Effect of corruption on economic growth

Francisco Legaspe

Universidad de San Andrés

ARTICLE HISTORY
Compiled August 23, 2023

ABSTRACT
Regardless of the substantial effort made by researchers, the effect that corruption has on economic growth is still under debate. Several studies identified a negative effect, but others support the idea that corruption is not in all cases harmful and that can even boost economic growth under certain conditions. This thesis is the first piece of literature using firm-level data on bribes to examine the effect that country-level corruption has on economic growth. To do so, I replicate the methodology used by Rajan & Zingales (1993). Specifically, I ask whether industrial sectors that are more sensitive to bribes develop disproportionally slower in countries with higher levels of corruption. I find this to be true in a large sample of countries over the period 2006-2021. This result supports the idea that corruption is indeed harmful to economic growth.

KEYWORDS
Corruption, Growth, Bribes, firm-level analysis.

1. Introduction

A large literature, dating at least as back as Leff (1964) has studied the influence of a country's corruption on the level and rate of economic growth. Corruption is a cultural, political, and social phenomenon, that affects every country in the world (Nye, 1967). The number of theories analyzing the effect that corruption has over economic growth is vast, as well as the empirical research that has been done based on them. However, the effect that corruption has on economic growth is still under debate.

There is a robust literature that argues that corruption hurts economic growth. Following this perspective, the World Bank has defined corruption as "the single greatest obstacle to economic and social development" (World Bank, 2003). The hypothesis that corruption is harmful to economic development is known as "Sand the wheels". It is based on the idea that corruption distorts the allocation of resources and reduces the productivity of public expenditure, generating a decrease in the rate of economic growth ((Mauro, 1995; Hauk & Saez, 2002; Dimant & Tosato, 2018; Romer, 1994).

On the other hand, some research had argued that corruption is not in all cases harmful. Corruption might even have a positive effect on economic growth in many cases. This hypothesis, commonly known as "Grease the wheels", is based on the idea that corruption acts as "speed money", allowing firms to circumvent inefficient regulations and reduce the negative impact of malfunctioning institutions. For example, by bribing government officials, a firm can significantly fasten the bureaucratic processes

(Dreher and Gassebner, 2013; Méon and Weill, 2010). Consequently, it facilitates the creation of new firms and direct investment into more prepared businesses. Therefore, having an overall positive effect on economic growth (Khan et al., 2021; Mo, 2001).

Therefore, the debate about the relationship between corruption and economic growth has not been completely resolved. Campos, Dimova, and Saleh (2010) conducted a survey of 41 quantitative studies with 410 estimates and found that 62 percent of the estimates do not find a significant relationship between growth and corruption, 6 percent support the "Grease the wheels" hypothesis, and 32 percent the "Sand the wheels" one. However, analyzing this relationship has significant challenges that have not been entirely addressed. Therefore, despite the significant and valuable effort done by researchers on the topic, the debate about the consequences of corruption is still open (Andwig, 1991; Ugur et al., 2011). There are three main challenges faced by previous empirical literature. First, in absence of an agreed-upon theory over which to base the empirical models, the list of potentially omitted variables that corruption may proxy for is large, and which explanatory variables to include is guesswork (Alt & Lassen, 2003). Second, studies commonly use country-level corruption indexes based on the perception of respondents, which can be biased by the overall performance of the economy (Treisman, 2000). Third, double causality may be present. Corruption may affect economic growth, but economic growth may affect corruption as well (Lipset, 1960).

A promising way to make progress on causality is to focus on the details of theoretical mechanisms through which corruption affects economic growth. This thesis tries to document how that works. Theorists argued that one main channel by which corruption affects a firm's performance is through bribes. Government officials demand bribes from firms to "get things done", having a direct impact on the firm's costs and benefits scheme. If some specific firm's characteristics determine the "probability of being asked for a bribe"; country-level corruption should disproportionally harm firms (and industries) that are more likely to be requested for these informal payments (Svensson, 2003; Fisman & Svensson, 2004).

This finding can act as a "smoking gun" in the causality debate for two simple reasons. Firstly, it can correct for both country and industry-level fixed effects, internalizing significant features that cannot be distinguished in the usual cross-country empirical analysis of corruption. Secondly, it studies a specific mechanism by which corruption affects economic growth, thus providing a stronger test of causality. Corrupt government agents requesting bribes from firms may not be the unique effect of corruption over economic growth. However, it is a relevant one, and the opportunity to analyze this key mechanism in isolation is highly relevant to the literature (Svensson, 2003; Fisman & Svensson, 2004).

The contributions and validity of my study will depend on how reasonable the microeconomic assumptions are, but are less dependent on specific macroeconomic models of corruption and growth (which are scarce and not agreed-upon). The main microeconomic assumption is that there are industry-specific technological factors that affect how sensitive an industry is to corruption.

It has been stated that firms that are more dependent and have more contact with the public sector, are more likely to be asked for informal payments by government officials. The main idea is that having a higher dependence on public infrastructure, government permissions, regulations, and international trade makes a firm more prone to being asked for a bribe (Svensson, 2003). Since these characteristics are highly industry-specific, can be assumed that there are industry-specific technological factors that determine the probability of being requested for an informal payment to get

things done (Fisman & Svensson, 2004). This would imply that ceteris paribus, a firm from an industry that is more sensitive to bribes should be relatively more affected by corruption than a firm from an industry that is relatively less sensitive to bribing. From now on, this probability of being asked for a bribe will be identified as "sensitive to bribe", which will be a key variable in the analysis.

The empirical methodology used in this thesis was first established in a seminal paper by Rajan and Zingales (1996). They studied the effect of financial development on economic growth, and their approach has been followed by many others (Diallo, 2017; Aghion et al., 2014; Claessens et al, 2003; Galindo & Micco, 2004). Furthermore, the methodology has been used to study the causal effect of different areas of study, such as sovereign default (Borenstein & Panizza; 2008), banking crises (Dell'ariccia et al., 2007), and pension systems (Bijlsma et al.; 2018). However, to the best of my knowledge, this methodology has never been applied to the analysis of corruption and its relationship with economic growth. This represents a valuable opportunity since this methodology provides the chance to overcome many of the problems that previous empirical research had left unsolved. Exploiting the mentioned micro-economic assumptions allows using sector-specific fixed effects that capture within-country factors that are relevant but unobserved on available datasets.

Furthermore, this thesis is one of the few studies using firm-level data to study the effect of corruption on country-level economic growth. Previous literature analyzing the relationship between corruption and economic growth based on firm-level data has focused the analysis on specific countries (Svensson, 2003; Svensson and Fisman, 2005). However, the relationship between corruption and economic growth may be sensitive to cultural, social, and institutional characteristics that are country-specific, and therefore the external validity issue does not allow to confidently generalize the analysis to other countries.

The first step of the empirical analysis is to determine the industry's sensibility to bribes, which reflects how likely is a firm from a specific sector to be requested a bribe from government officials. This measure is created based on the Enterprise Survey, which is an ongoing project of the World Bank that aims to collect firm-level data about the business environment. According to Birhanu et al. (2015), the Enterprise Survey is the best dataset to analyze bribes at a firm level. It has already been used to study the effect of corruption on other factors, such as sales growth (Gaviria, 2002; Seker and Yang, 2012) and returns on investment (O'Toole and Tarp, 2014; Birhanu et al. 2015).

There are two other pieces of data needed to replicate the methodology proposed by Rajan and Zingales (1996). First, a country-level measure of corruption. The two more commonly used indexes to measure country-level corruption are the Corruption Perception Index, and the International Country Risk Guide. Second, an indicator of economic growth for each industry in each country is needed. For this purpose, the compounded average real growth rate of the real value-added of each firm in each country is calculated based on data from the Industrial Statistics Yearbook database, produced by the United Nations Statistical Division. The results of my analysis support the hypothesis that corruption harms economic growth. The industries that are more sensitive to corruption present relatively lower economic development in countries with higher levels of corruption. In other words, the sensitive industries are significantly more affected by country-level corruption.

The coefficient estimates remain positive and statistically significant for all three different measures of industry's sensitivity to bribes created, and for the different indexes of country-level corruption. Furthermore, two robustness checks are conducted.

First, the model is estimated only over the subsample of firms that were reliable and not arbitrary (evaluated by the official conducting the survey). This strategy was used by Birhanu et al. (2013) and helped to reduce concerns about common source bias. Second, the model is estimated over the full universe of sectors, not just the manufacturing ones as in the main regressions. Having a sample that represents the overall economy, rather than just the manufacturing sector, supports the generalization of the results at a country level. One should be careful when interpreting the results. The findings suggest that ex-ante corruption damage ex-post economic growth. Therefore, this implies that the negative relationship between corruption and economic growth identified in previous literature is, at least partly, a consequence of bureaucratic corruption (type 1). My thesis presents casual evidence of the negative relationship that corruption has over economic growth but is agnostic regarding the mechanism (it is a reduced model).

The thesis is organized as follows: Section 2 introduces theoretical and empirical literature intending to provide an overlook of the main pieces of literature on the topic, its challenges, and why this thesis is relevant to current literature. Section 3 presents the different datasets used, proving an analysis of their strengths and weaknesses. In Section 4, the empirical analysis is presented and explained. Section 5 presents the result, and Section 6 offers conclusions, discuss the implications of the results, and proposes the next step of research on the topic.

2. Theoretical and empirical review

2.1. ¿What do we understand as corruption?

One of the most challenging things when analyzing corruption is to define what it is. While it might be interpreted as an unimportant semantic issue, it determines what is measured and how it is modeled (Jain, 2001). The debate about how to define corruption in detail still prevails in the literature, but in the last decades there has been some consensus in understanding corruption as an act in which: 1) government officials have personal gains, 2) that gain came from taking advantage of his/her public position.

Based on this consensus, Jain (2001) defined corruption as "the breaking of a rule by a bureaucrat (or elected official) for private gain" (Jain, 2001). This definition allows for different types of corruption. It includes the general bribing situation, where a government official receives a monetary payment in exchange to bend a rule. However, it also covers more broad forms of corruption such as nepotism (for example, when a government official provides a government contract to a person of their preference, rather than following the rule of conducting an open procurement process). Furthermore, it encompasses situations in which government officials "steal time", by not going to work their contract hours and still collecting their paycheck. This is the definition that I will use from now on since it has been commonly used in the literature and it allows us to easily identify whether an act is corrupt or not.

An important distinction of this definition, which makes it different from other broader definitions, is that whether an act is determined as corrupt depends on whether it is "breaking the rules" or not. Therefore, it depends on the rules at a specific time and place. Consequently, the same act may be seen as corrupt in a specific place (or time) but not in a different country (or decade). For example, in the United States, political parties cannot receive contributions in connection with federal elections. However, in

other countries such as Argentina, companies can make political contributions related to elections (under some conditions). An act that is corrupt in United Stated, does not break any rule in Argentina.

Regardless of the specifications of the definition, corruption is both pervasive and significant around the world. Nye (1967) stated that it is "endemic to all governments" and Glynn et al. (1997) argued that there has been no region or country that has been immune to it. Thus, corruption can be seen as old as government itself (Seldadyo et al., 2006). Furthermore, corruption plays a key role in the development of countries. The World Bank defined it as "the greatest obstacle to economic and social development. It undermines development by distorting the rule of law and weakening the institutional foundations on which economic growth depends" (World Bank, 2003). Because of the prevalence and importance of corruption, many different areas of social science had invested a lot of effort in studying it. It has been defined as a "meeting place for researchers of different disciplines of the social science and history:" (Andvig, 1991). However, the literature has not succeeded in informing policymakers about the best ways to tackle this issue. Research has lagged in anti-corruption policies; policymakers have in many cases worked based on trial and error (Banerjee et al., 2012).

In the field of economics, the literature analyzing corruption is vast. Both from a theoretical and empirical point of view, researchers have made a huge effort in trying to explain the causes and consequences of corruption, and how to treat them. The next sections will present the most influential ideas and developments about the topic. The aim is to provide an understanding of where the current research is standing, and where it is being headed.

2.2. Theoretical overview

Regardless of the large corpus of literature on the topic, there is still not a full-fledged theory of corruption (Alt & Larsson, 2003). This lack of a 'grand' theory of corruption is a consequence of a theoretical disagreement, rather than excessive empiricism (Hopkin, 2002). Early work on corruption and its relationship with growth argued that corruption was a consequence of excessive government regulation and intervention, which incentivize individuals and firms to offer informal payments to government officials to "get things done" (Leff, 1964; Huntington, 1968; Leys, 1965). In such a setting, the idea that corruption may have a positive effect on economic growth emerges. The basic hypothesis is that bribing government officials offers a way to circumvent unnecessary and inefficient regulations. However, the opposite view was also stated. The distortionary effects that corruption has on the economy could (likely) offset the positive effect (Myrdal, 1968).

As explained by Williams (2000), the decades that followed those early works emphasized the harms of excessive regulations, and the unproductive rent-seeking activities that it induces. Corruption started to be analyzed under principal-agent models that tried to capture the nature and incentives of the phenomenon. This was the first time that economists were analyzing corruption in a formal theoretical way.

Despite the huge progress and importance of the principal-agent model, the problem of corruption could not be solved. There was not a consensus about the causes or consequences of corruption at any level. The focus then shifted to the analysis of institutions (North 1990; North 1994).

Later, Mauro (1995) shifted the focus to the macroeconomic perspective, with the first attempt to estimate the effect of corruption on economic growth empirically. This

was possible thanks to the rapid development of country-level indexes of corruption that were suitable for cross-country comparison. Since then, the focus has shifted to empirical analysis. However, the lack of a widely accepted theory to use as a capstone for the specifications is still one of the main methodological complications in empirical work.

This aspect is partly circumvented by my thesis since I exploit an empirical methodology that is dependent on microeconomic assumptions that are not under significant debate, and less dependent on macroeconomic theory assumptions. This methodology has never been used, to the best of my knowledge, in the analysis of corruption. Consequently, it presents a strong contribution to the literature.

2.2.1. Microeconomic perspective – Principal-Agent models

Becker and Stigler (1974) introduced the idea of an economic analysis of institutions based on a principal-agent model. This was the kick-start of a new approach to the economic analysis of corruption. Based on this principal-agent framework, the first microeconomics models of corruption were developed. These models commonly focus on the relationship between the bureaucrat, or "agent", and the high-level government official, or "principal". The corruption act occurs when the bureaucrat abuses his/her public position to extract rent from the high-level government official, by whom the bureaucrat is employed to deliver a public service. The opportunity to take advantage arises from the cost that the high-level officials need to incur to control and evaluate the behavior of the bureaucrats. This "monitoring cost" is what allows the agent to break the rules and gain a private profit from its public position (Ugur et al., 2011). This monitoring cost led to 2 different types of corruption. The first type of corruption (Type I) refers to "Bureaucratic Corruption". It happens in scenarios where the agent is entrusted by the principal with the task of allocating public goods or services (passports, public property, contracts, permissions, etc.). If the principal has a cost or constraint in its capacity to 13 hold the agent accountable (i.e., there exists a monitoring cost), then the bureaucrat can exploit the situation, break the rules, and extract a private gain out of this task.

On the other hand, the second type refers to "Political Corruption" (Type II). In this scenario, the bureaucrat (agent) has discretionary power to decide over the allocation of government spending or sales of public assets. If the high-level government official (principal) has a cost to control this allocation (as expected in practice), the bureaucrat could manipulate this allocation in a way that generates a private economic rent.

The principal-agent frameworks allow the analysis of not only the causes of corruption but also its consequences. However, the literature on the topic tended to stay in the theoretical sphere. These microeconomic models were not often tested empirically, and thus the debate over them was vast and inconclusive. The lack of empiricism was driven by the lack of microdata at the time that these models were developed. Thus, the shift that researchers took on the 90s, to a perspective based on macroeconomics, is not surprising (Campos, Dimova, and Saleh; 2010). This shift allowed researchers to exploit the rising macroeconomic data about corruption and test the hypothesis empirically (Mauro, 1995; Peconio, 1992; Elbahnasawy et al., 2012), but the lack of a fully-fledged macroeconomic model of corruption has ever been a key problem for this type of research.

My thesis revitalizes the principal-agent application in the corruption analysis. It exploits the now available microdata about bribes to study empirically the "Bureaucratic Corruption" (type I).

2.2.2. Macroeconomic perspective – Effect of corruption on economic growth

The macroeconomic research of corruption aimed to analyze the causes and consequences that corruption has on different key variables such as investment (Enste et al., 2019), inequality (Gyimah-Brempong, 2002), and overall economic growth (Gründler et al., 2019; Mauro, 1995; Swaleheen, 2009). My thesis specifically focuses on the effect of economic growth, so this topic is going to be discussed in more detail.

The economic literature on the topic of corruption has identified different channels by which corruption affects economic growth. However, there are diverse opinions about whether corruption is overall beneficial or harmful for the rate of economic growth of a country. Both points of view are supported both theoretically and empirically.

As mentioned, Mauro (1995) was the first in changing the focus from the microeconomic principal-agent perspective into a macroeconomics framework and shifted into the empirical analysis of corruption. His conclusions support the idea that corruption hurts economic growth, a point that has been shared by many others (Wei, 2000; Hauk & Saez, 2002). The hypothesis that corruption harms economic growth is commonly known as the "Sand the Wheels" hypothesis (referring to the hindering effect that sand under the wheels has on vehicles).

The Sand the wheels' hypothesis argues that there are two main channels by which corruption disrupts economic growth. First, a firm that operates in a corrupt environment must invest a significant part of its resources into informal payments such as bribes. Consequently, fewer resources are invested in areas such as human capital which is one of the main engines of long-term economic growth (Peccorino, 1992). Second, higher levels of corruption generate uncertainty about business rules and private and social rights. Consequently, the returns on investment are more volatile and the incentives to invest are diminished. Once again, this implies a reduction of productive investment and therefore a negative effect on economic growth (Enste et al., 2019; Gründler et al., 2019).

On the other hand, there is strong theoretical literature supporting the idea that corruption is not always detrimental to economic growth. This hypothesis, known as "Grease the wheels", argues that, under certain conditions, corruption may boost economic growth by solving government inefficiencies. There are two may mechanisms by which corruption can generate a positive effect on economic growth. First, bribes can work as "speed money" that allows firms to circumvent inefficient bureaucratic processes and delays (Leff, 1964; Hunington, 1968; Acemoglu et al, 1998). Second, corruption may support market performance because only the "good firms" have the chance to survive the extra costs that corruption generates (Lui, 1985).

In this scenario, corruption is a channel to diminish the damaging effect of excessive and inefficient regulations. It is not beneficial due to the direct positive effect that corruption has by itself, but it rather works by reducing the negative effect of poor regulations. The main idea is that the first-best scenario for a country would always be to have an optimal regulation system and zero corruption. However, in scenarios with excessive and inefficient regulations (a reality in many countries), the second-best option would be to have some corruption that allows firms to work better with these regulations. Both the "Sand the wheels" and "Grease the wheels" hypothesis have been tested empirically. In most cases, the results support the negative effect on economic growth (Campos, Dimova, and Saleh; 2010)). Nonetheless, a significant number of studies also stand for the Grease the wheel hypothesis. In specific, a lot of Asian countries show a positive effect of corruption on economic development, a

phenomenon known as the "Asian Paradox" (Vial et al., 2009; Huang, 2012).

However, the debate about the effect of corruption on economic growth is far from being solved. Mainly due to the previously mentioned difficulties on the topic such as the intrinsic secrecy nature of corruption, the lack of a fully-fledged theoretical framework, and the scarcity of reliable data. My thesis tries to shed light on this debate, by offering an analysis that exploits micro-level data. The firm-level data is provided by the Enterprise Survey project of the World Bank. It has a robust methodological structure, data from firms from all around the world, and it asks specifically about bribes and corruption. However, it has never been used for research on this specific topic.

2.3. Empirical overview

As Ugur (2014) argues, the empirical work on corruption and economic growth combines the institutional approach to economic performance institutions (North 1990; North 1994) with the empirics of growth literature of Barro (1991), Mankiw (1992), and others.

With few exceptions, there are two common features of this empirical literature. First, they exploit data of corruption perception indexes. Second, they are based on cross-country analysis (Reinikka & Svensson, 2003). However, they are far from being homogeneous. These empirical studies vary significantly in different aspects, such as the estimation strategy used, ranging from Ordinary Least Square (OLS), though two-stage and three-stage least squares (2SLS and 3SLS) to Generalized Method of Moments (GMM) and simulation methods. They also differ in the sample of countries under analysis and the sampling period (Ugur & Dasgupta., 2014).

As mentioned, there has been a sharp increase in the number of empirical studies over the last decades, mainly thanks to the appearance of better measurements of country-level corruption. These indexes offered the opportunity to, somehow accurately, compare corruption levels across country and time (Gründler et al., 2019; Svensson, 2005; Ugur & Dasgupta., 2014). The main indexes used to conduct empirical analysis have been the Corruption Perception Index (produced by Transparency International), the International Country Risk Guide (produced by The PRS Group), and the Worldwide Governance Indicators (produced by the World Bank). They are known as "perception indexes" since they are created based on the perception that businessmen, experts, and ordinary people have of a specific country's situation regarding corruption. Nonetheless, the debate about the relationship between corruption and economic growth is still far from being solved. The empirical literature is still divided and there is not a consensus about either the causes or consequences of corruption. As happens in the theoretical literature, some empirical studies support the Sand the Wheels hypothesis (Aidt, 2009; Gupta et al., 2002; Pellegrini & Gerlagh, 2004), and others support the Grease the wheel hypothesis (Dreher & Gassebner, 2011; Vial & Hanoteau, 2010; Kahn et al., 2020).

Going a step back into the analysis, not even the causes of corruption are clear. Treisman (2007) reviewed the most relevant empirical literature in political science and economics looking to explain the cross-national variations. The analysis provided quite strong evidence supporting that there is a negative correlation between corruption and freedom of the press, a high share of women in government, a history of openness to trade, and developed long-lasting liberal democracies. Also, countries that are more dependent on fuel exports, unpredicted inflation, and have intrusive business

regulations, tend to be more corrupt. However, correlation does not necessarily mean cause. The causal relationship between corruption and these phenomena is still unclear. All those factors may be cause or consequence of corruption; or even have a double causality relationship, where it causes corruption but also is affected by it.

In conclusion, the empirical literature about corruption has experienced a boost in the last decades, but there is still a lot of controversy about both the causes and consequences of corruption. Some main challenges have not been solved yet. This lack of agreement about the topic has had a significant impact on policymaking since the ones in charge of tackling corruption cannot make informed policies if they do not know what is causing the problem and how bad it is. It has even been argued that research has lagged behind the policy, and in some cases, it has been a sort of "learning by doing" (Banerjee et al., 2012).

3. Main current challenges

The battle against corruption has taken important steps forwards, but it is far from fulfilling its mission. The statement that "corruption is endemic to all governments" (Nye, 1967) is as true as always.

Currently, the main theoretical challenge is the lack of a fully-fledged model of corruption (Alt & Larsson, 2003; Seldadyo et al., 2006). Such a model could provide the empirical literature with a framework that would work as a solid baseline. At the same time, would help to unify the direction of effort.

But the lack of a fully-fledged model of corruption is not the only challenge being faced. There are other important challenges, which affect the empirical analysis more directly. To start with, corruption indexes are far from perfect. As Elmukhtar Ertimi and Ali Saeh (2013) argued, "It is almost impossible to get an objective and precise measure for corruption due to the different faces and nature of corruption". Corruption acts are conducted in secrecy, and the corrupt agents have a strong incentive to hide their tracks due to the illegal nature of corruption. Consequently, there are no objective reliable measures of corruption. Empirical papers have been using the previously introduced perception-based indexes in their studies. However, these perception indexes are far from perfect. Things such as overall economic performance, political affiliation, or many other variables may bias the indexes. Also, fear of retaliation and untrust in entities conducting the surveys may as well bias the responses. However, as far from perfect, these perception indexes are the best available datasets, and they are good enough to allow researchers to study corruption in many valuable ways.

My thesis overcomes some of the weaknesses of these perception-based county-level indexes by exploiting firm-level data. The main idea is that corruption (especially type 1 corruption) needs both an agent asking for the bribe and another one paying the bribe. Analyzing the incentives scheme, the bureaucrat requesting the bribe would rarely provide a truthful answer to related questions since it is an illicit act. However, the firms that are requested for bribes do not have many incentives to avoid answering truthfully, besides the fear of retaliation. Consequently, if the questioning is focused on the firms being asked to pay bribes and they are provided with enough anonymity and protection against possible retaliation, the result could be a somehow objective and relatively less biased measure of type 1 corruption. Over the last years, some effort has been made to take advantage of this kind of firm data to analyze corruption (Svensson, 2003; Svensson & Fissman, 2004). Nonetheless, the empirical literature exploiting firm-level data to circumvent the issues of the perception-based indexes is scarce. There are

a lot of studies analyzing the effect of corruption based on country-level measures of corruption but there is yet not even a single study that aims to analyze cross-country growth rates based on this firm-level data.

Other challenges are not broad to the analysis of corruption but specific to the study of the relationship between corruption and economic growth. The main issue underlying this relationship is that corruption may affect economic growth, but the opposite direction of causality is also likely, i.e., a country's rate of economic growth may affect corruption levels. The Lipset hypothesis may be the main example of it, arguing that as societies get wealthier their capacities of monitoring public officials increase, and therefore there is a negative impact on corruption (Lipset, 1960). This double causality is a key challenge, which is not entirely solved. This thesis aims to tackle the double causality issue by exploiting an empirical methodology that, to the best of my knowledge, has never been used in the analysis of this topic. It was created by Rajan and Zingales (1996), and it solves many of the concerns about reverse causality by exploiting industry and country-fixed effects.

3.1. Bribes and their consequences

The seminal paper of Svensson (2003) is, to the best of my knowledge, the first of one of the few papers that analyzes the macroeconomic effects of corruption based on firm-level data. This approach has three main advantages compared to the commonly used aggregate measures of corruption. First, as mentioned previously, perception indexes have raised the concern of different biases. For example, they may predict economic growth simply because the perception of the corruption level may be highly biased by the enthusiasm (or pessimism) over the performance of the overall economy of a country (Treisman, 2000). Consequently, the levels of corruption shown by these indexes may be a simple indicator, rather than a causal factor. Second, aggregate data tells little about the relationship between corruption and individual agents (i.e., there is an aggregation problem). Lastly, macro-determinants cannot, by definition, provide information about within-country variation in corruption. This is of high importance when looking to analyze the causality of corruption to economic growth since this within-country variation is expected to cause significant variations that are relevant to the analysis, but at the same time unobserved.

The main question Svensson (2003) tries to answer is: "Which firms need to pay bribes and how much?". To answer it, he proposed a theoretical model and then test it empirically based on a rich dataset that contains detailed information about informal payments of Uganda's firms.

He focused the analysis on the bureaucrat's possibility to extract bribes in different situations. The possibility to obtain bribes is determined by the potential influence that bureaucrats have over a firm's decisions and cash flows. This potentiality to affect the firm's economic status is also known as "control rights". In the relationship between corrupt bureaucrats and private firms, the control rights stem from the regulatory system and from the discretionary power that the government officials have over it. In other words, the ability of bureaucrats to implement and enforce rules that affect the firms, such as exemptions, permissions, licenses, public-good provisions, etc.

The control rights are then what distribute the negotiation power between the government official requesting the bribe and the firm. Therefore, it determines whether a firm must pay the bribe or not. If a public officer has full control rights over the firm, then the manager of the firm must either pay the required informal payment or

exit the market. On the other hand, if the firm has full control rights, the government official cannot significantly harm the business so the manager can decide to deny the required informal payment without any significant impact on their business operation. In other words, the firms that present higher control rights have less probability of being forced to pay bribes, and vice versa.

As Svensson (2003) discussed, the control right of a specific firm is determined by the required dealings with the public sector, i.e., how much the firm depends on the public sector to function.

Then, the control right of a specific firm may be determined by whether the firm engages in international trade or not, a variable indicating the type and number of taxes a firm need to pay, and a measure of the firm's dependency on public infrastructure to operate. Presumably, the firms with higher control rights (less probable to be forced to pay informal payments) are the ones that do not engage in international trade, are reached by fewer taxes, and are less dependent on public infrastructure.

In this model, the control rights level regulates if a firm pays a bribe or not. But the magnitude of the bribes (i.e., how much the firm needs to pay) is determined by other factors, which are assumed to be independent of the control right. The magnitude of the bribe is determined by two factors: the firm's refusal power and its ability to pay.

The ability to pay can be proxied by the current and future expected profits. If a firm has higher profits (present or expected), the firm's bargaining position is weaker since the public official knows that the firm can afford to pay a higher bribe. The refusal power is determined by the alternative opportunities that the firm has in case of not paying the bribe. It is then determined by the alternative return for the firm's stock of capital.

In conclusion, this highly recognized model proposed by Svensson (2003) argues that firm-specific factors determine whether a firm needs to pay a bribe or not and the magnitude of it. The firm's dependence on public infrastructure, the engagement in international trade, and the amount of taxes confronted determine the probability to be forced to make an informal payment. The magnitude of the bribe is determined by both the alternative returns of the firm's capital stock and by the present and expected profits. This hypothesis is supported by data from firms in Uganda (Svensson, 2003).

This model is highly valuable to the economic literature on corruption since it gives the first approach to understanding how firms are affected by type 1 corruption and testing it bases on data. More importantly, it reflects the firm-specific characteristics that make a firm prone to be affected by corruption. This is critical to my thesis since it shows empirically that firm-specific characteristics are relevant for determining the presence and depth of the bribes. A critical assumption of my methodology is that there are industry-specific characteristics that affect bribing. Showing that there are firm-specific factors that alter the presence and depth of bribes is the first step to it, the next one is to analyze if these relevant factors tend to be common within an industry, e.i. if they are industry-specific.

Svensson & Fisman (2005) immersed in the industry-level analysis, both theoretically and empirically. In their highly cited paper, they analyzed the relationship between the corruption suffer by a specific firm and that firm´s economic growth. The paper is based on the identifying assumption that the level of corruption that a specific firm suffers from can be decomposed into two terms: one industry-specific and the other particular to each firm. The industry-specific part is determined by the underlying characteristics of that industry and it dictates to what extent bureaucrats can extract bribes from that specific industry.

In other words, they argue that the bribes that a firm is forced to pay are, at least

in part, determined by underlying technologies which are industry-specific. As these are exogenous to the firm, it is at the same time uncorrelated with the (unobserved) firm-specific factors that affect both firm growth and bribes.

They stated that the industry's specific technologies that determined the magnitudes of the bribes being forced to pay are: how much the sector relies on imports, the share of exports of produced goods, and the dependence on public infrastructure (Svensson & Fisman, 2005). These factors were demonstrated to determine the probability that a firm must pay bribes or not (Svensson, 2003). Therefore, the demand for bribes that corrupt governments exercise over firms is determined by technological factors that are industry-specific.

This is assumption is central to my analysis since it supports the assumption that there are industry-specific observable characteristics that determine how exposed an industry is to bribing. This is the key assumption needed to properly apply Rajan and Zingales (1996) empirical methodology to the analysis of corruption and economic growth. It allows using the aggregated industry-level data to capture both between and within-country variations relevant for the analysis of corruption.

4. Empirical Strategy

The methodology used in this thesis is based on the seminal work of Rajan & Zingales (1996). The main idea of the methodology is to aggregate firm-level data into industry-level measures and then exploit between and within-country variation to strengthen the search for a causality relationship. They studied the relationship between financial development and economic growth[1]. Consequently, the specifications are less dependent on macroeconomic assumptions. This is specifically interesting when analyzing the relationship between corruption and economic growth since previous literature has shown a hard time finding a strong theory to base their assumption, and to tackle concerns about reverse causality.

The ideal would be to work with firm-level data, without aggregating at the industry level. However, the firm's data is aggregated at an industry level because the most disaggregated available data on economic growth is at an industry level. Some datasets provide firm-level growth data, but they are usually restricted to large firms and available just for a few countries. Consequently, the best available option is to work at the industry level, and the Rajan and Zingales (1996) methodology suits perfectly this criterion. My main hypothesis is that industries that are more sensitive to bribes (i.e. more likely of being requested for an informal payment) will have relatively higher growth rates in countries with relatively lower corruption levels. On the other hand, industries that are less sensitive to bribing will perform better in countries with higher levels of corruption, compared to the firms that are more sensitive to bribes. This is in line with the "Sand the wheels" hypothesis since reflects that corruption harms economic growth. This is the expected hypothesis since most empirical worldwide studies support it.

Nonetheless, there is not a predisposed relationship between variables. The empirical strategy is open to identifying both positive and negative effects of corruption on economic growth.

Then, the study of the causal effect of corruption on economic growth is done based

[1]This methodology has been replicated by many researchers, in topics such as sovereign debt (Borensztein & Panizza, 2008), bank efficiency (Diallo, 2017), fiscal policy (Aghion et al., 2014), and others (Claessens et al, 2003; Galindo & Micco, 2004, Bijlsma et al., 2018)

on the following empirical specification. The dependent variable is the compounded annual real growth rate of value-added for industry j in country c, over the period 2006-2021. The main independent variable is the interaction between industry's j on country c sensitivity to bribes and country c corruption level.

If the measures of industry j's sensitivity to bribes and country k's corruption level are appropriate, then, after correcting for industry and country-specific effects, the interaction between industry sensitivity and corruption level is expected to be positive. That would support the main hypothesis that more sensitive industries perform relatively worst in countries with high corruption levels.

This would be in line with the principal-agent theories of type 1 corruption, indicating that corrupt bureaucrats indeed harm the growth of individual firms by extracting rents from them. A direct channel of casual effect between corruption and economic growth is then tested. Is important to mention that there might be other channels by which corruption affects economic growth, which are outside the scope of this analysis (for example, type 2 corruption is not studied). Nonetheless, bureaucrat corruption (type 1) is a strong and direct channel and is highly relevant to current literature since it is highly debated.

As Rajan and Zingales (1996) stated, the most accurate way of conducting the correction over industry and country characteristics is to include indicator variables. One for each country and industry. Only explanatory variables that vary both across country and industry would need to be added to the model. There is not any a priori reason to believe that there exists such a variable, and previous literature has not pointed out any. Consequently, the empirical model I want to estimate is then:

$$Growth_{k,j} = Constant + \sigma_j + \tau_k + \beta(SensitivityIndustry_j * CorruptionCountry_k) + \mu_{k,j}$$

Where $Growth_{k,j}$ refers to the growth rate of real value-added of industry j on country k over the period under analysis; σ_j and τ_k are industry and country indicators respectively, allowing for industry and country fixed effects. Furthermore, $((SensitivityIndustry_j * CorruptionCountry_k))$ refers to the interaction between the industry's sensitivity to bribes and country-level corruption. This specification has an important advantage relative to any previous cross-country analysis of the relationship between corruption and economic growth. This advantage arises from the simple fact that this model makes predictions about within-country differences between industries based on the interaction of industry and country characteristics. Consequently, I can correct for industry and country characteristics in a way that no previous 24 literature was able to test for. The results are thus less subject to the common criticism about omitted variable bias or model specification. For this specification to work properly, it is necessary to have adequate measures of the industry's sensitivity to bribes, and of the country-level corruption. This will be addressed next.

4.1. The measure of the industry's sensitivity to bribe

There is no available data informing about the number and magnitude of bribes being paid by firms. Firms are expected to be highly reluctant to report any bribe payment since it is illegal in many cases. However, even if that piece of data was available, it would not be useful for this thesis. It would reflect the equilibrium between the demand and supply of bribes. The supply-side represents the number of informal payments that firms are willing to pay for any unofficial benefit and the equilibrium

represents the actual bribe payments that take place. However, what is relevant for my analysis is not whether a firm is trying to gain extra benefits by paying bribes. What is relevant for my analysis is the requirement of informal payments that firms receive from government officials, i.e., the demand for bribes. This demand for bribes reflects the pressure that corrupt government officials put over individual firms, irrespectively of the firm's decision to finally pay the bribe or not. This pressure reflects the theoretical channel by which bureaucrat corruption (type 1) affects firms' development and thus economic growth (Tanzi, 1998). Once aggregated at an industry level, this can be interpreted as the industry's sensitivity to bribes. It directly reflects how likely is a firm to be requested for a bribe, and thus how much pressure corrupt government officials put on that specific industry. In other words, how sensitive an industry is to corrupt government officials.

What is needed then is a way to measure the demand for bribes that the firms are experiencing, and then aggregate it at the industry level to get the industry's sensitivity to bribes. Based on the previously explained works of Svensson (2003) and Svensson and Fisman (2005), I can assume that there are specific technical reasons why some industries are more sensitive to bribes than others. Factors such as dependence on public infrastructure, import reliability, and export of products or services are assumed to be industry-specific and to affect the probability that a firm is asked to pay a bribe, i.e., the demand for bribes. This assumption is necessary to replicate Rajan and Zingales (1996) methodology since it allows us to link the within-country differences in industrial demand for bribes with differences in the growth rate of those industries.

In their analysis of the effect of financial development on economic growth, Rajan & Zingales (1996) had difficulties disentangling their demand and supply determinants of external dependency (what determines how sensible is a firm to changes in financial development and directly relates to our measure of sensitivity to bribing). The problem arises because at the time there were no systematic studies of the external financing needs of different industries, either cross-sectionally or over time. To solve this issue, they measure external dependency as the sectoral gap between investment and operating cash flow. Furthermore, they assume that the industrial-technological differences that affect external dependence persist across countries, i.e. if a specific industry is highly dependent on one country, it is also dependent on any other countries of the sample. Based on this, they calculate the external dependency on the United States and extrapolate that index to other countries.

In practice, this assumption of common underlying external funding needs across countries has not been tested due to a lack of data (Balta et al, 2013), and had received significant critics (Balta et al, 2013; Bijlsma et al., 2018; Borenzstein & Panizza, 2018). The main concern is that there are reasons to suspect that the needs could vary across countries in some industries. For example, the development of industries related to non-tradable/service production is expected to be driven by country-specific factors. Furthermore, if the sub-sector composition of a specific sector varies across countries, the demand for external funding for that sector may vary over the country (Balta et al, 2013). Consequently, some studies have modified the methodology of Rajan & Zingales (1996) to relax this assumption. This has been done by measuring an index of sensitivity for each country in the analysis, rather than extrapolating the one for the United States to all the countries in the sample (Balta et al, 2013; Bijlsma et al., 2018; Borenzstein & Panizza, 2018).

Basing my empirical strategy on the original Rajan & Zingales's methodology would imply measuring the industry's sensibility to bribe in just one country and extrap-

olating that index to industries of all other countries. In that case, I would need to implicitly assume that the demand for bribes is only affected by technological industry-specific factors, and does not significantly depend on any country-specific characteristics. However, it is too strong an assumption to make. Relevant factors such as regulations regarding foreign trade, usage of infrastructure, and tax policies are very country-specific and relevant for bribing.

Therefore, I follow the modified methodology and base the analysis on country-specific indexes of industry sensitivity to bribe methodology (Borenzstein & Panizza, 2018; Balta et a., 2013). In other words, each country has its measures of the sector's sensitivity to bribes. Consequently, I allow country-specific factors to affect this measure and do not need to rely on the strong assumption that the technological differences of industries are persistent across countries. This can be done based on the firm-level data on the demand for bribes available for many countries by the Enterprise Survey, conducted by the World Bank.

The World Bank's Enterprise Survey collects data about a firm's experiences with different aspects of the business environment. It is conducted around the world and for firms of different sizes and industries. The Enterprise Survey provides a unique opportunity for my analysis since it asks firms specifically about the demand for bribes that they faced. They are not asked about how many bribes they paid, but about how many bribes they have been requested for. Furthermore, in many cases, the firms are asked to share the amount of the requested bribes as a percentage of total annual sales. This is highly beneficial since it allows me to measure the responses relative to the business size. The effect of paying a bribe of ten thousand dollars is not the same for a large firm as for a small business.

Then, the industry's sensitivity to bribes is the aggregation of the firm's reported experiences about the percentage of total annual sales required to pay as informal payment to corrupt government officials to "Get things done" on each firm in each country.

The analysis is restricted to manufacturing sectors. As Rajan and Zingales (1996) pointed out, including non-manufacturing sectors would generate concerns about reverse causality. The argument of reverse causality would go as follows. Suppose there are some underlying country-specific factors or endowments (like natural resources) that favor certain industries (such as Mining, Tourism, or Construction) that happened to be highly suitable for corruption to occur. For example, big projects, such as the ones often taking place in mining and construction, are a good opportunity for tax evasion, embezzlement, or public infrastructure fraud. Then, countries abundant in these factors should experience a higher growth rate in industries that encourage corruption and therefore should become relatively more corrupt countries. If these factors persist, then the growth rate in sensitive industries would raise and the interaction term would be significant. This problem is solved by restricting the sample to manufacturing industries, following Rajan & Zongales (1996).

Nonetheless, the non-manufacturing sectors are later added to the analysis since they are a good complement that strengthens previous results. However, the magnitudes should be analyzed cautiously and the regression should not be interpreted in isolation. Lastly, using micro-level data on firms may raise the concern about the extent and impact of "noisy data". Measurement errors on firm-level data are usually a concern, particularly when working with bribe data, due to the secret nature of the topic. Nonetheless, the World Bank took many considerations and precautions to reduce the biases as much as possible. Furthermore, as the data is averaged at an industry level the measurement error tends to be mitigated, due to the error being

highly idiosyncratic to the firm and consequently uncorrelated with the average values (Svensson & Fisman, 2005).

To tackle potential concerns about results being driven by the noise of the data, the Enterprise Survey data is aggregated using alternative techniques. As mentioned, the basic industry's sensitivity to bribe measures is created based on the firms that answered directly in terms of percentage. Nonetheless, the model is also calculated based on two different measures of the industry's sensitivity to bribes. The first considers all the firms in the sample, not only the ones reporting the demand for bribes directly as a percentage of total annual sales. This way the number of observations is increased. The second is based on a firm's dummy indicators, reflecting whether a firm was required to pay bribes or not. This measure captures the presence of the demand for a bribe but neglects the magnitude of the bribe. Thus, the noisy data concerns are diminished, since the sensibility of the reported magnitudes is not relevant.

5. Data

5.1. Data on firms

My thesis combines firm-level, industry-level, and country-level data. To the best of my knowledge, this is the first study exploiting firm-level data to study the effects of corruption on economic growth for a worldwide sample of countries. The reason for this is probably the scarcity of reliable data. The only worldwide indexes that directly and reliably measure corruption are made at a country level; they are not based on firm-level data.

Given the secretive nature of corrupt acts, the common view has been that is not possible to collect reliable quantitative information on corruption at the firm level. However, Kauffman (1997) argued that this is an incorrect presumption, and that under appropriate survey methods and interview techniques managers would be willing to discuss corruption with remarkable candor. Furthermore, he stated that if they design an empirical strategy that provides managers with enough incentives, they would cooperate and truthfully report their experiences with corruption.

The first attempt to collect such data was made by Reinikka & Svensson (2001). They used a standard firm-level survey tool to collect quantitative data on bribe payments across firms in Uganda. Based on this dataset they were able to conduct a unique study about the causes and consequences of corruption in Uganda. Despite being a highly valuable dataset, it is just available for Uganda´s firms. This dataset was exploited for very valuable studies about corruption and its effect on economic growth (Reinikka & Collier, 2001). However, external validity is a potential issue. Corruption is highly dependent on cultural and social factors, so generalizing results from Uganda to the rest of the world cannot be done without significant concerns.

What would be needed then is a dataset that captures the firm's experiences with corruption, but for a worldwide sample. Luckily enough, there is a survey conducted by the World Bank that reports this exact piece of data. The World Bank's project "Enterprise Survey" aims to collect firm-level to inform about the business environment in countries around the world. The survey asks about the relationship between the firm and government officials. Therefore, it presents a unique opportunity to study the effects of corruption at a firm-level on a large set of countries.

5.2. Enterprise Survey

The Enterprise Survey is an ongoing project of the World Bank that aims to collect firm-level data about the business environment. Despite not being its main purpose, the Enterprise survey provides valuable information about the relationship between corrupt government officials and firms, in many countries around the world. According to Birhanu et al. (2015), the Enterprise Survey is the best dataset to analyze bribes at a firm level. It has also been used to study the effect of corruption on other factors, such as sales growth (Gaviria, 2002; Seker and Yang, 2012) and returns on investment (O'Toole and Tarp, 2014; Birhanu et al. 2015). The survey includes a representative sample of firms in the private sector, chosen from all the eligible firms listed by each country's statistics agencies, tax authorities, or business associations. The universe of firms included is then constituted by nonagricultural, formal, and private sector firms.

It follows a global approach with a standardized questionary that allows cross-country assessments. The questionary has been translated into local languages and back-translated into English to control its accuracy. Furthermore, the translation was checked by the staff of the World Bank.

Due to the sensitivity of the survey questions, addressing business-government relations and topics related to bribery, the survey is conducted by private contractors on behalf of the World Bank. They are in no case government agencies or organizations/institutions associated with the government. Furthermore, to ensure confidentiality, the information is not shared even with government agencies cooperating with the survey. The confidentiality of the survey is a key factor since it affects the degree of participation and the integrity and quality of the responses.

Measurement error of my proxy for firm-level demand for bribes is a potential concern. Ideally, I would need an objective and precise measure of the informal payments that corrupt government officials request from each firm in the survey. This represents a very challenging task since such actions rarely leave a paper trail, and managers of the firms may not be prone to share truthful information in case they paid a bribe. However, the fear of measurement error is significantly reduced by the several tactics used by the World Bank to avoid these issues. First, the anonymity of the respondents is ensured by keeping their identities secret.

Second,the questions regarding bribery appear later in the survey, once the respondent has had time to build a rapport with the data collector. Third, the question regarding bribes does not refer to blame directly to the respondent. Fourth, the survey is collected by independent data collectors, not related to the government. Lastly, the questions about a firm's experiences paying bribes are framed indirectly to reduce self-censoring bias. Specifically, they are asked: "On average, what percentage of total annual sales do firms like yours pay in informal payments to public officials to get things done?".

This indirect formulation of the question might be seen as less precise, but this is a common approach in socially sensitive questions because it reduces the self-censorship bias (Fisher, 1993; Neeley & Cronley, 2004). It has been argued that this type of questioning allows respondents to project their unconscious biases into ambiguous response situations and reveal their attitudes (Sherwood, 1981; Campbell, 1950). More importantly, this indirect questioning can also be interpreted as reflecting the concern of firms about the prevalence of bribery in the environment (Birhanu et al., 2016). This means that what respondents are measuring is the perception of the pressure that government officials put over firms with similar characteristics, i.e., the demand for bribes for firms such as theirs. This is exactly what is needed in my analysis. Not

the actual payment of bribes, but the pressure that government officials put over firms (the sensibility of the firms to bribery).

Despite being a highly well-deployed survey and adjusting appropriately to the specific needs of my empirical analysis, the Enterprise Survey is not perfect. The survey-based measures approach to corruption has its intrinsic weaknesses. Golden and Picci (2005) pointed out two main intrinsic weaknesses of this approach. First, the reliability of survey information about corruption is largely unknown. The respondents directly involved in corruption acts may have incentives to under-report due to fear of retaliation from corrupt government officials, or of being legally sanctioned. Second, the reliability of the index may deteriorate over time. The perception that managers have of corruption may be shaped by previous high publicized results of the index rather than reflecting objectively the "real" experiences with corruption.

These weaknesses should be kept in mind when interpreting the results of my analysis. However, the essence of the results holds strong even after these considerations. There is not a perfect measure of corruption, but the Enterprise Survey is certainly the best existing dataset for this job and is, without doubt, good enough. There are two rounds of surveying. The first wave took place from 2001 to 2005, but the survey was not conducted using a standardized methodology. From 2006 onwards the data is standardized and comparable across countries. This second wave is based on a different survey and the questions are not directly comparable. Therefore, this thesis exploits solely the data from the second wave and set the period of analysis from 2006 to 2021.

5.2.1. Sampling methodology, stratification, and sample size

The World Bank's Enterprise Survey collects information from manufacturing and service sectors in every region of the world. The survey uses standardized survey instruments and a uniform sampling methodology to create data that is comparable across countries around the world (World Bank, 2015).

The sampling methodology of the Enterprise Survey determines the sample sizes based on two objectives. First, to be able to compare the investment climate of different economies across the world. Second, to conduct firm performance analysis focusing on how investment climate constraints affect job creation and productivity. Industries are selected at the two-level ISIC (according to ISIC, version 3.1), depending on certain sector characteristics of each country: the number of firms from that sector, the number of employees, and their contribution to value-added. However, the selection of industries tries to keep similar industries across countries to facilitate cross-country comparability.

The sampling methodology implemented is stratified random sampling. Therefore, population units are grouped within homogeneous groups and simple random samples are selected within each group. The strata chosen on the Enterprise Survey are business sector, firm size, and geographical region within a country. For more detail on the sample size, sample methodology, and stratification, see Appendix 1.

5.2.2. Missing data

In the countries surveyed, large-scale data collection at the firm level is less than systematic. Despite the sample size conditions claimed, certain industries in some countries have a significantly low number of observations. As my empirical strategy aggregate the firms at the industry level, this can cause problems. Therefore, the data about an industry in a specific country is considered only if there are at least five firms

of that industry answering the questions.

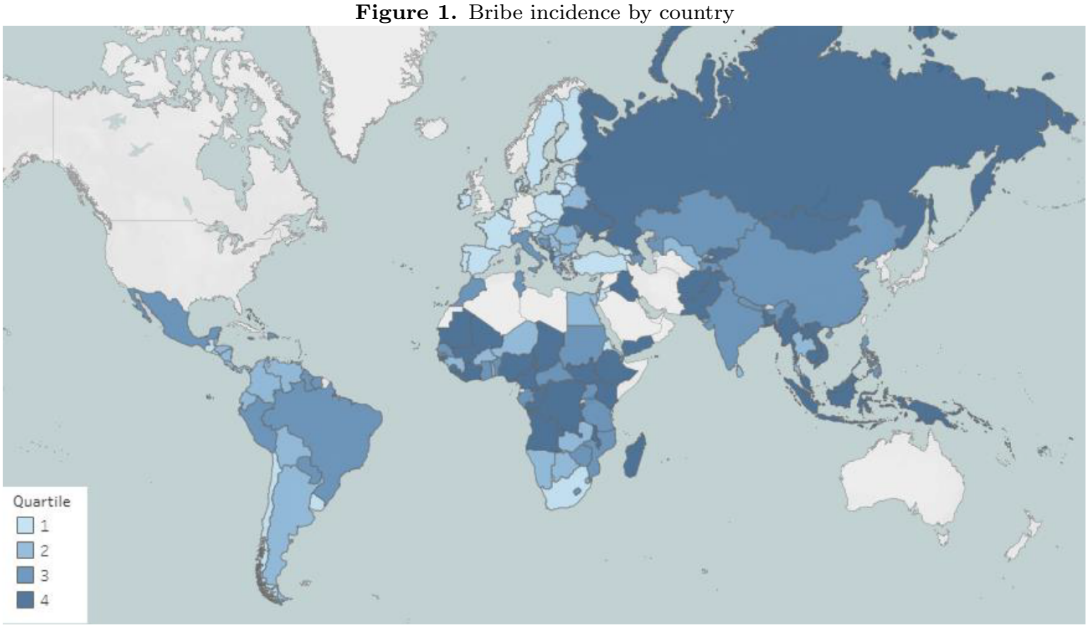
Furthermore, the methodology is based on the comparison of the same industry across different countries. Therefore, only sectors that are present in a large enough number of countries are useful for the analysis. Based on this, two-level ISIC is analyzed only if the industry was surveyed in more than five countries

5.2.3. Variation across countries

The number of bribes suffered by firms in varies widely from country to country. There are countries where being requested for a bribe from a government official is highly uncommon. For example, in Netherland and Ireland, the percentage of firms experiencing at least one bribe payment request is 0.3% and 0.4% respectively. However, some countries are experiencing the opposite scenario. In Iraq, 39.3% of firms reported being requested for bribes, and in Afghanistan, the number goes up to 46.8% (World Bank, 2022). Figure 2 shows how the percentage of firms experiencing at least one bribe request is divided by quartiles. As can be seen, bribing varies significantly both between and within regions.

In some areas, like Asia or Africa, the bribe request tends to be higher than in others, like South America and Europe. However, there are significant variations within continents. In Europe, 11.8% of Italian firms were asked for bribes, while only 0.3% of Portuguese firms suffer from that. The same happens in Africa, where 1.5% of South African firms were asked for bribes, and 28.9% of Nigerian ones (World Bank, 2022).

Figure 1. Bribe incidence by country



Note: Figure subtracted from World Bank (2022). Surveys were conducted in different years. For countries where more than one survey was conducted, the latter value is considered.

5.2.4. Calculation of the industry's sensibility to bribe – Robustness check

The aim of exploiting the Enterprise Survey firm-level data is to generate a measure of the industry's sensitivity to bribes. In other words, what I try to measure is the

pressure that corrupt government officials put over the industries when requesting informal payments on each country, i.e., the industry's demand for bribes.

As explained, the questioning regarding informal payments is done indirectly. Firms are asked: "On average, what percentage of total annual sales do firms like yours pay in informal payments to public officials to get things done?". This percentage serves as a proxy for the pressure that the firms receive from corrupt government officials. To obtain an industry-level measure, the firm-level data is then aggregated by industry in each country.

Is important to remark that having information about the bribes as a percentage of total annual sales is highly beneficial since it allows us to measure the presence and depth of the problem. However, a minor number of firms did not answer the question directly as percentages of sales, but they reported the absolute values asked as informal payment.

I addressed this situation by differentiating between two measures of the industry's sensitivity to bribes. The first one is generated based solely on firms that directly report the values of required bribes as a percentage of total sales. The second one also considers the firms that reported informal payment in absolute values. In these cases, the measure is calculated by dividing the absolute value of informal payments over the reported value of total sales. This second variable has the benefit of adding new observations to the analysis but has the downside of adding uncertainty to the calculations since the values of total annual sales and absolute levels of required bribes are answered based on the perception of the respondent.

As informal payments of bribes as a percentage of a firm's total sales are not commonly registered on record, the magnitudes reported by the firm's managers might not be sufficiently precise. Furthermore, the required informal payment by the corrupt government official might not be of a specific amount in all cases. In other words, there may not be a specific price that a firm is being asked to pay, but it is rather asked to pay something. A subjective amount that is then accepted or not by the government official. If that is true, and the reported magnitudes of bribes needed to be paid are significantly unprecise, then the results based on the percentage of sales paid as bribes may be biased (without a certain direction of the bias).

To address this, I study an additional measure of the industry's sensitivity to bribes that is based on whether a firm was asked about bribes or not, without considering the magnitude of it. In specific, each firm is given a value of 1 if it has been asked for a significant informal payment to "Get things done", and 0 otherwise. Then, the industry's sensitivity to bribes is calculated as the average of all the responses for each industry in each country. Is important to notice that results based on this measurement will limit to inform whether corruption has or not an effect on economic development but are less informative about the magnitude of those effects.

5.3. Data on industries

To determine the real economic growth of each industry in each country, I exploit the value added by industry data published in the Industrial Statistics Yearbook database, produced by the United Nations Statistical Division. It publishes the level of value-added by industry yearly, starting from 1963. Industries are individualized based on the two-level ISIC code; the same industry reference used on the Enterprise Survey.

The real compounded growth rate in real values added for the period 2006-2021 determined for each year in each country is used as a measure of the industry's real

economic growth over the period. As the Industrial Statistics yearbook's data is published in current United States dollars, the data is transformed from nominal to real values. This is done based on the World Banks' data of the United States yearly Consumer Price Index (using 2010 as the base year).

However, the data about value-added is not present for all the industries in all countries in the analysis. Since my empirical strategy compares industries within countries, not having data about the compounded average rate of value-added growth for enough firms may difficult the comparability and thus generate poor quality results. Therefore, the condition to keep observations on the analysis is that the country must have available information regarding the real compounded average rate of value-added growth for at least nine different industries. In other words, only countries with at least nine different industries are considered in the analysis.

5.4. Data on countries

Without a doubt, measuring country-level corruption in a way that is comparable across countries has been one of the main challenges in the empirical literature on the topic. As mentioned, corruption cannot be objectively observed due to its illegal and secrecy nature. To tackle this issue, during the last decades' economists and political scientists began to develop and based the research on indexes of "perceived" corruption. These perceived indexes of corruption are usually based on surveys made to residents and businessmen.

Measure such variables based on subjective measures is not free of criticism, but two main things support this approach. First, the main country-level indexes tend to be highly correlated over countries and across time. This indicates that what is being measured by the different survey methodologies is not too different and that its essence is being captured. Also, surveys responded by businessmen tend to be highly correlated to the ones made to the public.

This suggests that what is being measured is a broad country-level definition of corruption and not just a specific type or scope of it. The second support of perception indexes is that what affects economic growth is not the actual objective level of corruption, but rather the corruption perceived by the agents of the economy. When taking an economic or political decision, agents are based on their perception of the world, the objective level of some variables is often unknown and thus irrelevant for decision-making (Svensson, 2005). It is assumed that if this is the case, and individuals make decisions based on how they perceived corruption, then the subjective index might be even more valuable than the objective ones (Treisman, 2000; Mauro, 1995).

Consequently, perceived measures of corruption are commonly seen as the best measures of country-level corruption. However, empirical literature shows a significant disagreement on which is the proper corruption index to be used. The main indexes used in the empirical literature are the Corruption Perception Index by Transparency International, and the International Country Risk Guide, by the PRS Group. Both are positively and statistically significantly correlated with each other, showing that the essence of what they measure is the same (Lederman et al., 2005). Nonetheless, there are significant differences in the methodologies and scopes of each. Using both on the analysis works as a robustness check, avoiding results being due to specific characteristics of one of the indexes.

5.4.1. Corruption Perception Index

The Corruption Perception Index (CPI) is the most common measure of country-level corruption in economic literature (Swaleheen, 2011). It has been used both in cross-country as well as in panel data analysis. For example, Aidt (2009), Goel and Nelson (2010), Lessmann and Markwardt (2010), Swaleheen (2011), Méon and Weill (2010), Bjørnskov(2012), Cooray and Schneider (2018), Huang (2016), Debski et al. (2018), and many others.

The CPI was established in 1995 as a composite indicator used to measure perceptions of overall corruption in the public sector in different countries around the world. During the past 20 years, both the sources used to compile the index and its methodology of it have been adjusted and refined (Transparency International, 2020).

The methodology is divided into four steps: selection of source data, rescaling of selected data, aggregating the rescaled data and lastly reporting a measure of uncertainty of the data. To improve the quality of the dataset and avoid biases, the calculation process also incorporates quality control, which consists of parallel independent calculations made by external academic advisors and in-house researchers. Regarding the selection of data sources, the CPI is based on thirteen different data sources that capture the assessment of experts and business executives on different aspects of corrupt behavior. Some of these aspects are the diversion of public funds, bribery, nepotism, and the use of public office for personal gains. The "perception" label raises from the fact that the index is calculated based on the perception of these experts and business executives, rather than on an objective measure of corruption.

To carry out a quality evaluation, Transparency International reaches out to every one of the institutions providing the data and verifies the methodology that each of them uses to conduct their scores (Transparency International, 2020).

After obtaining and evaluating the data, each of the datasets needs to be rescaled to allow the aggregation into the CPI score. The standardization converts all the datasets to scale them from 0 to 100, where 0 represents the highest level of perceived corruption. Then, each country's CPI score is calculated based on a simple average of all the rescaled scores for that country. However, a country will be given a score only if there are at least three data sources for that country in that year.

5.4.2. Comparability across countries before 2012

In 2012 Transparency Index conducted an important update to the CPI dataset. Before then, the perceived corruption level of a specific year was calculated based on the average values of the previous seven years. Consequently, before 2012, the CPI was not comparable over time.

In the 2012 methodological note, Transparency Index stated that from there on the index would allow to compate scores over time. However, they also recognized that such comparison is not accurate with data previous to 2012

Including period fixed effect does not solve the problem with the CPI's incomparability over time, because the bias is distributed heterogeneously across continents and countries. For this reason, using longitudinal data that includes country-year observations of the CPI for periods before 2012 will produce biased results (Gründler et al., 2019).

However, the scores that countries receive are highly stable over time. Is not common for a country to have sudden or significant changes in its CPI score. Furthermore, the core factors and characteristics that determine corruption tend to be stable or change subtly. The surface changes, that can be relevant for a country's corruption level in the

short term, are not as relevant. The bureaucrats and their behavior are affected by the institutional organization and the government power, and not that much by specific events that may shift corruption perceptions temporarily (Shleifer and Vishny, 1993). Therefore, the nine years between 2012 and 2021 are used to determine the country-level corruption for the period under analysis, without much fear of the missing years causing a relevant bias on the measurement (2006 to 2021).

5.4.3. International Country Risk Guide

As the CPI is a subjective index, estimating the model using a different corruption index is a way of providing evidence that results are not driven by the subjectivity of a specific index. Therefore, all the estimations are additionally performed using the International Country Risk Guide (ICRG) index.

The ICRG model was created in 1980 by the editors of International Reports, a respected weekly newsletter on international finance and economics. In 1992, its editor and analysts moved from International Reports to The PRS Group, becoming an integral part of the company's services to the international business community.

The index is based on a set of 22 components grouped into three major categories of risk: political, financial, and economic, with political risk comprising 12 components, and financial and economic risk each comprising five components. A separate index is created for each component. Out of the political risk components, there is a specific one measuring corruption in the government. This index is used in this paper and on many other empirical pieces of literature as a measure of country-level corruption (Keefer, 1995; Braun & Di Tella ,2004; Swaleheen, 2009). Knack and Keefer (1995) argued that it can be seen as a proxy for the general efficiency with which government services are provided, and for the extent and damage of government officials' rent-seeking behavior. Its values range from zero to one, with higher values indicating lower corruption and vice versa.

5.5. Descriptive Statistics

This thesis exploits different datasets to combine information about a firm's experiences with corruption, industry-level economic growth, and country-level corruption. Table 1 provides the operational description of the main variables used in the analysis. The analysis exploits two different perceptions-based measures of corruption, three created measures of industry's sensibility to bribe, and one measure of industry economic growth.

The variable ISIC refers to the two-digit level industry reference, made by the United Nations Statistics Division. This reference is the one adopted by the World Bank on the Enterprise Survey. Consequently, it serves as a tool to combine smoothly the Enterprise Survey dataset with the Industrial Statistics Yearbook. The variable Growth refers to the industry's real compounded growth rate of real value-added values, for the period 2006-2021.

The three different measures of the industry's sensitivity to bribes are Bribe, Bribe Absolute, and Bribe Dummy. Bribe and Bribe Absolute report the request of informal payments that government officials made to firms to "Get things done". Both of them are presented as a percentage of total annual sales. The difference is that bribes only consider firms that directly reported the percentage of sales paid as bribes. Bribe Absolute additionally considers firms that did not provide that information directly but provided the bribe requests in absolute values. For these firms, the percentage of

Table 1: Description of variables

Variables	Operational definition of the variable
ISIC	Industry reference number at 2 digits label. Based on ISIC 3.1
ICRG	International Country Risk Guide by The PRS Group
CPI	Corruption Perception Index by Transparency International
growth	Industry's real compounded growth rate in real Value Added.
bribe	Percentage of sales paid by firms to government officials to "Get the things done". Calculated based only on firms that reported the value directly.
Bribe dummy	Dummy variable with value 1 if the firm paid significant bribes to government officials to "Get things done", and 0 otherwise.
Bribe absolute	Percentage of sales paid by firms to government officials to "Get the things done". Calculated based on firms that reported directly plus firms that reported paid bribes in absolute values.

sales paid as a bribe is calculated by dividing the absolute value of the bribe over the total sales of the period. Lastly, the variable Bribe Dummy indicates whether a firm was asked for an informal payment or not, without exposing the magnitude of those payments. It, therefore, takes the value 1 if a firm has been asked for an informal payment to "Get things done", and 0 otherwise.

CPI and ICRG are both perception measures of country-level corruption. Their focus is not on the effect of corrupt government officials on individual firms, but rather on the effect of the different levels of corruption (type 1 and type 2 corruption) on the overall economic development of a country.

Table 2 presents descriptive statistics for the main variables of the analysis. The values are aggregated at an industry level, so each observation reports values for an industry in a specific country. There are a total of 74 countries under analysis, and 21 different industries. However, there is no information for all the 21 industries in the 74 countries. The Enterprise Survey varies the industries under analysis based on the size of the country and country-specific characteristics. This leads to the analysis of 973 different industries around the world.

Both ICRG and CPI are constructed in a way that higher values indicate lower levels of corruption. In other words, a country with high corruption would score low in both CPI and ICRG, relative to a country that does not experience high levels of corruption, which would score high in both perception indexes. They are thus a measure of how "good" a country is doing with corruption.

The ICRG index allows values that range from 0 to 1. The lowest score is Iraq, with a score of 0.25. On the opposite corner, the highest score on the ICRG is received by Finland, with a score of 0.98. The mean value of 0.54 is close to the median of 0.50,

indicating that the values are symmetrically distributed over the sample of countries under analysis. Since the ICRG is not done for all the countries of the words, some countries such as Nepal, North Macedonia, and Kyrgyzstan are not present on the dataset. Therefore, the total number of observations is 917, out of the possible 973.

The CPI index takes values that range between 0 and 100. The lowest score is received by Yemen, with a value of 16.89. Followed closely by Iraq, with 17.78. On the contrary, the highest score was received by Denmark, with 89.44. The mean value of the CPI for the sample is 44.02, a value greater than the median of 38.89. This indicates that the distribution of the CPI is slightly positively skewed, so there is a larger number of countries with relatively low scores and a reduced sample of countries with disproportionate high scores of CPI. The last thing to notice is that the CPI is reported for all countries on the analysis, having, therefore, a total of 973 observations.

The growth variable is measured as the real compounded average growth rate in real values added for the period 2006-2021. Some sectors experienced a negative growth rate over the period, the minimum being a rate of -1%. This indicates that during the period these sectors reduced their value added by an average of 1% per year. On the other hand, some sectors presented a significant increase in their value-added over the period. The highest growth rate presents values of a real compounded average growth rate of 5.61 %. As happens with ICRG, the information of industry-level value-added is not present for all industries of all countries under analysis, leading to a total of 905 observations.

The three different measures of industry's sensibility to bribe are present for the 973 industries of the 74 countries under analysis. Bribe and Bribe absolute show the percentage of sales paid as a bribe for "Getting things done" by government officials. The industry's average requirements for informal payments reach 31.56% of the firm's annual sales. Both variables have a positively skewed distribution, indicating that there are a high number of industries being asked for bribes of relatively small magnitudes, and a reduced group of firms being asked for bribes of significantly greater value. Bribe Dummy aims to measure the percentage of firms that were asked for bribes in each industry of each country. The industry with the highest percentage of firms being asked for bribes is the "Manufacturing of other non-metallic mineral products" in Malaysia, with 27.05% of the firms being required for informal payments by government officials. The last important thing to notice is that in both categories there are industries that do not suffer from being asked for informal payments, having thus values of 0.

6. Results

Table 3 reposts the estimates of my empirical specifications. Since all the specifications control for country-specific effects and industry-specific effects, the effects that are identified are the ones that have a variation both across industry and country. Therefore, following Rajan & Zingales (1996), Table 3 reports only the interaction term between the measure of industry's sensitivity to bribes and the corresponding corruption perception index. The dependent variable is in all cases the real compounded average rate of value-added growth for the period 2006-2021. In all regressions, and throughout other regressions, the standard errors are robust to heteroscedasticity.

The first two columns of Table 3 report the estimates of the basic specification using the different measures of country-level corruption. The measure of the industry's sensitivity to bribes, "Bribe", is then based solely on firms that directly reported bribe values as a percentage of total annual sales. As can be seen in Table 3, the

Table 2: Summary Statistics

	Number of observations	Mean	Median	Standard deviation	Minimum	Maximum
ICRG	917	0.54	0.50	0.16	0.25	0.98
CPI	973	44.02	38.89	16.17	16.89	89.44
Growth	905	1.59	1.91	1.62	-1	5.61
Bribe	973	1.04	0.25	2.59	0	31.56
Bribe dummy	973	0.04	0	.08	0	0.8
Bribe absolute	973	0.91	0.23	2.28	0	27.05

coefficient estimates for the interaction term are in both cases positive and statistically significant at a 5% level. Besides its statistical importance, the impact of corruption is also quantitative important for economic growth. In specific, the regression-based on Corruption Perception Index reports that by moving from the industry with sensitivity to bribes at the lower 25th percentile (0.00) to one to the higher 75th percentile (1.03) of the distribution, the differential impact of an increase of 10 points in the CPI has over sector's value-added growth is of 1.13%. In other words, an increase of 10 points on the CPI (the country being less corrupt) have a higher impact on firms that are more sensitive to corruption. In specific, the effect on the 75th percentile more sensitive industry is 1.13% higher than the impact on the 25th percentile sensitive industry. As a country became more corrupt, the industries that are more prone to be affected by bureaucratic corruption show the lowest development compared to the industries that are not as sensitive to corrupt government officials. This way of interpreting the results is in line with the interpretations made by Rajan and Zingales (1996).

This supports the "Sand the wheels" hypothesis since it shows that industries are indeed negatively affected by corruption. In specific, industries that are more sensitive to corruption show less economic growth in countries with a higher level of corruption. This reinforces the idea that country-level corruption hurts firm-level growth, and thus over economic development (at least by the bureaucrat corruption channel).

The third and fourth columns present the results for the extended sample of firms. On these estimations, the firms that reported the informal payment requirements in absolute values instead of as a percentage of total annual sales are also included in the sample. For these sets of firms, the request for informal payments as a percentage of sales is calculated by dividing the absolute value of required informal payments over the reported value of total annual sales for the period. Therefore, the set of firms that directly reported the bribe as a percentage of sales and the ones that were calculated indirectly are considered in the variable "Bribe Absolute". The same regression specification is conducted based on this extended variable, for both perception measures of corruption.

The results of columns 3 and 4 are also positive and statistically significant. This supports the idea that industries that are sensitive to corruption grow more in less corrupt countries, compared to how they develop in corrupt countries.

Since the magnitudes reported by managers can be argued to be unprecise, columns

Table 3: Regression Results

	Country-level corruption measured by:					
Dependent variable: Growth	CPI	ICRG	CPI	ICRG	CPI	ICRG
Bribe*CPI	0.00110**	-	-	-	-	-
	(2.28)					
Bribe*ICRG	-	0.0934**	-	-	-	-
		(2.40)				
Bribe Absolute*CPI	-	-	0.00104**	-	-	-
			(2.11)			
Bribe Absolute*ICRG	-	-	-	0.0903**	-	-
				(2.26)		
Bribe Dummy*CPI	-	-	-	-	0.0335**	-
					(2.08)	
Bribe Dummy*ICRG	-	-	-	-	-	3.059**
						(2.27)
Constant	2.008***	2.034***	2.014***	2.039***	2.008***	2.031***
	(17.75)	(16.87)	(17.85)	(16.96)	(17.78)	(16.86)
Number of observations	905	849	905	849	905	849
R-Squared	0.693	0.687	0.693	0.687	0.693	0.688
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

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

Notes: The dependent variable is the annual compounded growth rate in real value-added for the period 2006-2021 for each ISIC in each country. Bribe represents the average required informal payment that firms are asked from government officials to "Get things done" on each ISIC in each country. Bribe Absolute adds firms that do not directly report as a bribe as a percentage of firm but is calculated based on absolute values reported of bribes and sales. Bribe Dummy is the percentage of firms being asked for bribes in each industry in each country. All regression includes both country-specific and industry-specific fixed effects (coefficient estimates not reported). The t-statistic is represented in parenthesis. All regressions have heteroscedastic robust standard errors.

5 and 6 of Table 3 conduct the regression by measuring bribe sensitivity based on a dummy variable. A firm's value of Bribe Dummy equals 1 if the firm was required to pay a significant amount of money to "Get things done", and 0 otherwise. Therefore, when aggregated by industry in each country, it shows the proportion of firms that were asked for informal payments on that industry in that country. Once again, the estimated coefficients are positive and statistically significant for both country-level measures of corruption.

6.1. Robustness Checks

Table 4 presents additional regressions that aim to support the results from previous regressions. There are two main extra groups of regression. The first looks to assess whether the effect is solely for the subset of manufacturing industries, or if it holds when other important sectors, such as service sectors, are included in the analysis. The second aims to evaluate whether the results are a consequence of common source bias.

Maintaining the regression specifications of Table 3, country and industry fixed effects are present in every regression and the standard errors are heteroscedastic robust. In all cases, the industry's sensitivity to bribes is calculated based on the variable "Bribe", since is the main and more reliable measure

To test whether the effect is consistent over the broader sample of sectors, all the available sectors on the Enterprise Survey are added to the analysis. In specific, the sectors added are Construction, Land Transport, Hotels and Restaurants, and Retail Trade. The coefficients estimated for this extended sample of industries are reported in columns 1 and 2 of Table 4. Once again, the coefficients estimated are positive and statistically significant at the 5% level. These results prevail independently of which corruption perception index is used.

Adding non-manufacturing sectors into the analysis has the obvious benefit of having a broader representation of a country's production matrix. However, in isolation, it confronts issues regarding reverse causality. The argument would be as follows. Suppose there are some underlying country-specific factors or endowments (like natural resources) that favor certain industries (such as Mining or Construction) that happened to be highly suitable for corruption to occur. For example, big projects, such as the ones often taking place in mining and construction, are a good opportunity for tax evasion, embezzlement, or public infrastructure fraud. Then, countries abundant in these factors should experience a higher growth rate in industries that encourage corruption and therefore should become relatively more corrupt countries. If these factors persist, then the growth rate in sensitive industries would raise and the interaction term would be significant. In conclusion, adding the non-manufacturing sector into the analysis is a good complement since it strengthens previous results, but the magnitudes should be analyzed cautiously, and the regression should not be interpreted in isolation.

The last two regressions check whether previous results are driven by common source bias. The Enterprise Survey took place on firms all around the world with the same methodology. However, the respondents do not always answer based on embellishment records. Therefore, the Enterprise Survey's interviewers rated whether the estimated answers were made with reasonable precision, or they were arbitrary and unreliable numbers.

Columns 3 and 4 of Table 4 show the regression result for the basic specification,

Table 3: Regression Results

	Country-level corruption measured by:					
Dependent variable: Growth	CPI	ICRG	CPI	ICRG	CPI	ICRG
Bribe*CPI	0.00110**	-	-	-	-	-
	(2.28)					
Bribe*ICRG	-	0.0934**	-	-	-	-
		(2.40)				
Bribe Absolute*CPI	-	-	0.00104**	-	-	-
			(2.11)			
Bribe Absolute*ICRG	-	-	-	0.0903**	-	-
				(2.26)		
Bribe Dummy*CPI	-	-	-	-	0.0335**	-
					(2.08)	
Bribe Dummy*ICRG	-	-	-	-	-	3.059**
						(2.27)
Constant	2.008***	2.034***	2.014***	2.039***	2.008***	2.031***
	(17.75)	(16.87)	(17.85)	(16.96)	(17.78)	(16.86)
Number of observations	905	849	905	849	905	849
R-Squared	0.693	0.687	0.693	0.687	0.693	0.688
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

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

Notes: The dependent variable is the annual compounded growth rate in real value-added for the period 2006-2021 for each ISIC in each country. Bribe represents the average required informal payment that firms are asked from government officials to "Get things done" on each ISIC in each country. Extra sectors refer to the addition of non-manufacturing sectors into the analysis. Common source refer to the elimination of observations that were characterized as arbitrary or unrealistic. All regression includes both country-specific and industry-specific fixed effects (coefficient estimates not reported). The t-statistic is represented in parenthesis. All regressions have heteroscedastic robust standard errors.

but solely for the subset of survey answers that are not considered to be arbitrary and unreliable. The veracity of the is evaluated by the officer responsible for conducting each survey. 23% of the observations were not answered with enough veracity and are therefore not considered in these cases. The coefficients estimated for the interaction terms are positive and statistically significant. This strengthens the previous regression results, by undermining the potential criticism of coefficients being driven by common source bias.

7. Conclusion

In this thesis, I applied the methodology of Rajan and Zingales (1996) to investigate whether corruption influences industrial growth. This methodology had never been applied to the analysis of the relationship between corruption and economic growth. In doing so, I partially circumvent some of the main problems faced by cross-country methodology which is commonly used in the empirical literature on the topic (Swaleheen, 2009).

First, it is difficult to interpret observed correlations in cross-country regressions in a casual sense. Here, I pushed the causality debate one step further by finding evidence of a specific channel by which corruption affects economic growth. This channel is known as Bureaucratic Corruption (Type I corruption), which focuses on how the request for bribes from government officials affects a firm's performance. Also, since I have multiple observations per country, the analysis adds within-country fixed effects, reducing the concerns about reverse causality.

A second problem faced by traditional methodology is that explanatory variables are multicolinear and are measured with error. The combination of these problems may generate a coefficient to be significant when it is merely a proxy for other variables measured with error. Consequently, observed correlations may be biased and misleading. This is solved by looking at the interaction effects (between country and industry indicators) rather than direct effects. By doing so, the number of variables that are needed to rely on and the range of possible alternative explanations is reduced.

The application of this methodology is possible thanks to the highly valuable dataset provided by the Worlds Banks´ Enterprise Survey. It is the only reliable firm-level data that provide direct experiences with corruption in a worldwide setting. Furthermore, the analysis exploits unique data on country-level corruption and industry-level economic growth.

Based on such strong and suitable methodological, my thesis makes valuable contributions to current research on the topic. First, the results suggest that country-level corruption has a substantially detrimental and economically significant effect on the rate of economic growth. This works, at least in part, by corrupt government officials demanding more bribes and consequently increasing the cost that firms face. The ¨Sand the Wheels¨ hypothesis is then supported[2].

My thesis is the first approach to exploring the high benefits of applying the suitable and precise methodology proposed by Rajan and Zingales (1996) for the study of the effect that country-level corruption has on economic growth. However, rather than offering a hard conclusion to the discussion, my thesis opens a window of new possibilities. The methodology and datasets used can be adapted to complement my analysis. For example, previous data from the Enterprise Survey could be used to

[2]However, this may not be the only channel by which corruption affects economic growth. "Political Corruptio (also known as type II corruption) may be another mechanism, and this is not directly analyzed in this thesis.

create a longer period of analysis and therefore a panel data set rather than a one-period analysis. Furthermore, alternative measures of corruption could be used, or the analysis could be done for countries from specific regions.

8. Bibliography

Aghion, P., Hemous, D. and Kharroubi, E., 2014. Cyclical fiscal policy, credit constraints, and industry growth. Journal of Monetary Economics, 62, pp.41-58.

Alt, J.E. and Lassen, D.D., 2003. The political economy of institutions and corruption in American states. Journal of Theoretical Politics, 15(3), pp.341-365.

Andwig, J.C., 1991. The economics of corruption: A survey. Studi economici. Arif, I., Khan, L. and Waqar, S., 2020. Does corruption sand or grease the wheels? A case of BRICS countries. Global Business Review, p.0972150920927370.

Balta, N. and Nikolov, P., 2013. Financial dependence and growth since the crisis. Quarterly Report on the Euro Area (QREA), 12(3), pp.7-18.

Balta, N. and Nikolov, P., 2013. Financial dependence and growth since the crisis. Quarterly Report on the Euro Area (QREA), 12(3), pp.7-18.

Banerjee, A., Mullainathan, S. and Hanna, R., 2012. Corruption (No. w17968). National Bureau of economic research.

Becker, G.S. and Stigler, G.J., 1974. Law enforcement, malfeasance, and compensation of enforcers. The Journal of Legal Studies, 3(1), pp.1-18.

Birhanu, A.G., Gambardella, A. and Valentini, G., 2016. Bribery and investment: Firm-level evidence from A frica and L atin A merica. Strategic Management Journal, 37(9), pp.1865-1877.

Borensztein, E. and Panizza, U., 2010. Do sovereign defaults hurt exporters?. Open Economies Review, 21(3), pp.393-412.

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

Campbell, D.T., 1950. The indirect assessment of social attitudes. Psychological Bulletin, 47(1), p.15. 50 Campos, N.F., Dimova, R.D. and Saleh, A., 2010. Whither corruption? A quantitative survey of the literature on corruption and growth. Claessens, S. and Laeven, L., 2003. Financial development, property rights, and growth. the Journal of Finance, 58(6), pp.2401-2436.

Dell'Ariccia, G., Detragiache, E. and Rajan, R., 2008. The real effect of banking crises. Journal of Financial Intermediation, 17(1), pp.89-112.

Diallo, B., 2017. Corporate governance, bank concentration and economic growth. Emerging Markets Review, 32, pp.28-37.

Dimant, E. and Tosato, G., 2018. Causes and effects of corruption: what has past decade's empirical research taught us? A survey. Journal of Economic Surveys, 32(2), pp.335-356.

Dreher, A. and Gassebner, M., 2013. Greasing the wheels? The impact of regulations and corruption on firm entry. Public Choice, 155(3), pp.413-432.

Dreher, A. and Gassebner, M., 2013. Greasing the wheels? The impact of regulations and corruption on firm entry. Public Choice, 155(3), pp.413-432.

Drury, A.C., Krieckhaus, J. and Lusztig, M., 2006. Corruption, democracy, and economic growth. International political science review, 27(2), pp.121-136.

Elbahnasawy, N.G. and Revier, C.F., 2012. The determinants of corruption: Cross-country-panel-data analysis. The Developing Economies, 50(4), pp.311-333.

Elbahnasawy, N.G. and Revier, C.F., 2012. The determinants of corruption: Cross-

country-panel-data analysis. The Developing Economies, 50(4), pp.311-333.

Enste, D. and Heldman, C., 2017. Causes and consequences of corruption: An overview of empirical results. Fisher, R.J., 1993. Social desirability bias and the validity of indirect questioning. Journal of consumer research, 20(2), pp.303-315.

Fisman, R. and Svensson, J., 2007. Are corruption and taxation really harmful to growth? Firm level evidence. Journal of development economics, 83(1), pp.63-75. 51 Galindo, A. and Micco, A., 2004. Do state owned banks promote growth? Cross-country evidence for manufacturing industries. Economics letters, 84(3), pp.371-376.

Gaviria, A., 2002. Assessing the effects of corruption and crime on firm performance: Evidence from Latin America. Emerging Markets Review, 3(3), pp.245-268.

Gründler, K. and Potrafke, N., 2019. Corruption and economic growth: New empirical evidence. European Journal of Political Economy, 60, p.101810.

Gupta, S., Davoodi, H. and Alonso-Terme, R., 2002. Does corruption affect income inequality and poverty?. Economics of governance, 3(1), pp.23-45.

Gyimah-Brempong, K., 2002. Corruption, economic growth, and income inequality in Africa. Economics of governance, 3(3), pp.183-209.

Hauk, E. and Saez-Marti, M., 2002. On the cultural transmission of corruption. Journal of Economic theory, 107(2), pp.311-335.

Hauk, E. and Saez-Marti, M., 2002. On the cultural transmission of corruption. Journal of Economic theory, 107(2), pp.311-335.

Hopkin, J., 2002. States, markets and corruption: a review of some recent literature. Review of International Political Economy, 9(3), pp.574-590.

Huang, C.J., 2012. Corruption, economic growth, and income inequality: evidence from ten countries in Asia. International Journal of Economics and Management Engineering, 6(6), pp.1141-1145.

Jain, A.K., 2001. Corruption: A review. Journal of economic surveys, 15(1), pp.71-121.

Knack, S. and Keefer, P., 1995. Institutions and economic performance: cross-country tests using alternative institutional measures. Economics & Politics, 7(3), pp.207-227.

Lederman, D., Loayza, N.V. and Soares, R.R., 2005. Accountability and corruption: Political institutions matter. Economics & politics, 17(1), pp.1-35.

Leff, N.H., 1964. Economic development through bureaucratic corruption. American behavioral scientist, 8(3), pp.8-14.

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

Lipset, S.M. and Man, P., 1960. The social bases of politics. Baltimore: The Johns Hopkins UniversityPress. Mauro, P., 1995. Corruption and growth. The quarterly journal of economics, 110(3), pp.681-712

Méon, P.G. and Weill, L., 2010. Is corruption an efficient grease?. World development, 38(3), pp.244-259.

Mo, P.H., 2001. Corruption and economic growth. Journal of comparative economics, 29(1), pp.66-79.

Myrdal, G., 1968. Corruption: Its causes and effects. Asian drama: An inquiry into the poverty of nations, 2, pp.953-961.

Neeley, S.M. and Cronley, M.L., 2004. When research participants don't tell it like it is: pinpointing the effects of social desirability bias using self vs. indirect-questioning. ACR North American Advances.

Nye, J.S., 1967. Corruption and political development: A cost-benefit analysis. American political science review, 61(2), pp.417-427.

O'Toole, C.M. and Tarp, F., 2014. Corruption and the efficiency of capital investment in developing countries. Journal of International Development, 26(5), pp.567-597.

Pecorino, P., 1992. Rent seeking and growth: The case of growth through human capital accumulation. Canadian Journal of Economics, pp.944-956.

Pellegrini, L. and Gerlagh, R., 2004. Corruption's effect on growth and its transmission channels. Kyklos, 57(3), pp.429-456. Rajan, R. and Zingales, L., 1996. Financial dependence and growth.

Reinikka, R. and Collier, P. eds., 2001. Uganda's recovery: the role of farms, firms, and government. World Bank Publications.

Reinikka, R. and Svensson, J., 2003. Survey techniques to measure and explain corruption (Vol. 3071). World Bank Publications.

Romer, P.M., 1994. The origins of endogenous growth. Journal of Economic perspectives, 8(1), pp.3-22.

Seker, M. and Yang, J.S., 2012. How bribery distorts firm growth: Differences by firm attributes. World Bank Policy Research Working Paper, (6046).

Seldadyo, H. and De Haan, J., 2006, April. The determinants of corruption: A literature survey and new evidence. In EPCS Conference, Turku, Finland (pp. 20-23).

Sherwood, G.G., 1981. Self-serving biases in person perception: A reexamination of projection as a mechanism of defense. Psychological bulletin, 90(3), p.445.

Tanzi, V., 2002. Pitfalls on the road to fiscal decentralization. Carnegie Endowment for International Peace..

Transparency International (2012): Corruption Perception Index 2012: Short Methodological Note. Berlin.

Transparency International, 2020. Corruption Perceptions Index 2020: Technical Methodology Note Treisman, D., 2000. The causes of corruption: a cross-national study. Journal of public economics, 76(3), pp.399-457.

Ugur, M. and Dasgupta, N., 2011. Corruption and economic growth: A meta-analysis of the evidence on low-income countries and beyond.

Ugur, M. and Dasgupta, N., 2011. Evidence on the economic growth impacts of corruption in low-income countries and beyond: a systematic review. University of London, EPPI-Centre.

Vial, V. and Hanoteau, J., 2010. Corruption, manufacturing plant growth, and the Asian paradox: Indonesian evidence. World Development, 38(5), pp.693-705.

Vial, V. and Hanoteau, J., 2010. Corruption, manufacturing plant growth, and the Asian paradox: Indonesian evidence. World Development, 38(5), pp.693-705.

Wei, S.J., 2000. How taxing is corruption on international investors?. Review of economics and statistics, 82(1), pp.1-11. World Bank, 2003. World Development Report 2004 Overview.

World Bank Publications World Bank, 2015. Understanding the sample methodology. www.enterprisesurveys.org

World Bank, 2022. https://www.enterprisesurveys.org/en/graphing-tool.*Aghion, P., Hemous, D. and Kharroubi, E., 2014. Cyclical fiscal policy, credit constraints, and industry growth. Journal of Monetary Economics, 62, pp.41-58.*

Aghion, P., Hemous, D. and Kharroubi, E., 2014. Cyclical fiscal policy, credit constraints, and industry growth. Journal of Monetary Economics, 62, pp.41-58.

Alt, J.E. and Lassen, D.D., 2003. The political economy of institutions and corruption in American states. Journal of Theoretical Politics, 15(3), pp.341-365.

Andwig, J.C., 1991. The economics of corruption: A survey. Studi economici. Arif, I., Khan, L. and Waqar, S., 2020. Does corruption sand or grease the wheels? A case of BRICS countries. Global Business Review, p.0972150920927370.

Balta, N. and Nikolov, P., 2013. Financial dependence and growth since the crisis. Quarterly Report on the Euro Area (QREA), 12(3), pp.7-18.

Balta, N. and Nikolov, P., 2013. Financial dependence and growth since the crisis. Quarterly Report on the Euro Area (QREA), 12(3), pp.7-18.

Banerjee, A., Mullainathan, S. and Hanna, R., 2012. Corruption (No. w17968). National Bureau of economic research.

Becker, G.S. and Stigler, G.J., 1974. Law enforcement, malfeasance, and compensation of enforcers. The Journal of Legal Studies, 3(1), pp.1-18.

Birhanu, A.G., Gambardella, A. and Valentini, G., 2016. Bribery and investment: Firm-level evidence from A frica and L atin A merica. Strategic Management Journal, 37(9), pp.1865-1877.

Borensztein, E. and Panizza, U., 2010. Do sovereign defaults hurt exporters?. Open Economies Review, 21(3), pp.393-412.

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

Campbell, D.T., 1950. The indirect assessment of social attitudes. Psychological Bulletin, 47(1), p.15. 50 Campos, N.F., Dimova, R.D. and Saleh, A., 2010. Whither corruption? A quantitative survey of the literature on corruption and growth. Claessens, S. and Laeven, L., 2003. Financial development, property rights, and growth. the Journal of Finance, 58(6), pp.2401-2436.

Dell'Ariccia, G., Detragiache, E. and Rajan, R., 2008. The real effect of banking crises. Journal of Financial Intermediation, 17(1), pp.89-112.

Diallo, B., 2017. Corporate governance, bank concentration and economic growth. Emerging Markets Review, 32, pp.28-37.

Dimant, E. and Tosato, G., 2018. Causes and effects of corruption: what has past decade's empirical research taught us? A survey. Journal of Economic Surveys, 32(2), pp.335-356.

Dreher, A. and Gassebner, M., 2013. Greasing the wheels? The impact of regulations and corruption on firm entry. Public Choice, 155(3), pp.413-432.

Dreher, A. and Gassebner, M., 2013. Greasing the wheels? The impact of regulations and corruption on firm entry. Public Choice, 155(3), pp.413-432.

Drury, A.C., Krieckhaus, J. and Lusztig, M., 2006. Corruption, democracy, and economic growth. International political science review, 27(2), pp.121-136.

Elbahnasawy, N.G. and Revier, C.F., 2012. The determinants of corruption: Cross-country-panel-data analysis. The Developing Economies, 50(4), pp.311-333.

Elbahnasawy, N.G. and Revier, C.F., 2012. The determinants of corruption: Cross-country-panel-data analysis. The Developing Economies, 50(4), pp.311-333.

Enste, D. and Heldman, C., 2017. Causes and consequences of corruption: An overview of empirical results. Fisher, R.J., 1993. Social desirability bias and the validity of indirect questioning. Journal of consumer research, 20(2), pp.303-315.

Fisman, R. and Svensson, J., 2007. Are corruption and taxation really harmful to growth? Firm level evidence. Journal of development economics, 83(1), pp.63-75. 51 Galindo, A. and Micco, A., 2004. Do state owned banks promote growth? Cross-country evidence for manufacturing industries. Economics letters, 84(3), pp.371-376.

Gaviria, A., 2002. Assessing the effects of corruption and crime on firm performance: Evidence from Latin America. Emerging Markets Review, 3(3), pp.245-268.

Gründler, K. and Potrafke, N., 2019. Corruption and economic growth: New empirical evidence. European Journal of Political Economy, 60, p.101810.

Gupta, S., Davoodi, H. and Alonso-Terme, R., 2002. Does corruption affect income inequality and poverty?. Economics of governance, 3(1), pp.23-45.

Gyimah-Brempong, K., 2002. Corruption, economic growth, and income inequality in Africa. Economics of governance, 3(3), pp.183-209.

Hauk, E. and Saez-Marti, M., 2002. On the cultural transmission of corruption. Journal of Economic theory, 107(2), pp.311-335.

Hauk, E. and Saez-Marti, M., 2002. On the cultural transmission of corruption. Journal of Economic theory, 107(2), pp.311-335.

Hopkin, J., 2002. States, markets and corruption: a review of some recent literature. Review of International Political Economy, 9(3), pp.574-590.

Huang, C.J., 2012. Corruption, economic growth, and income inequality: evidence from ten countries in Asia. International Journal of Economics and Management Engineering, 6(6), pp.1141-1145.

Jain, A.K., 2001. Corruption: A review. Journal of economic surveys, 15(1), pp.71-121.

Knack, S. and Keefer, P., 1995. Institutions and economic performance: cross-country tests using alternative institutional measures. Economics & Politics, 7(3), pp.207-227.

Lederman, D., Loayza, N.V. and Soares, R.R., 2005. Accountability and corruption: Political institutions matter. Economics & politics, 17(1), pp.1-35.

Leff, N.H., 1964. Economic development through bureaucratic corruption. American behavioral scientist, 8(3), pp.8-14.

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

Lipset, S.M. and Man, P., 1960. The social bases of politics. Baltimore: The Johns Hopkins UniversityPress. Mauro, P., 1995. Corruption and growth. The quarterly journal of economics, 110(3), pp.681-712

Méon, P.G. and Weill, L., 2010. Is corruption an efficient grease?. World development, 38(3), pp.244-259.

Mo, P.H., 2001. Corruption and economic growth. Journal of comparative economics, 29(1), pp.66-79.

Myrdal, G., 1968. Corruption: Its causes and effects. Asian drama: An inquiry into the poverty of nations, 2, pp.953-961.

Neeley, S.M. and Cronley, M.L., 2004. When research participants don't tell it like it is: pinpointing the effects of social desirability bias using self vs. indirect-questioning. ACR North American Advances.

Nye, J.S., 1967. Corruption and political development: A cost-benefit analysis. American political science review, 61(2), pp.417-427.

O'Toole, C.M. and Tarp, F., 2014. Corruption and the efficiency of capital investment in developing countries. Journal of International Development, 26(5), pp.567-597.

Pecorino, P., 1992. Rent seeking and growth: The case of growth through human capital accumulation. Canadian Journal of Economics, pp.944-956.

Pellegrini, L. and Gerlagh, R., 2004. Corruption's effect on growth and its transmission channels. Kyklos, 57(3), pp.429-456. Rajan, R. and Zingales, L., 1996. Financial dependence and growth.

Reinikka, R. and Collier, P. eds., 2001. Uganda's recovery: the role of farms, firms, and government. World Bank Publications.

Reinikka, R. and Svensson, J., 2003. Survey techniques to measure and explain corruption (Vol. 3071). World Bank Publications.

Romer, P.M., 1994. The origins of endogenous growth. Journal of Economic perspectives, 8(1), pp.3-22.

Seker, M. and Yang, J.S., 2012. How bribery distorts firm growth: Differences by firm attributes. World Bank Policy Research Working Paper, (6046).

Seldadyo, H. and De Haan, J., 2006, April. The determinants of corruption: A literature survey and new evidence. In EPCS Conference, Turku, Finland (pp. 20-23).

Sherwood, G.G., 1981. Self-serving biases in person perception: A reexamination of projection as a mechanism of defense. Psychological bulletin, 90(3), p.445.

Tanzi, V., 2002. Pitfalls on the road to fiscal decentralization. Carnegie Endowment for International Peace..

Transparency International (2012): Corruption Perception Index 2012: Short Methodological Note. Berlin.

Transparency International, 2020. Corruption Perceptions Index 2020: Technical Methodology Note Treisman, D., 2000. The causes of corruption: a cross-national study. Journal of public economics, 76(3), pp.399-457.

Ugur, M. and Dasgupta, N., 2011. Corruption and economic growth: A meta-analysis of the evidence on low-income countries and beyond.

Ugur, M. and Dasgupta, N., 2011. Evidence on the economic growth impacts of corruption in low-income countries and beyond: a systematic review. University of London, EPPI-Centre.

Vial, V. and Hanoteau, J., 2010. Corruption, manufacturing plant growth, and the Asian paradox: Indonesian evidence. World Development, 38(5), pp.693-705.

Vial, V. and Hanoteau, J., 2010. Corruption, manufacturing plant growth, and the Asian paradox: Indonesian evidence. World Development, 38(5), pp.693-705.

Wei, S.J., 2000. How taxing is corruption on international investors?. Review of economics and statistics, 82(1), pp.1-11. World Bank, 2003. World Development Report 2004 Overview.

World Bank Publications World Bank, 2015. Understanding the sample methodology. www.enterprisesurveys.org

World Bank, 2022. https://www.enterprisesurveys.org/en/graphing-tool.

9. Appendices

9.1. Appendix 1 - Sampling methodology, stratification, and sample size

The Enterprise Survey aims to archive the following objectives:

- *provide statistically significant investment climate indicators that are comparable across countries.*
- *assess the constraints to private sector growth and job creation.*
- *build a panel of firm-level data that will make it possible to track changes in the business environment over time, thus allowing impact assessments of reforms.*
- *stimulate dialogue on reform opportunities.*

The sampling methodology of the Enterprise Survey determines the sample sizes based on two objectives. First, to be able to compare the investment climate of different economies across the world. Second, to conduct firm performance analysis focusing on how investment climate constraints affect job creation and productivity.

To achieve those objectives, the sampling methodology:

- *Is constructed to generate a sample that is representative of the whole economy that corroborates assertions about the country (not only the manufacturing sector). Therefore, the sample includes manufacturing sectors, service industries,*

Table 5: Number of industries and sample size based on GNI.

Size	GNI as of 2005	# of manuf. industries	# of services industries	Rest of the economy	Total sample size
Small	$5-25 billion	2	1	1	480
Medium	$25-80 billion	3	1	1	600
Large	$80-200 billion	4	2	1	840
Very large	>$200 billion	6	2	1	1080

Note: Table subtracted from World Bank (2015). Size refers to the differentiation of countries based on the GNI level. The total sample size refers to the number of firms surveyed in each industry.

and other relevant sectors.

- *Is formed of a sample size that is large enough to conduct statistically robust analysis with levels of precision of at least 7.5% precision for 90% confidence intervals about estimates of the mean of the logarithm of sales at the industry level and estimates of population proportions, also at the industry level.*

To accomplish the objective of having a sample representative enough to make statements about the whole economy, the population of industries included in the Enterprise Survey are (according to ISIC, version 3.1): construction (Group F), all manufacturing (Group D), services (Groups G and H), and transport, storage, and communication (group I).

The universe of industries is stratified into several manufacturing industries, two service industries, and a residual. The number of industries defined as an individual stratum in each country is chosen based on the Gross National Income (GNI) of that specific country. The number of strata by type can be seen in Table 5.

To improve comparability across countries, two manufacturing industries were surveyed in all countries of the sample: manufacturing of food and beverage (ISIC 15), and manufacturing of wearing apparel and fur (ISIC 18). Additional industries are also selected at the two-level ISIC, depending on certain sector characteristics of each country: the number of firms from that sector, the number of employees, and their contribution to value-added. However, the selection of industries tries to keep similar industries across countries to facilitate cross-country comparability. Some very small countries did not have enough establishments to stratify at the two-level ISIC. In those cases, the sample of 240 manufacturing firms was randomly chosen from the whole manufacturing sector.

To constrain the survey to the formal economy, only firms with more than 5 employees are included in the basic Enterprise Survey's sample. A separate extra module for firms with less than 5 employees has been generated as a complement, allowing to study of micro-enterprises that tend to operate under the veil of the informal economy. However, the relationship that those micro-firms have with corrupt government officials may differ significantly from one of the firms in the formal sector. For example, they do not usually depend on foreign trade so the need for permission to import or export is not necessary to them. Consequently, I only focus on the basic sample of formal firms and do not analyze micro firms.

The sampling methodology implemented is stratified random sampling. Therefore, population units are grouped within homogeneous groups and simple random samples are selected within each group. This allows computing the estimates of each stratum

with a specific level of precision while estimating population estimates by properly weighting individual observations.

The strata chosen on the Enterprise Survey are business sector, firm size, and geographical region within a country. The sector is divided between manufacturing, retail, and other services. For larger economies, sub-sectors are selected as additional strata based on value-added, employment, and the total number of establishments.

The firm's size levels are 5-19 (small), 20-99 (medium), and more than 100 (large). Geographic regions within a country are selected based on which regions enclose most of the economic activity. The Enterprise Survey reports the weight for each observation. The data used in this analysis is then weighted based on these factors.

9.1.1. Sample size

As mentioned, the sample size is determined under the condition that is large enough to conduct statistically robust analysis with levels of precision of at least 7.5% precision for 90% confidence intervals about estimates of the mean of the logarithm of sales at the industry level and estimates of population proportions, also at industry level. The sample size required for proportions may differ from the size required for the mean of log sales. To guarantee that both conditions are fulfilled, the maximum sample size out of the two conditions is selected.

Table 6 shows the minimum sample size for different population sizes for estimates of proportions with 5% and 7.5% precision in a 90% confidence interval, assuming maximum variation. With the 5% precision, the minimum sample size tends to a sample size 270, as the population size increases; with 7.5% precision, the sample size tends to be 120. Note that if the population size of an industry falls below 1,500, the required sample size for proportions may be reduced considerably. Although a 5% precision would be most desirable, a precision of 7.5% is used, since is in line with the budget that the World Bank determined in the Enterprise Survey.

To determine the minimum sample size for the estimates of means of a specific variable it is necessary to have an estimate of the variance of that variable. The variable under analysis, in this case, is the sales of each specific establishment. For areas where the survey has been implemented in the past, the information about the variance of sales can be determined based on historical data.

As sales usually present a skewed distribution, the sample size required to do conclusions about its means is typically too large. Consequently, it is common practice to transform sales data to logarithm, and therefore significantly reduce its variability. After this transformation, the sample size requirements for inferences about the mean are lower than the one required for proportions. In other words, working with the minimum sample size required for proportions with 7.5% precision and 90% confidence intervals generally meets the requirement of the minimum sample size for inferences about the mean logarithm of sales with a precision of 5% (even more demanding than for proportions). Whenever the data of previous sales are available, the minimum sample size required for both conditions is analyzed and the largest one is chosen. However, when there is no information about previous sales, the Enterprise Survey relies on the previous condition, focusing just on the minimum sample size for proportions.

Table 6: Sample Sizes Required with 5% and 7.5% Precision and 90% Confidence

Population size	Sample Size 5%	Sample Size 7.5%
50	42	36
100	73	55
200	115	75
300	143	86
400	162	93
500	176	97
600	187	100
700	195	103
800	202	105
900	208	106
1000	213	107
1250	223	110
1500	229	111
1750	234	113
2000	238	113
2500	244	115
3000	248	116
5000	257	117
10000	263	119
50000	269	120
100000	270	120

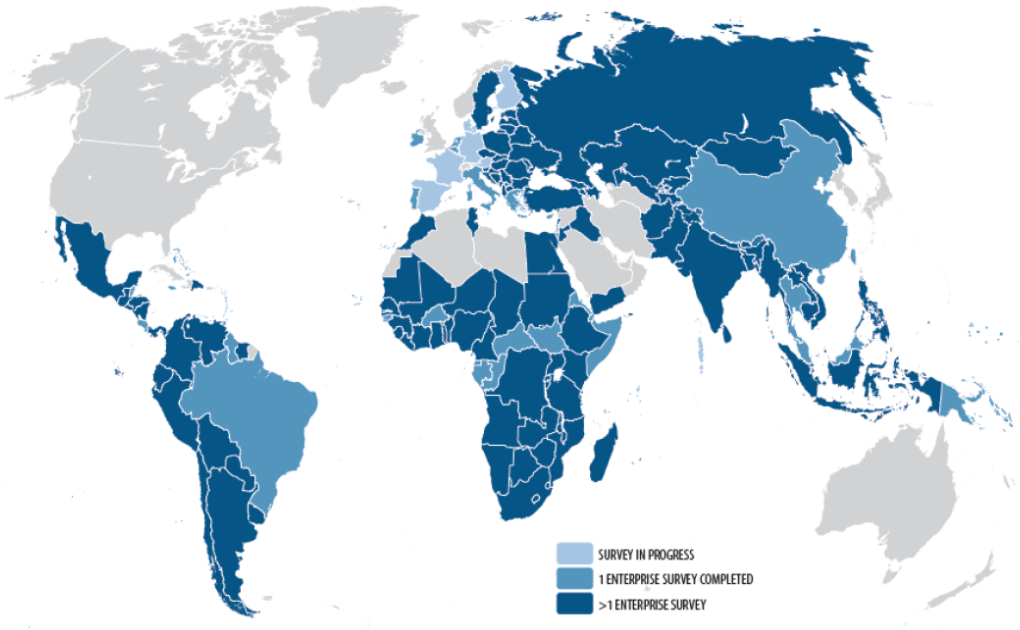
Note: Table subtracted from World Bank (2015)..

10. Appendix 2 - Enterprise Survey – Country coverage

The Enterprise Surveys are conducted across all geographic regions and cover small, medium, and large firms. It covered more than 210.000 firms in 159 countries around the world. Figure 2 shows the countries currently being surveyed, the ones that have one completed survey, and the ones with more than one completed Enterprise Survey.

Unsurprisingly, the data contain missing observations (especially in questions regarding bribes), and therefore, my analyses will use fewer observations than the ones on the full sample.

Figure 2: Enterprise Surveys around the world.



Notes: Figure subtracted from World Bank (2022). Surveys were conducted in different years.