/20140128134348-w6xb1/ (accessed April 4, 2023)

Shleifer, A., & Vishny, R. W. (1993). *'Corruption'*. The Quarterly Journal of Economics, 108(3), 599–617

Serra, D. (2006). *'Empirical Determinants of Corruption: A Sensitivity Analysis'*. Public Choice, 126(1/2), 225–256.

Silva, J.S., & Tenreyro, S. (2010). *'On the existence of the maximum likelihood estimates in Poisson regression'*. Economics Letters, 107(2), pp.310-312.

Skidmore, M., & Toya, H. (2007). '*Economic Development and the Impacts of Natural Disasters.*' Economic Letters 94:20–25

Smarzynska, B. K., & Wei, S.J. (2000).'*Corruption and the Composition of Foreign Direct Investment : Firm-Level Evidence.*', Policy Research Working Paper;No. 2360. World Bank, Washington, DC. © World Bank.

Stock, J. H., & Watson, M. W. (2020). *'Introduction to Econometrics',* Fourth Edition-Global Edition. Pearson Education Limited

Tanhueco, R., & Velasquez, G. (2005). '*Know Risk'*. United Nations '*World Conference on Disaster Reduction*', Chapter: Incorporating social issues in disaster risk assessment.

The PRS Group. (2014). *'ICRG Methodology'*. https://www.prsgroup.com/wp-content/uploads/2014/08/icrgmethodology.pdf (accessed March 18, 2023)

The World Bank. (2020). '*Anti-corruption-Fact Sheet*'. https://www.worldbank.org/en/news/factsheet/2020/02/19/anticorruption-fact-sheet (accessed April 3, 2023)

The World Bank. (2021). '*Anti-corruption*'. https://www.worldbank.org/en/topic/governance/brief/anti-corruption (accessed April 3, 2023)

The World Bank. (2022). '*Pakistan: Flood Damages and Economic Losses Over USD 30 Billion and Reconstruction Needs Over USD 16 Billion - New Assessment*' . Press Release October 28, 2022., https://www.worldbank.org/en/news/press-release/2022/10/28/pakistan-flood-damages-and-economic-losses-over-usd-30-billion-and-reconstruction-needs-over-usd-16-billion-new-assessme (accesed April 4, 2023)

The World Bank. (2023a). '*Subnational Population*'. [Database]. https://databank.worldbank.org/source/subnational-population (accessed March 18, 2023)

The World Bank. (2023b). '*World Development Indicators (WDI)*'. [Database]. https://databank.worldbank.org/source/world-development-indicators# (accessed March 18, 2023)

The World Bank. (2023c). '*World Governance Indicators (WGI)*'. [Database]. https://databank.worldbank.org/source/worldwide-governance-indicators (accessed March 18, 2023)

Transparency International. (2022), *'Corruption Perception Index'.* https://www.transparency.org/en/cpi/2022 (accessed March 18, 2023)

Treisman, D. (2000). *'The Causes of Corruption: A Cross-national Study'.* Journal of Public Economics 76:339–547

Tobin J. (1958). *'Estimation of relationships for limited dependent variables'*. Econometrica 26: 24–36

Ukomadu, A. (October 20, 2022). '*Nigeria's flooding spreads to the Delta, upending lives and livelihoods*'. Reuters, Thomson Reuters Foundation. https://www.reuters.com/world/africa/nigerias-flooding-spreads-delta-upending-lives-livelihoods-2022-10-19/ (accessed April 6, 2023)

United Nations Sustainable Development Goals (UN). (2023). '*Goal 11 - Make cities and human settlements inclusive, safe, resilient and sustainable*'. United Nations Department of Economic and Social Affairs, Sustainable Development. https://sdgs.un.org/goals/goal11 (accessed March 28, 2023)

United Press International (UPI). (September 2, 2010). '*Pakistan floods and the timber mafia*'. https://www.upi.com/Business_News/Energy-Industry/2010/09/02/Pakistan-floods-and-the-timber-mafia/UPI-92421283453799/ (accessed April 4, 2023)

United States Bureau of Labor Statistics (U.S. BLS). (2023). *'Consumer Price Index'.* https://www.bls.gov/cpi/ (accessed March 18, 2023)

United States Census Bureau. (2010). *'State Area Measurements and Internal Point Coordinates'.* [Database]. https://www.census.gov/geographies/reference-files/2010/geo/state-area.html (accessed March 18, 2023)

United States Census Bureau. (2022). *'Current versus Constant (or Real) Dollars'.* https://www.census.gov/topics/income-poverty/income/guidance/current-vs-constant-dollars.html (accessed March 18, 2023)

United States Census Bureau. (2023). *'Population'*, [Database]. https://www.census.gov/topics/population.html (accessed March 18. 2023)

Vinet, F. (2017). '*Flood impacts on loss of life and human health*', in: Floods, Vol. 1: Risk Knowledge, 33–51

Wenzel, D. (2021). *'Droughts and corruption'*, Public Choice 189, 3–29

White, G.F., & W.A.R. Brinkmann, H.C. Cochrane, N.J. Ericksen (1975), '*Flood Hazard in the United States: A Research Assessment*', Boulder, CO: Institute of Behavioral Science, University of Colorado.

Woodberry, R.D. (2012). *'The Missionary Roots of Liberal Democracy'*. The American Political Science Review, 106(2), 244–274.

Wooldridge, J.M. (1999). *'Distribution-free estimation of some nonlinear panel data models*' Journal of econometrics, Vol.90 (1), p.77-97

Wooldridge, J. (2010). '*Econometrics of cross section and panel data*'. Second Edition Cambridge, MA: MIT Press.

Wooldridge, J.M. (2020). '*Introductory Econometrics: A Modern Approach*'. Seventh Edition. Cengage Learning, Inc.

World Health Organization (WHO). (2023). *'Floods'.* https://www.who.int/health-topics/floods#tab=tab_1 (accessed March 18, 2023)

Yamamura, E. (2014). *'Impact of natural disaster on public sector corruption'*, Public Choice, 161(3/4), 385–405.

Appendix

Appendix 1: Descriptive Statistics

TABLE A1. Sample Countries and Average Statistics, 1985-2017

Country	Country Code	Number of Floods	Average Flood Fatalities	Average Corruption (COR-ICRG)
Algeria	DZA	3	1	1.7
Angola	AGO	5	8	1.8
Argentina	ARG	25	10	2.7
Australia	AUS	58	5	4.7
Azerbaijan	AZE	1	0	2.0
Bangladesh	BGD	11	550	1.3
Belarus	BLR	3	2	3.3
Bolivia	BOL	17	16	2.1
Botswana	BWA	2	6	3.5
Brazil	BRA	52	40	2.9
Bulgaria	BGR	3	16	2.7
Burkina Faso	BFA	4	5	3.0
Canada	CAN	12	1	5.1
Chile	CHL	7	8	4.2
China	CHN	138	167	2.7
Colombia	COL	22	82	2.6
Congo, DR	COD	5	12	1.2
Costa Rica	CRI	3	1 347	2.3
Czech Republic	CZE	3	23	3.0
Dominican Republic	DOM	1	142	2.0
Ecuador	ECU	6	26	2.8
Egypt	EGY	1	12	2.0
El Salvador	SLV	1	2 000	2.5
Ethiopia	ETH	13	36	2.0
France	FRA	9	11	4.7
Germany	DEU	4	18	4.6
Ghana	GHA	3	101	2.2
Greece	GRC	1	3	2.5
Guatemala	GTM	1	17	2.0
Guinea	GIN	1	11	3.0
Guyana	GUY	1	76	2.0
Haiti	HTI	1	3 006	1.5
Honduras	HND	3	3 707	2.0
Hungary	HUN	1	0	3.0
India	IND	109	399	2.5
Indonesia	IDN	14	57	2.3
Iran	IRA	14	76	2.8
Iraq	IRQ	2	0	1.2
Italy	ITA	1	83	3.0
Japan	JPN	1	38	5.0
Kazakhstan	KAZ	7	8	1.7
Kenya	KEN	14	58	1.6
Madagascar	MDG	11	73	3.5
Malawi	MWI	6	68	2.1
Malaysia	MYS	6	15	2.6

Sources: Authors' rendering of flood fatalities and corruption measures from the DFO (2023) & ICRG (2017)

Cont. TABLE A1. Sample Countries and Average Statistics, 1985-2017

Country	Country Code	Number of Floods	Average Flood Fatalities	Average Corruption (COR-ICRG)
Mali	MLI	2	20	2.5
Mexico	MEX	23	74	2.4
Mongolia	MNG	1	20	2.0
Morocco	MAR	3	30	2.7
Mozambique	MOZ	14	101	2.2
Myanmar	MMR	13	7 706	1.4
Namibia	NAM	5	23	2.6
New Zealand	NZL	1	0	5.0
Nicaragua	NIC	5	31	2.8
Niger	NER	14	21	1.7
Nigeria	NGA	15	44	1,4
Norway	NOR	1	1	6.0
Oman	OMN	2	80	2.5
Pakistan	PAK	24	346	1.9
Papua New Guinea	PNG	1	0	2.0
Paraguay	PRY	9	14	1.4
Peru	PER	21	40	2.6
Philippines	PHL	12	162	2.2
Poland	POL	2	29	3.6
Romania	ROU	12	12	2.8
Russia	RUS	43	7	1.9
Saudi Arabia	SAU	4	7	2.5
Senegal	SEN	2	6	2.5
Serbia	SRB	2	2	2.0
Slovakia	SVK	1	3	2.5
Somalia	SOM	3	2	1.0
South Africa	ZAF	8	23	3.5
Spain	ESP	1	4	4.0
Sudan	SDN	19	30	1.1
Suriname	SUR	2	2	2.0
Tanzania	TZA	9	30	2.5
Thailand	THA	27	100	2.1
Tunisia	TUN	1	25	3.0
Turkey	TUR	4	14	2.4
Uganda	UGA	3	71	2.0
Ukraine	UKR	3	5	2.3
United Kingdom	GBR	7	8	4.5
United States	USA	111	31	4.3
Uruguay	URY	7	2	3.9
Venezuela	VEN	5	4 008	2.0
Vietnam	VNM	19	97	2.5
Yemen	YEM	3	45	2.3
Zambia	ZMB	6	4	3.0
Zimbabwe	ZWE	4	68	0.8

Sources: Authors' rendering of flood fatalities and corruption measures from the DFO (2023) & ICRG (2017)

TABLE A2. Definition and Source of Variables − Main Regression

Variable Name	**Definition**	**Source(s)**
FATALITIES	The number of casualties caused by a flood event	The DFO Flood Archive Database (2023)
COR-ICRG	Corruption index from ICRG (PRS Group), annual surveys with 6 indicating least corruption and 0 most corruption.	International Country Risk Guide (2017)
COR-TI	Corruption index from Transparency International, annual surveys with 10 indicating least corruption and 0 most corruption.	Transparency International (2022)
COR-WB	Corruption index from World Bank, annual surveys with 5 indicating least corruption and 0 most corruption.	World Bank, World Governance Indicators (2023c)
MAGNITUDE	The magnitude of a flood event estimated as the log(Duration x Severity x Area)	The DFO Flood Archive Database (2023)
FREQUENCY	The number of flood events occurring in a specific country in a specific year	The DFO Flood Archive Database (2023)
RAINFALL	Mean annual national precipitation of a country in a specific year	Harris et al. (2020), CRU CY Dataset
POPULATION	Population of the most affected province by a flood event	World Bank, Subnational Population Database (2023a), Eurostat (2023), European Commission's Joint Research Center and its GHS built-up grid (2023),
POPDENSITY	Population of the most affected province per square kilometre	National Bureau of Statistics and Economic measures, National Census bureaus
GDPPC	Real GDP per capita, expressed in constant (2015) U.S. dollars	World Bank, World Development Indicators (2023b)

TABLE A3. Definition and Source of Variables ─ Preliminary Regression

Variable Name	**Definition**	**Source(s)**
DEMOCRACY	Index scaled 0-10, with higher values indicating more thoroughgoing democratic and transparent institutions	PolityV Database (2022)
INTCONFL	Index scaled 0-12, with higher values indicating less risk of political violence and more stability in a specific country	International Country Risk Guide (2017)
PROTESTANT	Share of a country's total population that is Protestant as of 1980	La Porta et al. (1999)
ENGLISH	Dummy variable with value 1 for indicating that the legal origin of a specific country is English i.e., common law	La Porta et al. (1999)
FEDERAL	Dummy variable with a value of 1 when a country is a federal state and 0 otherwise for unitary states	Treisman (2000)

TABLE A4**.** Descriptive Statistics – Preliminary Regression Data

Variable	N	Mean	Median	St. Dev.	Min	Max
DEMOCRACY	1 080	5.93	8.00	3.81	0	10.0
INTCONFL	1 080	8.75	9.06	1.99	1.58	12.0
PROTESTANT (%)	1 080	9.83	1.90	15.28	0	97.8
ENGLISH	1 080	0.434	0.00	0.496	0	1
FEDERAL	1 080	0.486	0	0.50	0	1
AMERICAS	1 080	0.310	0	0.463	0	1
AFRICA	1 080	0.163	0	0.370	0	1
EUROPE	1 080	0.090	0	0.286	0	1
OCEANIA	1 080	0.056	0	0.229	0	1
ASIA	1 080	0.382	0	0.486	0	1

Note: The sample period covers 1985-2017 for 89 countries for floods of magnitude 6.08 or greater.
Sources: See Table A3 above

Appendix 2: Data & Data Adjustments

A2.1 Current to Constant

The United States Census Bureau uses the U.S. Bureau of Labor Statistics's (2023) consumer Price Index for all urban consumers Retroactive series (R-CPI-U-RS) for all items, not seasonally adjusted, for 1947 through 2021. To use the R-CPI-U-RS to inflation adjust an income estimate from 1995 dollars to 2021 dollars, one multiplies the 1995 estimate by the R-CPI-U-RS from 2021 (399,0) divided by the R-CPI-U-RS from 1995 (225,3). In this case the R-CPI-U-RS from 2015 (348,9) is used as the denominator.

TABLE A5**.** Conversion from Current to Constant

Country	Year	R-CPI-U-RS	GDP per capita (Current U.S. Dollars)	GDP per Capita (Constant 2015 U.S. Dollars)
Canada	1994	220,0	19935,3815	31615,7027
Venezuela	1995	225,3	3501,45768	5422,36388
Venezuela	1998	239,5	3885,8027	5660,77897
Venezuela	2001	260,1	4939,82948	6626,32259
Venezuela	2003	270,2	3243,3688	4188,05098
Venezuela	2009	315,2	11641,7991	12886,4966

Source: Authors' rendering of R-CPI-U-RS (U.S. BLS, 2023) and GDP per capita Current to Constant U.S. Dollars (World Bank, 2023b)

A2.2 Accuracy of Population Estimates

Examples of data used by previous studies are from the Gridded Population of the World Version 4 or Lahmeyer's Population Statistics. Even though using the population grid method could yield a more reasonable population estimate, it has its limitations with regard to extrapolation as do all numerical approximations of the true value. Because of the five-year intervals, the numerical method chosen, depending on its degree of accuracy and convergence towards the true value, will produce an approximation based on the two data points five years apart. Depending on the chosen threshold of error, the output could be misleading since one does not know for sure if the method takes account for irregular movements or what the actual population patterns are of the years within the interval e.g., 2000 and 2005 (Sauer, 2012). Furthermore, due the starting year of 1990, extrapolation has to be used to calculate an estimation beyond the range of observations which is subject to a higher level of uncertainty and risk of yielding meaningless output.

A2.3 Share of Forest Coverage

Despite common belief that forests can reduce flood events, the effect of forests remains rather contentious (CIFOR & FAO, 2005). Whilst Bradshaw et al. (2007) utilised global data on floods and showcased that forests are linked with lower flood severity and risk, Bruijnzeel et al. (2009) suggested a different proposal using the same data, namely that flood events are less prone to be reported if the events affected smaller populations. However, Bradshaw et al. (2007) – unlike Ferreira et al. (2013) – did not control for population which could have bewildered the estimations since share of forest cover tends to be more in less populated areas. Even if deforestation can have an impact, as for instance in the Pakistan case mentioned in section 2.1.3, it is deemed as entity-specific and not present at an general level. In fact, the most fitting previous empirical work (Ferreira et al., 2013) reported that a country's percentage forest cover has no effect on fatalities, magnitude and frequency of flood events when looking at 2 171 large floods in 92 countries over a 25 year-period, which is coherent with the remaining evidence (Ferreira & Ghimere, 2012) disproving forest mitigating the occurrence and consequences of floods. Since forest does appear to not have an impact on flood fatalities, and since controls of magnitude, frequency and country fixed effects should account for it, share of forest coverage was not included in our regressions.

Appendix 3: Empirical Methodology

A3.1 The Negative Binomial Model and its shortcomings

The Negative Binomial model generalises the Poisson estimation by purposely relaxing the assumption of the conditional variance equaling the conditional mean. Even though the model is constructed to manage overdispersion and situations where the dependent variable has a broad, nonnegative distribution, it is susceptible to extreme outliers. If the sample population shows existence of a great degree of skewness in the dependent variable, it could lead to concerns regarding a misleading effect of outliers. This can be handled by providing a sensitivity test, but then the question of the selection of outliers to drop arises which is a further issue to handle that depends from sample to sample. However, most related previous studies have indeed employed the Negative Binomial model over the Poisson model as it was considered the preeminent alternative of count data models. The reason being the common issue of overdispersion (Cameron & Trivedi, 1986).

Previous attempts controlled bluntly for unobserved country effects, be it by including continent dummy variables, vector of time-invariant controls for mean national elevation, mean national latitude and national coastal percentage, or using random/fixed effects estimators. Firstly, using a vector of time-invariant controls in order to account for country fixed effects can be deemed as insufficient since all unobserved heterogeneity is not accounted for. Secondly, to manually create a FE estimator by estimating a standard Negative Binomial model with dummies for every single year and country to emulate a one-way or two fixed effects model to account for unobserved heterogeneity is not immune to shortcomings.
One runs the risk of suffering the incidental parameter problem which yields more biassed and inconsistent estimates (Greene, 2004). With every additional dummy, it is more likely

that the standard numerical solution for the sought inference estimates will work insufficiently (Lancaster, 2000).

Although all previous studies have analysed cross-country, only two to our knowledge have undertaken attempts to utilise the data structure to control for time invariant unobserved country effects. One of them, Kellenberg and Mobarak (2008), employed the Fixed Effect Negative Binomial (FENB) model to analyse if rising income and economic development could decrease deaths and economic losses from natural disasters. However, the issue with their choice of model is that the conditional FENB model employed is not a true fixed effects model (Allison & Waterman, 2002; Greene, 2007). Guimarães (2008) reasserts the result of Allison and Waterman (2002) that the CMLE of the FENB model does not necessarily remove individual fixed effects, and hence is not controlling for all unobserved heterogeneity unless under a very restrictive assumption.

The conclusion of Guimarães's revisitation of the original model constructed by Hausman et al. (1984) is that the individual fixed effects must be related to the individual overdispersion parameter in a certain way i.e, there must exist a specific functional relation between the two. Hausman et al. (1984) employed a parameterization that implicitly assumed that the fixed effects would equal the logarithm of the overdispersion parameter but Guimarães (2008) showcases that there is no reason for this suggestion to be true. When the original FENB model was first introduced by Hausman et al. in 1984, it was believed that it allowed for two forms of heterogeneity. However, Wooldridge (1999) showed that, in fact, the model collapsed to depend on only one heterogeneity parameter which is the first major shortcoming. Continuing with the second, the FENB imposes a very specific overdispersion of the form *(1 + C(i))* where the mean effect is *C(i)*, with many economists, including Wooldridge and Guimarães, arguing that it is difficult to see why this relation would ever be true. Third, the FENB imposes conditional independence with serial correlation not being allowed and time constant variables do not drop out of the FENB estimation, hence explaining econometricians saying it is not a true fixed effects procedure (Hilbe, 2012; 2014). Lastly, the actual estimation of the FENB model often fails to converge, highly likely because of the odd overdispersion it requires for every unit *i* in the cross section, as explained by its specific form previously mentioned. Even though Allison and Waterman's (2002) simulation study illustrates that an unconditional Negative Binomial estimator with dummy variables could represent a fixed effects estimator, as far as it is understood, their results do not prove that there is no incidental parameters problem in the Negative Binomial model. The more favourable results they obtain in the undertaken simulations is highly likely to be specific to the particular simulation design the authors considered.

To conclude, the FENB is not known to be robust to failure of any of its assumptions and shortcomings, and employing a Negative Binomial model with dummy variables will not account for all unobserved heterogeneity, with its estimator suffering from the incidental parameter problem. Estimating the average effect of *x* on *Y* becomes more severe if one must simultaneously estimate the effects for multiple *x*'s. Therefore, we reject employing both as our main econometric model.

A3.2 FEP and its Conditional Mean Assumption

What follows is the econometric definition of the FEP model and the employed conditional mean assumption.

$$Y_{ij} \mid \eta_i, x_{ij}, \phi_i \sim Multinomial\{\eta_i, p_1(x_{ij}, \beta_{j0}),...., p_T(x_{ij}, \beta_{j0})\} \qquad (A)$$

where $$p_t\left(x_{ij}, \beta_{j0}\right) \equiv Exp\left(x_{ijt,}, \beta_{j0}\right) / \left(\sum_{r=1}^{T} Exp\left(x_{ijr,}, \beta_{j0}\right)\right). \qquad (B)$$

Because the utilised distribution is not dependent on ϕ_i , Equation (A) is also the distribution of Y_{ij} conditional on η_i and x_{ij}. Hence β_{j0} can be estimated by using standard conditional ML procedures.

Let $\{(Y_i, x_{ij}, \phi_i):i = 1,2,...\}$ be the sequence of the independent and identically distributed random variable, where Y_{ij} is the $T \times 1$ vector of count variables, x_{ij} is the $T \times K$ vector of independent variables, and ϕ_i being the unobserved scalar effect. Consider now the FEP model as

$$E(Y_{it}| x_{ij}, \phi_i) = E(Y_{it}| x_{ijt}, \phi_i) = \phi_i \mu(x_{ijt}, \beta_{j0}) = \phi_i Exp(x_{ijt}\beta_{j0}), \ t = 1,2.... T, \qquad (C)$$

where ϕ_i is the unobserved, time-constant effect with the observed independent variables, x_{ij}, may be time varying or time constant. The term β_{j0} is the unknown $P \times 1$ vector of regression coefficients, and estimated by maximising the log-likelihood associated with the multinomial density that is given for each i by

$$\sum_{i=1}^{N} L_i(\beta_j) = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} log\left[p_t\left(x_{ij}, \beta_{j0}\right)\right] \qquad (D)$$

Thus, the estimator $\widehat{\beta}_j$ maximising Equation (D) is called the Fixed Effects Poisson Estimator. Since the multinomial distribution is included in the linear exponential family (Gourieroux et al., 1984), the previous results of the QCMLE within the linear exponential family implies a certain degree of robustness of the FEP estimator. If

$$E(Y_{it} |\eta_i, x_{ij}) = p_t(x_{ij}, \beta_{j0})\eta_i, \qquad (E)$$

is fulfilled then the FEP estimator is consistent and asymptotically normal, even if misspecification of the multinomial distribution exists. Therefore, using Equation (C) (the conditional mean assumption), we assume that the expected value of $L_i(\beta_j)$ is maximised at β_{j0} which yields that FEP is generally consistent for β_{j0} provided that Equation (C) holds[32].

[32] One can also demonstrate consistency by showcasing that the FEP log-likelihood has a zero expectation when evaluated at β_{j0}.

TABLE A6. Overdispersion Test

	Data: Poisson
Z = 1.7404	P-value = 0.04089
Alternative Hypothesis:	True dispersion is greater than 1
Sample estimate:	Dispersion para. 10 966.73
True Dispersion parameter:	11 051.01

Source: Authors' rendering of regression seen in Table A8 using data from Table A2

The true Dispersion Parameter is calculated by dividing the Sum of Squared Pearson Residuals with Degrees of Freedom which yields a result of 11 051.01 >> 1.

TABLE A7. Pearson Goodness of Fit Test

	GoF Statistic	**DF**	**P.value**
Deviance	1 379 400	1 072	0.00
Pearson	11 846 896	1 072	0.00

Source: Authors' rendering of regression in Table A8 using data from Table A2

One way to detect overdispersion is by conducting a variant of Pearson goodness-of-fit test. If the residual deviance and Perason goodness-of-fit statistics are much larger than the residual degrees of freedom, then either the fitted models are insufficient or the data is overdispersed. However, if lack of fit remains even after employing the maximal possible explanatory model (and removing any potential outliers), then overdispersion is the alternative reason which is accurate in our case given our empirical approach.

TABLE A8. Ordinary Poisson Model without Fixed Effects

	Dependent variable:
	FATALITIES
COR-ICRG	-0.167^{***}
	(0.002)
FREQUENCY	-0.118^{***}
	(0.001)
MAGNITUDE	-0.239^{***}
	(0.004)
log(POPULATION)	-0.010^{***}
	(0.001)
POPDENSITY	0.001^{***}
	(0.0000)
log(GDPPC)	1.395^{***}
	(0.026)
$\log(\text{GDPPC})^2$	-0.129^{***}
	(0.002)
Constant	4.772^{***}
	(0.102)
Observations	1 080
Log Likelihood	−691 886.100
Akaike Inf. Crit.	1 383 788.000

Note: *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sources: Authors' rendering of data from sources listed in Tables A2 in Appendix 1