



INSTITUTIONS AND DEFORESTATION IN THE BRAZILIAN AMAZON: A GEOGRAPHIC REGRESSION DISCONTINUITY ANALYSIS

Ingvild Engen Bogetvedt & Mari Johnsrud Hauge

Supervisor: Torfinn Harding

Master thesis: MSc Economics and Business Administration
Majors in Economics / Energy, Natural Resources & the Environment

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

ABSTRACT

This study explores the impact of institutional quality at the municipal level on deforestation in the Legal Amazon. We add to this insufficiently understood topic by implementing a geographic regression discontinuity design. By taking advantage of high-resolution spatial data on deforestation combined with an objective measure of corruption used as a proxy for institutional quality, we analyse 138 Brazilian municipalities in the period of 2002-2004. Our empirical findings show no causal effect of institutional quality on deforestation, suggesting that other unobserved factors are more important drivers of deforestation. A supplementary analysis indicates that location is a vital underlying factor in the interplay between deforestation and institutional quality. However, further research is needed to establish inference with respect to the relationship between deforestation, institutional quality and location.

TABLE OF CONTENTS

1. INTRODUCTION.....	3
2. LITERATURE REVIEW.....	7
2.1 POPULATION GROWTH.....	7
2.2 ECONOMIC FACTORS	7
2.3 INSTITUTIONAL FACTORS	9
2.4 CORRUPTION.....	12
2.5 CORRUPTION AND DEFORESTATION IN THE LEGAL AMAZON	15
3. BACKGROUND	18
3.1 TROPICAL DEFORESTATION	18
3.2 DEFORESTATION IN THE LEGAL AMAZON	19
3.3 CORRUPTION IN BRAZIL	20
4. DATA	22
4.1 DEFORESTATION DATA.....	22
4.2 CORRUPTION DATA.....	22
4.3 CREATING OUR DATASET.....	23

5. EMPIRICAL STRATEGY	25
5.1 IDENTIFICATION.....	26
5.2 SETUP AND NOTATION.....	30
5.3 THREATS TO IDENTIFICATION.....	34
6. FINDINGS AND DISCUSSION: MAIN SAMPLE	38
6.1 GRAPHICAL PRESENTATION.....	38
6.2 ECONOMETRIC ESTIMATION.....	39
6.3 ROBUSTNESS TESTS.....	41
6.4 INTERPRETATION.....	48
7. FINDINGS AND DISCUSSION: EXTENDED SAMPLE.....	52
7.1 GRAPHICAL PRESENTATION.....	53
7.2 ECONOMETRIC ESTIMATION.....	54
7.3 LOCATION OF MUNICIPALITIES.....	56
7.4 WITHIN-STATE ANALYSIS: PARÁ.....	59
7.5 INTERPRETATION.....	61
8. CONCLUSION.....	63
9. RECOMMENDATIONS FOR FUTURE RESEARCH.....	65
9.1 IMPROVED CONTROL OF UNOBSERVED FACTORS.....	65
9.2 INCREASED SAMPLE.....	66
9.3 PREDETERMINED INSTITUTIONS.....	67
9.4 QUALITY CHECKING THE DATA.....	68
9.5 OTHER ISSUES TO CONSIDER.....	70
REFERENCES.....	72
APPENDICES.....	78

1. INTRODUCTION

The world's forests produce global environmental goods that are essential to society at all scales and locations (Hargrave and Kis-Katos, 2013). Due to its size and nature, the Brazilian Amazon plays a particularly crucial role in the provision of such goods, and is therefore considered one of the earth's most valued treasures (Pailler, 2016). Nevertheless, it has been subject to extraordinary deforestation over the past half-century (Nepstad et al., 2009). This leads to intensification of climate change (Assunção et al., 2015), interruption of global ecosystem services (Foley et al., 2007), mass-extinction of species (Burgess et al., 2012), severe dislocations and even extinction of indigenous groups (Laurance, 1999). Hence, understanding how to effectively combat deforestation in the Amazon, and hinder the corresponding effects from reaching the level of irreversibility, has become a key priority on the global environmental policy agenda (Assunção et al., 2015; Burgess et al., 2012; Cisneros et al., 2013; Hargrave and Kis-Katos, 2013). In order to appropriately respond to the issue, policy makers must understand the effects of the full set of potential drivers of deforestation (Pfaff, 1999).

There have been remarkable efforts in several disciplines to identify the determinants of deforestation (Koyuncu and Yilmaz, 2009). Since the 1980s, the economic literature has focused on various human activities that cause deforestation (Barbier and Burgess, 2001). Initially, there was extensive focus on population growth as the prominent determinant (Cropper and Griffiths, 1994; Ehrhardt-Martinez, 1998; Laurance, 1999). However, newer literature exhibits results suggesting that the effects of population growth are highly exaggerated (Geist and Lambin, 2002; Koyuncu and Yilmaz, 2009), and that deforestation is more closely linked to the state of the economy and market forces (Assunção et al., 2015; Hargrave and Kis-Katos, 2013; Nepstad et al., 2009). Specific factors such as the economics of forestry and agriculture (Andersen, 1996; Assunção et al., 2015; Barbier and Burgess, 2001; Bhattarai and Hammig, 2001), as well as proximity from forest areas to infrastructure and economic markets (Angelsen and Kaimowitz, 1999; Burgess, 1993; Pfaff, 1999) are frequently cited as drivers of deforestation.

A topic that has received less attention in the empirical literature and is not adequately understood is the effect of more underlying forces on deforestation, such as institutional factors (Cisneros et al., 2013; Mendes and Porto Jr., 2012; Pailler 2016). Although several scholars have argued that institutions are likely to be important drivers, they have often been neglected in economic models

of deforestation (Bhattarai and Hammig, 2004; Umemiya et al., 2010). Nonetheless, of the empirical studies that have been conducted on the topic, many confirm that institutional factors have a substantial impact on deforestation (Barbier and Burgess, 2001; Geist and Lambin, 2002; Umemiya et al., 2010). Thus, the relationship between institutional quality and deforestation deserves further scrutiny, which is therefore the purpose of this study.

Due to the complex nature and difficulty of obtaining appropriate data on institutions (Barbier and Burgess, 2001; Umemiya et al., 2010; Vanclay, 1993), we deem it fruitful to focus on one of its key sub-components. Corruption is an essential aspect of institutional quality affecting deforestation (Barbier, 2004; Barrett et al., 2006), as forests possess certain characteristics that make them a potential breeding ground for illegal and corrupt activities (FAO, 2001). Forest activities often revolve around large and remote areas far from public scrutiny and government agencies, where local politicians are granted broad discretionary power to make decisions on highly subjective matters. The insight that corruption may affect forests has led some authors to make use of the expression *forest corruption* (Sundström, 2016, p. 781), and some have argued that deforestation is merely a symptom of societal problems such as corruption (Vanclay, 1993).

Similarly to institutional factors in general, there is limited empirical research on the impact of corruption on deforestation. However, much of the existing studies on the topic find that corruption leads to increased deforestation (Barbier et al., 2005; Koyuncu and Yilmaz, 2009; Meyer et al., 2003; Smith et al., 2003; Wright et al., 2007). Nonetheless, there are noteworthy problems with these findings regarding the quality of the corruption data utilised, and that they are based on cross-country analysis. The corruption measures are based on perceived rather than actual corruption, and refers to the performance of federal governments (Ferraz and Finan, 2011). The vast majority of tropical forests, including the Brazilian Amazon, are locally governed and national corruption measures are unable to capture sub-national variation and country-specific contexts in-depth (Angelsen, 2009). This limits the ability to make meaningful comparisons between countries, accurately assess the effects of corruption on deforestation, and may even lead to invalid conclusions. Inference with respect to this important and complex relationship requires far more nuance than what commonly appears in published studies (Barrett et al., 2006).

Consequently, we will proceed with a within-country analysis of the interaction between corruption and deforestation in the Legal Amazon, which requires a local-level corruption

measure (Mendes and Porto Jr., 2012). In 2003, in response to serious concerns about the extent of corruption in Brazil, the government created an autonomous federal agency that executes audits on municipal governments, randomly selected in a national lottery (Ferraz and Finan, 2011). This has made it possible to construct a new and objective measure of corruption in Brazilian municipalities (Cisneros et al., 2013; Ferraz and Finan, 2011), which enables us to overcome the data limitations observed in the existing literature.

Moreover, after auditing around 2,000 municipalities and billions worth of governmental funds (Avis et al., 2016), it has been revealed that corruption at the local level is a serious problem in Brazil. As the Brazilian Amazon is de facto controlled by local agents, combined with the many ways in which corruption can lead to deforestation, the urgency for improved understanding of the mechanisms affecting the management of this forest is exacerbated. To our knowledge, only three studies analyse the relationship between deforestation and corruption at the municipal level in the Legal Amazon, using the corruption data made available by the federal audits (Cisneros et al., 2013; Mendes and Porto Jr., 2012; Pailler, 2016). However, they are looking at the issue from different perspectives, and there is a great need for more work to supplement their findings. Therefore, we intend to add to this branch of literature. Nevertheless, we deem it unlikely that we can isolate the effect of corruption on deforestation, as it is interdependent on other institutional factors. Consequently, we will rather regard corruption as a proxy for institutional quality.

We combine high-quality datasets on deforestation and corruption levels across Brazilian municipalities and utilise an innovative geographic regression discontinuity design (Keele and Titiunik, 2016; Lee and Lemieux, 2010) to investigate the relationship between institutional quality and deforestation. Our study can be regarded as a natural experiment, where we observe neighbouring municipalities that have, for natural reasons, different levels of corruption. We exploit this discontinuity in institutional quality at municipal borders and assess whether there is systematically corresponding variation in the level of deforestation. The most crucial aspect of our identification strategy is that institutional quality should be the only feature that changes, whereas all other factors that may affect deforestation remain continuous when crossing a municipal border (Galiani et al., 2017; Keele and Titiunik, 2014; Sekhon and Titiunik, 2012). We perform several tests to ensure that this assumption of continuity holds, and that the municipalities are comparable along the most salient dimensions that may influence deforestation.

Our main empirical finding contradicts the notion voiced by the literature where institutions are central drivers of deforestation. We find no statistically significant discontinuity in the amount of deforestation when moving from high to low institutional quality municipalities in the Legal Amazon, which implies that other features may play a more crucial role in determining the level of deforestation. This is an unexpected finding, considering the high levels of deforestation and corruption seen in Brazilian municipalities, combined with the fact that the Legal Amazon forest is controlled by local agents, and the many mechanisms through which poor institutional quality likely leads to increased deforestation. However, as the empirical literature on this relationship is in its infancy and definite evidence is scarce, our study is adding novel insight to this insufficiently understood topic.

In addition to our main research, we conduct a supplementary analysis where we find indications of location being a vital underlying factor in the interplay between deforestation and institutional quality. Location is likely to be a key influencer, as it affects economic factors such as access to markets, transportation costs, economies of scale and agglomeration effects (Angelsen and Kaimowitz, 1999; Alves, 2002; Burgess, 1993; Hargrave and Kis-Katos, 2013; Krugman, 1991; Pfaff, 2007). We cannot, however, say anything about the direction in which the possible relationship between the three variables, deforestation, institutional quality and location, works. Overall, complementary research is much needed before any grand conclusions should be drawn.

The remainder of this paper is organised as follows: In section 2 and 3 we present a review of the literature and background information on deforestation and corruption, with specific focus on the Legal Amazon. These sections are recommended for readers who are interested in the topic and the specific context we study, but are not necessary for understanding our research, and can thus be regarded as optional. In section 4 we describe our data. We formulate our empirical strategy in section 5. In section 6 and 7 we present our findings and discussion. Section 8 concludes, before section 9 provides recommendations for future research.

2. LITERATURE REVIEW

2.1 POPULATION GROWTH

The early economic literature on the drivers of deforestation in the Legal Amazon characterises it as a state-driven process. The military government in power from the 1960s to 1980s implemented large scale settlement projects that initiated a surge of deforestation (Alston et al., 2000; Hargrave and Kis-Katos, 2013). This clearing of the region was justified by the government as a release valve for pressures arising from a growing population. Accordingly, population growth has been widely researched and frequently cited as a prominent determinant of deforestation (Cropper and Griffiths, 1994; Ehrhardt-Martinez, 1998; Laurance, 1999). However, only a few scholars have confirmed this hypothesis empirically, and newer literature has exhibited results of alternative drivers of deforestation, which suggests that the early literature has overemphasised population growth as a primary cause of deforestation (Geist and Lambin, 2002; Koyuncu and Yilmaz, 2009). This newer stream of literature, which generally considers deforestation from the 1980s onwards, explores the issue at a much finer scale and usually finds deforestation to be more closely linked to the state of the economy and market forces (Assunção et al., 2015; Hargrave and Kis-Katos, 2013; Nepstad et al., 2009).

2.2 ECONOMIC FACTORS

ECONOMIC GROWTH AND THE ENVIRONMENTAL KUZNETS CURVE

A particular focus in the debate on the relationship between economics and deforestation is whether there is an Environmental Kuznets Curve (EKC) for forests analogous to that found for air and water quality (Foster and Rosenzweig, 2003). The EKC hypothesis predicts an inverted-U-shaped relationship between economic growth and environmental degradation, where environmental degradation tends to increase in early stages of economic development, before it decreases in the later stages (Koyuncu and Yilmaz, 2009). Evidence on the existence of such relationships would be of considerable interest to policy makers. According to Koyuncu and Yilmaz (2009), there are several studies finding strong evidence for the presence of EKC, but there is no agreement on where the turning point takes place. Barbier and Burgess (2001) on the other hand, emphasise that income effects tend to vary between regions, and do not always display an EKC relationship. Furthermore, other studies using cross-country regressions find no evidence to support the existence of an EKC relationship (Barbier, 2004; Meyer et al., 2003). Hence, there

is still uncertainty with regards to both the features and the existence of the EKC relationship. Moreover, this literature is focusing on the impact of the overall state of the national economy on deforestation. Finding causal relations between these two forces can be very challenging, as macroeconomic variables influence society through complex and often indirect paths (Angelsen and Kaimowitz, 1999). Furthermore, tropical deforestation is an outcome that is likely to stem from interaction between many factors and mechanisms (Vanclay, 1993). Therefore, it has been argued that it is more sensible to attempt to single out more specific factors that influence tropical deforestation one at the time, in order to get a more nuanced picture of the phenomenon.

FORESTRY AND AGRICULTURE

Understandably, timber harvesting and commercial logging are frequently cited drivers of deforestation (Bhattarai and Hammig, 2001). In this regard, the prices on timber is expected to be of significant importance. However, the literature is not conclusive on this topic, although it seems to be a direct source of deforestation in some contexts and an indirect source in others (Angelsen and Kaimowitz, 1999). Economics of agriculture and agricultural development has also been recognised as a predominant cause of forest loss in tropical regions (Barbier and Burgess, 2001). Assunção et al. (2015) find empirical evidence that deforestation responds to agricultural output prices in the sense that falling agricultural commodity prices lead to less deforestation. Certain types of agriculture, such as soybean and cattle, have been given considerable attention and are characterised as central determinants of deforestation (Andersen, 1996). During the last decade, deforestation rates have become closely correlated with the prices of these two commodities, both in spatial and time dimensions. For instance, using panel data analysis, Hargrave and Kis-Katos (2013) found statistically significant evidence that deforestation rates were positively affected by increases in soybean and meat prices.

ACCESS TO MARKETS

Analytical models and empirical studies find that increased access from forest areas to economic markets accelerates deforestation (Angelsen and Kaimowitz, 1999). Several studies have showed a strong relation between deforestation and proximity to infrastructure, as it facilitates access to markets (Burgess, 1993) and reduces transport costs (Pfaff, 1999). Improved transportation systems also tend to encourage economic development activities that accelerate deforestation further (Hargrave and Kis-Katos, 2013).

In essence, while economic factors certainly are intertwined and directional impact of each of these factors have not been set in stone, it has been widely documented that there is a clear link between economics and deforestation.

2.3 INSTITUTIONAL FACTORS

A topic that has received less attention in the empirical literature and is not adequately understood is the effect of more underlying forces on deforestation, such as institutional factors and governance quality (Bhattarai and Hammig, 2004; Cisneros et al., 2013; Mendes and Porto Jr., 2012; Pailler, 2016). Institutions can be defined as humanly devised constraints that structure human interaction and incentives in society. They are made up of formal constraints such as laws and constitutions, and informal constraints such as conventions and norms of behaviour. In other words, institutions are the *rules of the game* in society (North, 1994). Governance refers to the way in which power is exercised in the management of a country's resources (World Bank, 1992). Both terms relate to politics and political economy.

Several scholars have argued that the quality of governance and underlying institutions are likely to be important determinants of deforestation (Bhattarai and Hammig, 2004; Umemiya et al., 2010). According to Eliasch (2008), it is widely recognised that the way in which forests are governed influences the extent of deforestation. Yet, we have limited in-depth understanding of the exact roles of the quality of governance affecting deforestation (Umemiya et al., 2010), and institutional factors are usually neglected in economic models of deforestation (Bhattarai and Hammig, 2001). Thus, important questions regarding the impact of such factors on tropical deforestation remain unanswered, and should be addressed promptly.

Nonetheless, of the empirical studies that have been conducted on the topic, many confirm that institutional factors and governance quality have a substantial impact on deforestation (Barbier and Burgess, 2001; Geist and Lambin, 2002).

In a global study, Umemiya et al. (2010) examined the relationship between governance quality and deforestation and found statistically significant indications of a negative association, even when controlling for other intervening variables, such as forest cover ratio, population growth, and expansion of agricultural land. In a similar style study, Bhattarai and Hammig (2001; 2004) found that strengthening the quality of governance and institutions plays a vital role in tropical

forest management. Their results also suggest that underlying institutional factors are relatively more important in explaining the tropical deforestation process than other frequently cited factors like population and macroeconomic conditions (Bhattarai and Hammig, 2001). This emphasizes the importance of governance on forest dynamics and management, which is why they recommend that institutional dimensions of the deforestation problem need to be better scrutinised (Bhattarai and Hammig, 2001; 2004).

In a less macro-oriented study, Gibson et al. (2000) discovered that the combination of national, regional, and local institutions play a critically important role in the consumption of forest resources. Furthermore, in researching Brazil's efforts to reduce deforestation in the Amazon, Assunção et al. (2015) found that changes to conservation policies implemented by the government in 2004 and 2008 significantly contributed to curbing deforestation, even after controlling for a variety of economic indicators.¹ In fact, counterfactual simulations suggest that conservation policies avoided approximately 56% of total forest clearings that would have occurred between 2005 and 2009 had the policies not been implemented. Similarly, Hargrave and Kis-Katos (2013) presented evidence that deforestation rates were significantly affected by government policies concerning designated settlement and protection areas as well as the local presence of IBAMA in the period of 2002 to 2009. Burgess et al. (2012) also established a link between deforestation and political economy in their study of changed borders of political jurisdictions in Indonesia, and corresponding deforestation outcomes. They found that an increase in the number of political jurisdictions lead to more competition and a following surge of deforestation, which serves as evidence that incentives faced by local politicians are a key determinant affecting tropical deforestation.

However, there is not unanimous confirmation that policies best explain changed deforestation rates in the literature. In studying the conservation zones established as a part of a larger policy initiative to curb deforestation in the Legal Amazon, Anderson et al. (2016) present difference-in-difference estimates indicating that these are not the reason for the large reduction in deforestation rates. This somewhat contradicts the results of Assunção et al. (2015). Anderson et al. (2016) rationalise their findings with the fact that conservation zones are located in areas where

¹ See appendix 5 for a figure created by Assunção et al. (2015), which illustrates deforestation rates, agricultural price trends and years indicating the implementation of the two major conservation policies.

agricultural production is unlikely to be profitable. The study also reveals that the conservation zones only reduce deforestation if the incentives for municipalities to reduce deforestation are high. Nevertheless, this indicates that there is an important relationship between institutional factors and deforestation. Forests have distinct characteristics that make them subject to policies and governance quality. In fact, according to Ascher (1999), no resource is immune to institutional forces or political processes, be it direct or indirect. Accordingly, the World Bank emphasises the role of good governance within its forest sector (Barrett et al., 2006). Yet, the empirical basis for linking deforestation to institutional factors is scanty (Bhattarai and Hammig, 2001; Pailler, 2016). Thus, the impact of institutional quality on deforestation deserves increased attention.

A potential reason for the lack of empirical research on the topic could be that the mechanisms of deforestation and institutions and the linkages between the two are so complex and diverse that finding causal relationships is problematic (Umemiya et al., 2010). As previously mentioned, tropical deforestation is an outcome that is likely to stem from interaction between many factors (Vanclay, 1993). Moreover, identifying the causal effect of institutions is challenging, as they are usually intertwined with other factors and because institutional characteristics are themselves endogenous equilibrium outcomes (Cust and Harding, 2014). Furthermore, there is a problem of retaining appropriate institutional data (Barbier and Burgess, 2001), so information to precisely define the reported quality of governance is not available. Consequently, the exact role of governance quality in relation to deforestation remains unclear (Umemiya et al., 2010), and there is a need for more and improved research on this topic.

Hitherto, terms such as institutional factors, institutions, institutional quality, governance, governance quality, politics, political economy have been used interchangeably. Although they are highly related and interdependent, they are not synonymous. Hence, for the sake of clarity, we will focus and base our research on *the effect of institutional quality on deforestation*.

Due to the challenges faced when trying to estimate the effect institutional quality, it can be fruitful to examine one or more of its key sub-components, such as government effectiveness, corruption and rule of law, and consider these as measures that represent the overall institutional quality (Michalopoulos and Papaioannou, 2013; Williams and Siddique, 2008). FAO (2001) contends that one of the most important aspects of institutional quality affecting forests is crime and corruption, and that ongoing efforts to improve forest management will have limited value

unless accompanied by measures to reduce it. Similarly, in studying key institutional factors thought to influence deforestation, Barbier (2004) found that only corruption has a direct impact, more so than political stability and rule of law. In relation to deforestation outcomes, corruption can thus be deemed a sensible sub-component of institutional quality to base our research on.

2.4 CORRUPTION

Corruption can be defined as the abuse of entrusted power for private gain (Transparency International, 2017). According to FAO (2001), corruption and forest crime is widespread in the countries that house our largest tropical forests. Important reasons for this is that forest activities often concern large and remote areas, far from public scrutiny and government agencies, where local politicians in office for relatively short periods are granted broad discretionary power and are largely unsupervised to make decisions on highly subjective matters. This creates a favourable environment for illegal activities, and the prospects of public officers to enhance their private wealth by supporting the overharvesting of natural resources are great. According to Barrett et al. (2006), it is undisputable that corrupt politicians and bureaucrats play a vital role in environmental degradation.

Corruption opportunities are plenty and arise in many stages of forest management which may directly affect the speed of deforestation (Barrett et al., 2006; Søreide, 2007). For instance, in the land-use allocation process, corrupt public officials may facilitate logging without permits, issue logging permits that are not in accordance with ecological criteria, intentionally under-enforce laws that conserve resources and turn a blind eye towards activities that increase deforestation (Ascher, 1999; Fearnside, 2003; Wibowo and Byron, 1999). Corruption can also lead to deforestation via more indirect channels. For instance, farmers who bribe corrupt politicians in exchange for agricultural subsidies tend to have lower productivity, which triggers the expansion of agricultural land, and in turn, the depletion of forests (Bulte et al., 2007). In collusion with local economic interests, corrupt local governments can also lobby the state legislatures against the enlargement of protected areas (Cisneros et al., 2013). Consequently, Vanclay (1993) argues that unsustainable deforestation is merely a symptom rooted in more serious societal problems, such as corruption. Illustratively, as it has been estimated that 80% of the deforestation in the Brazilian Amazon is illegal, it seems clear that this would be impossible without substantial corruption taking place (Kolstad and Søreide, 2009). The insight that corruption may affect forest management has led some authors to make use of the expression *forest corruption* (Sundström,

2016, p. 781). Furthermore, the Forest Integrity Network has proposed to establish a Forest Sector Corruption Transparency Index, to make the problem more visible and identify countries and actors that are leading the way in good forest governance (Sundström, 2016).

Nevertheless, despite the scale of the problem, and the fact that FAO (2001) has warned that immediate attention must be given to corruption and its impact on the world's remaining forests, very little empirical research has been done on the topic (Kolstad and Søreide, 2009). There is considerable lack of knowledge about the actual extent of deforestation that might be directly or indirectly attributed to corruption (Angelsen, 2009). Empirical research on the impacts of corruption on the forest sector is in its infancy, and more in-depth understanding is needed to provide firm recommendations on policies and prioritisation areas in different countries (Angelsen, 2009; Sundström, 2016). Yet, in this relatively new and growing field of scholarly research, several authors have found empirical evidence that largely supports the claim that corruption increases deforestation directly (Sundström, 2016).

A cross-country econometric study assessing expansion of agricultural land areas in tropical low and middle income countries in Africa, Asia and Latin America in the period 1961–1999, found that a 1% reduction in corruption may decrease deforestation by 0.17% to 0.30% (Barbier et al., 2005). Similarly, a study of 117 countries in the period 1990-2000 found that countries with less corruption are less likely to liquidate forest assets (Meyer et al., 2003). Both studies utilise corruption data from the World Bank. Focusing on 823 forest reserves in a smaller number of tropical countries in the period 2002-2004, Wright et al. (2007) came to a similar conclusion. Their findings suggest that protected areas are more effectively managed where corruption is low, using the Corruption Perception Index (CPI) by Transparency International. To improve the robustness of the analysis, Koyuncu and Yilmaz (2009) made use of three different measurements of corruption; the CPI, the International Country Risk Guide Index (ICRG), the Business Intelligence Index (BI), and official figures on deforestation for three different time periods from 100 countries. Their findings reveal a highly statistically significant positive correlation between corruption and deforestation, and that corruption generally has more explanatory power than other determinants of deforestation. Their findings remain valid in both univariate and multivariate models. Similarly, using CPI and ICRG, Smith et al. (2003) find strong evidence that an increase in corruption leads to deforestation. They therefore stress the need for developing policies that reduce the effects of corruption in order to conserve the remaining tropical forests of the world.

QUALITY OF CORRUPTION DATA

However, a significant problem with all the abovementioned studies relates to the quality of their corruption data. As it represents illicit and hence often hidden activity, obtaining objective and complete data on corruption is a difficult task (Ferraz and Finan, 2011; Mendes and Porto Jr., 2012). To overcome this challenge, many international corruption indices have been developed by non-governmental organisations, like the ones utilised in the studies discussed above (Koyuncu and Yilmaz, 2009). However, the problem with these indices is that they are based on subjective perceptions and accusations, rather than actual corruption (Angelsen, 2009; Ferraz and Finan, 2011), as they are constructed with surveys completed by business people and economic leaders (Williams and Siddique, 2008). It is also common that corruption is defined differently from one survey to another, and that many indices use imprecise measures for corruption. For instance, one of the most widely used indices of institutional quality, the ICRG, assigns corruption scores to countries based on how the government came to power and their length of time in office. Although this relates to corruption, it is neither a precise nor direct measure of corruption per se, a point often missed by researchers who are seeking to identify the effect of corruption on various outcomes (Williams and Siddique, 2008). Corruption definitions can also change between surveys conducted by the same agency in different years. Additionally, the measures in different countries are influenced by tax rates, which bias the estimates (Timmons and Garfias, 2015).

Inconsistencies and lack of complete information indicate that data on corruption are still far from adequate. One cannot be certain whether these indices provide representative data for any given country or that the data is comparable between countries. Indeed, according to Williams and Siddique (2008, p. 143), there exists *no definite* governance measure based on survey data. Naturally, the lack of objective and complete corruption data limits the ability of analysts to perform an accurate study on the effects of corruption (Ferraz and Finan, 2011) on deforestation, and may even lead to invalid conclusions about how corruption may influence forests (Barrett et al., 2006). Inference with respect to this important and complex relationship requires far more nuance than what commonly appears in published studies (Barrett et al., 2006).

NATIONAL VS. LOCAL CORRUPTION

Another issue in the existing literature on the relationship between corruption and deforestation is that it focuses on cross-country analysis, using national and not within-country estimations of corruption (Mendes and Porto Jr., 2012). National corruption scores are largely generalised and

based on the performance of federal governments. Corruption that affects deforestation may appear at different levels of government, and by using national estimations we risk assessing countries whose central government is relatively non-corrupt, yet in remote regions local officials contribute to deforestation and biodiversity loss (Barrett et al., 2006). In fact, this is most probably the case, as the vast majority of tropical forests are de facto controlled by local agents. In addition, there may be a limited reach of the state, where areas located further from capital centres are less affected by national institutions (Michalopoulos and Papaioannou, 2013). National levels of corruption will in this case not reflect the actual significance of institutions on all areas in the country. Therefore, in studying corruption and its impact on deforestation, it is essential to understand institutional quality at the local level (Burgess et al., 2012). Similarly, in a summary of lessons learned in natural resource management, the US Agency for International Development emphasises the importance of good governance at the local level (Barrett et al., 2006).

Generalised measures of national corruption are not able to capture and analyse such sub-national variation and country-specific contexts in-depth (Angelsen, 2009). Consequently, studies comparing different countries based on such estimations may fail to capture the true relationship between corruption and deforestation (Barrett et al., 2006; Ferraz and Finan, 2011). Compared to within-country analysis, great variation between countries imposes additional difficulties when analysing differences across nations, as many factors must be considered in order to isolate the effect of corruption. Unobservable features may therefore cause an omitted variable problem, as they cannot be controlled for, even though they are influencing the outcome. Hence, essential components of this complex relationship may be lost in attempts to generalise in current studies (Barrett et al., 2006). The potential inaccuracy of analysis should be a cause for concern, and international studies need careful assessment before one can draw conclusions on the relationship between corruption and deforestation (Angelsen, 2009).

2.5 CORRUPTION AND DEFORESTATION IN THE LEGAL AMAZON

Due to the limitations of cross-country analysis on the relationship between corruption and deforestation, it is important to redirect the focus towards within-country estimations of the relationship, and use credible corruption data. To our knowledge, there are three studies that make

the connection between corruption and deforestation on the municipal level in the Legal Amazon, using objective corruption data.²

Mendes and Porto Jr. (2012) analyse the relationship between deforestation, corruption and economic growth in 538 Brazilian municipalities in 2004. Initially, they study municipalities in all nine Amazon states, and find no evidence on the relationship between corruption and deforestation. However, in the two states with the highest levels of deforestation, Mato Grosso and Pará, they find statistically significant estimates suggesting that municipal corruption is a key variable in explaining the higher level of deforestation. An important reason for performing within-state analysis, is that it naturally controls for important characteristics influencing the outcome, such as timber and meat prices, culture and legal issues. Consequently, by keeping these factors constant, the researchers are able to isolate the effects of corruption and economic growth.

Cisneros et al. (2013) investigate the relationship between federal corruption audits, local governance quality and deforestation dynamics in 209 Brazilian municipalities in the period 2002 to 2009. Remarkably, they find that deforestation in Amazon municipalities increase as a consequence of the otherwise successful anti-corruption programme, suggesting that deforestation may be an unintended consequence of the audits. Deforestation increased on average by at least 11% in the aftermath of public fiscal audits, with larger increases in more corrupt municipalities. With regards to the relationship between corruption and deforestation, descriptive analysis shows that municipalities with about one standard deviation higher measured corruption levels experienced up to 20% higher deforestation between 2002 and 2009. However, a significant share of this effect disappeared once other initial socio-economic determinants are controlled for. Cisneros et al. (2013) attribute the increase in deforestation to mayors shifting illegal behaviour to activities that are difficult to monitor through public audits.

In studying the interaction between corruption, election incentives and deforestation at the municipal level in the Amazon, Pailler (2016) finds strong links between electoral cycles, corruption and deforestation. In the 12-month period during and immediately following elections, deforestation rates increase by 8-11%, and this appears to be driven by corruption rather than

² These objective corruption data will be elaborated on in section 3.3.

agricultural activities and policies. Corrupt municipalities have nearly 50% more deforestation in election periods than non-corrupt municipalities and municipalities without incumbent mayors running for re-election. The corruption effect accounts for all the observed increase in deforestation during and after elections. In resemblance to Cisneros et al. (2013), Pailler (2016) rationalises this finding by the notion that politicians may shift or re-direct their corrupt activities to deforestation, and exploit forest resources rather than other resources, as such behaviour is less observed than other forms of rent-extraction. In addition to being difficult to detect (Burgess et al., 2011; Kolstad and Søreide, 2009), deforestation generates high localised benefits, while its costs are widely dispersed across the voting population and time periods. Hence, increased deforestation is not likely to be punished, but rather rewarded with voters.

Both Pailler (2016) and Cisneros et al. (2013) contend that the combination of the authority given to Brazilian mayors over resource allocation and responsibility to ensure local economic growth is the key explanation for why they are likely to encourage deforestation activities. Furthermore, economic interests have a significant influence on local politics as they are usually very tightly linked, and sometimes even overlap. Mayors may have economic interests in, or close personal ties to, local logging, farming and sawmill facilities (Cisneros et al., 2013). If the local government is corrupt as well, this will likely exacerbate deforestation rates.

Hence, the relationship between corruption and deforestation has been identified in the Legal Amazon, albeit in a relatively small number of studies. There is a great need for more work to supplement these findings. Therefore, we intend to add to this branch of literature. Nevertheless, we deem it unlikely that we can isolate the effect of corruption on deforestation, as it is strongly interdependent on other institutional factors. Consequently, we will rather regard corruption as a proxy for institutional quality in our study. This has also been done by others (Driffield et al., 2016; Esiyok and Ugur, 2017).

3. BACKGROUND

3.1 TROPICAL DEFORESTATION

The world's forests produce global public goods in the form of ecosystem services, biodiversity conservation and carbon sequestration, which counteracts climate change (FAO, 2016). Due to their size and nature, tropical forests play a particularly crucial role in the provision of such goods, and are therefore essential to society at all scales and locations (Hargrave and Kis-Katos, 2013).

Nevertheless, forests have diminished globally over the past 400 years (Foster and Rosenzweig, 2003). It has been estimated that the world's original forest cover was approximately 6 billion hectares (Bryant et al., 1997). According to the Global Forest Resources Assessment Report (FAO, 2010), the earth's remaining forest area is just over 4 billion hectares, implying that the world has lost about one third of its forest areas. This is to a large degree a result of extensive deforestation activities. Deforestation is defined as a process where forests are cleared by human activities or destroyed by natural disasters, and then permanently converted into other land uses such as wasteland, cropland and pasture or left as abandoned land (Koyuncu and Yilmaz, 2009; Meyer et al., 2003).

The adverse effects of tropical deforestation are vast. It has been estimated that it accounts for 7-14% of the global CO₂ emissions, which is the main driver of climate change (European Commission, 2014). In addition, tropical deforestation interrupts the functioning of the global ecosystem services, for example by leading to changed evapotranspiration, pollination and cloud cover, which affects other natural environments (Foley et al., 2007). While occupying a mere 7% of the earth's land surface, tropical rainforests sustain more than half of the planet's life forms, making them our most biodiverse environments (Wilson, 1988). Hence, their disappearance results in a mass-extinction of species, whose value is associated with genetic diversity (Burgess et al., 2012). Moreover, tropical deforestation is causing severe dislocations and even extinction of indigenous groups living in these areas (Laurance, 1999). Hence, understanding how to effectively combat tropical deforestation and hinder the corresponding effects from reaching the level of irreversibility, has become a topic of major global concern and a key priority on the international environmental policy agenda (Assunção et al., 2015; Burgess et al., 2012; Cisneros et al., 2013).

3.2 DEFORESTATION IN THE LEGAL AMAZON

The three greatest areas of tropical forest in the world are located in Indonesia, Brazil and the Democratic Republic of Congo (Burgess et al., 2012). As the largest rainforest on the planet, the Brazilian Amazon sustains around 40% of the world's remaining tropical forest system (Kirby et al., 2006) and is consequently one of the earth's most valued treasures (Pailler, 2016).

At the same time, it has been subject to extraordinary deforestation over the past half century, and Brazil has been characterised as a world leader in deforestation (Nepstad et al., 2009). The 1990s and early 2000s exhibited record-breaking deforestation rates in the Amazon, with an average of 19,500 km² per year from 1996 to 2005 (Anderson et al., 2016). It has therefore been referred to as the world's most active agricultural frontier in terms of forest loss and CO₂ emissions by FAO (2006). These massive levels of deforestation raise all the mentioned environmental issues, making it a matter that not only affects its home country Brazil, but the entire earth. The conservation of this forest has therefore become a major concern from both a national and international perspective (Hargrave and Kis-Katos, 2013).

At the international level, large-scale initiatives have been established. A significant one is the Reducing Emissions from Deforestation and forest Degradation (REDD+) initiative that emerged from negotiations under the United Nations Framework Convention on Climate Change in 2005 (den Besten et al., 2014). The central idea is to enhance forest carbon stocks in the tropics with an incentive-based system, where finances could be generated for the protection of forests by creating forest carbon credits. Today, more than 40 countries participate and hundreds of projects have been initiated across the tropics, but attention has recently been moved to national and local levels (Angelsen, 2009).

At the national level, the Brazilian government has made substantial commitments to reduce deforestation (Nepstad et al., 2009). For example, by establishing the federal environmental agency, Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) who is responsible for the enforcement of environmental law. Furthermore, the Amazon Fund has been established to raise investments for efforts to prevent, monitor and combat deforestation of the Brazilian Amazon. It is administered by the Brazilian Development Bank, and supports around 90 projects worth US\$620 million (Amazon Fund, 2017). Moreover, Brazil has considerably increased its law enforcement activities during the last decade and implemented a number of

conservation policies designed to protect its forests (Arima et al., 2014; Cisneros et al., 2013). The two most important actions are the 2004 Action Plan for the Prevention and Control of Deforestation in the Legal Amazon and the 2008 initiative of blacklisting of municipalities with critical deforestation rates (Assunção et al., 2015).

In the second half of the 2000s, the pace of forest clearings in the Amazon was slowed down substantially. After peaking in 2004, with an annual deforestation rate of 27,000 km², it fell sharply over the following years to about 7,000 km² in 2009 (Assunção et al., 2015). Deforestation rates have continued to fall since then, and in 2013 they were a mere 10% of the levels observed in 2000 (Pailler, 2016). Nevertheless, although the rates of deforestation have significantly slowed down, it remains a serious global environmental problem, and there is still a need for improved knowledge about determinants in order to appropriately respond to the issue.

3.3 CORRUPTION IN BRAZIL

In addition to being the home to the largest tropical forest on earth, Brazil is one of the most decentralised countries in the world (Avis et al., 2016; Ferraz and Finan, 2011). The political jurisdictions under the federal government are divided into 26 states and 5,570 municipalities, and these enact their own constitutions and laws and collect taxes (Pailler, 2016). Furthermore, each year, billions of dollars are transferred from the federal government to municipal governments (Ferraz and Finan, 2011). With minimal federal oversight, local officers decide how to allocate these large influxes of funds (Fundo de Participação dos Municípios) on a significant share of public services. In this system, corruption at the local level has become prevalent, and positively correlated with the size of federal transfers (Brollo et al., 2013).

In response to serious concerns about the extent of corruption in Brazil, the federal government created in 2003 Controladoria Geral da União (CGU) (Avis et al., 2016). It is a functionally autonomous federal agency that possesses constitutional powers resembling that of a ministry. The same year as it was established, the CGU introduced an ambitious, unprecedented anti-corruption program that audits municipalities for their use of federal funds (Ferraz and Finan, 2011). The audits are executed on municipalities that are randomly selected by a national lottery system held on a monthly basis at the Caixa Economica Federal in Brasilia (Ferraz and Finan, 2011). To assure a fair and transparent process, representatives of the press, political parties, and members of the civil society are invited to witness the lottery.

Around 50-60 municipalities are selected in each lottery (Ferraz and Finan, 2011). Once a municipality is chosen, approximately 10 to 15 CGU auditors are sent to the municipality to examine information on all federal funds transferred to the municipal government from 2001 and onwards. After approximately one week of inspections, a report (50-150+ pages) of detailed substantive and procedural information about the manner in which funds were spent is made (Timmons and Garfias, 2015). They contain the total amount of federal funds transferred to the municipal administration and the amount audited, as well as explicit identification of corruption, theft, and other improper expenditure, and in most cases the amount of funds involved. The reports are submitted to the central CGU office in Brasilia and then sent to the Tribunal de Contas da Uniao (TCU), to public prosecutors, and to the legislative branch of the municipality. For each municipality audited, a summary of the main findings is posted on the internet and disclosed to media sources (Ferraz and Finan, 2011).

As of February 2015, there has been 2,241 audits across 40 lotteries in 1,949 municipalities and over US\$7.7 billion worth of federal funds audited (Avis et al., 2016). The audits have uncovered that 79% of Brazilian municipalities have had at least one incidence of corruption and virtually every municipality (99%) has committed some act of mismanagement. Based on estimates by Ferraz and Finan (2011), corruption in local governments is responsible for losses of approximately US\$550 million per year. Thus, corruption at the local level in Brazil is a serious problem.

In the Legal Amazon alone, there are more than 600 municipalities (Cisneros et al., 2013). Although Brazil's environmental regulation and conservation policies are centralised, Brazilian mayors have authority over resource allocation and policy implementation at the municipal level (Ferraz and Finan, 2011). Furthermore, economic interests are tightly linked to the politicians, and they have substantial power to influence local deforestation levels. Due to the high corruption rates and the many mechanisms through which this likely exacerbates deforestation, we contend that urgent attention must be given to the link between corruption and deforestation in the Legal Amazon.

4. DATA

To identify differences in deforestation and institutional quality across Brazilian municipalities, we take advantage of two independent datasets. The outcome of interest in our analysis is deforestation, and the main explanatory variable is corruption, which functions as a proxy for institutional quality.

4.1 DEFORESTATION DATA

Data on deforestation is obtained from high-resolution NASA satellite images (Anderson et al., 2016). These have been processed into objective spatial deforestation data by the Project for Monitoring Deforestation in the Legal Amazon (PRODES) at the Brazilian Institute for Space Research (INPE). Anderson et al. (2016) aggregate these high-resolution data up to grid-cells of 1km^2 , and this is what we use in our analysis. The resulting dataset covers the entire Legal Amazon, with deforestation data on each individual cell. Furthermore, we have information on what municipality each cell is located in, which allows us to zoom in and compare deforestation rates on both cell and municipal level (Anderson et al., 2016).

The deforestation data corresponding to each grid-cell in the dataset is measured as the proportion of each cell deforested. Amount of forest loss for each observation is recorded over the 12 months leading up to August of a given year (Anderson et al., 2016). Using the coordinate of the centre of the grid cells, each cell is also assigned other geo-specific information, such as distances to cities, rivers, roads and IBAMA offices. In addition, the data contains information on initial forest cover, share of non-forest and whether the cell is in a conservation zone (Anderson et al., 2016).

4.2 CORRUPTION DATA

As mentioned, the unique policy experiment of public audits done by the CGU has made it possible to construct a new and objective measure of corruption levels in Brazilian municipal governments (Cisneros et al., 2013; Ferraz and Finan, 2011). As the process of selecting

municipalities to be audited happens through a fully randomised national lottery,³ the resulting sample can be considered representative of all Brazilian municipalities.

The reports were first coded by Ferraz and Finan (2008; 2011), who calculated the number of irregularities and share of federal resources affected by corruption (Cisneros et al., 2013). They were the first ever to present an objective measure of municipal corruption (Ferraz and Finan, 2008). Subsequently, Brollo et al. (2013) classified each occurrence in the audit reports in the spirit of Ferraz and Finan (2008; 2011), and the resulting dataset is what we use in this study.⁴ Brollo et al. (2013) created categorical variables for different corruption levels and amounts. The corruption category referred to as *broad* includes a great set of irregularities, and some of these can also be interpreted as bad administration rather than overt corruption. The *narrow* category only includes severe irregularities that are related to corruption. We will only employ this corruption measure, as it is stricter and thus avoids inclusion of general violations that may not be considered corruption. When a municipality is considered corrupt in the *narrow* category, we can assume that it has worse institutional quality than a municipality without irregularities of great severity. The corruption variable is a dummy that takes on the value one if the municipality had any irregularity reported, while it is zero when the auditors did not find any irregularities.

Brollo et al. (2013) analyse 1,202 municipalities that were randomly selected to be audited by the CGU in lottery 2 to 29, in the time period 2001-2008. This period consists of two full mayoral terms; 2001-2004 and 2005-2008. The corruption score indicates level of corruption in a municipality for a full mayoral term. There were 802 audited municipalities in the first term and 400 audited municipalities in the second term. Because we use an econometric approach that does not require several periods, we use cross-sectional data and focus on the first mayoral term. We choose this term as it leaves us with more observations, and hence greater statistical power.

4.3 CREATING OUR DATASET

The two individual datasets described above are merged together. As mentioned, we focus on the political term 2001-2004, where the municipal mayor and corruption level in each municipality

³ See section 3.3 for detailed information about the CGU audits and the national lotteries.

⁴ See Brollo et al. (2013) for an overview and explanation of the irregularities.

stay constant during the entire period. Therefore, time specific variability in the data can be ignored, and we are dealing with cross-sectional data. We follow Anderson et al. (2016), and exclude the year 2001 in the deforestation data. Because this is the first year where deforestation data is available, it suggests excessively high levels of deforestation, and the maps report some patterns yet to understand (Anderson et al., 2016). Hence, the period we observe is in fact 2002-2004.⁵ Furthermore, we generate an aggregate deforestation measure, which is the accumulated amount of deforestation observed throughout this period. Moreover, we only include the nine states that are housing parts of the Legal Amazon. In addition, we exclude all municipalities in the deforestation data that have not been audited between 2001 and 2004. Hence, we have ensured that each grid-cell in the dataset has one observation for deforestation and one for corruption.

As will be discussed in the empirical strategy section, it is crucial to analyse neighbouring municipalities. These will be referred to as municipal *pairs*. Therefore, we need information on the specific location of each grid-cell, which we obtain from the same satellite images as the actual data on deforestation. By using these data, we know which municipality border that lies closest to each cell, and the distance to this border measured from the centre of the cell. In other words, the neighbouring municipality for each grid-cell is identified. Furthermore, as will be discussed, in a geographic discontinuity analysis we need to identify an abrupt change in the main explanatory variable (Galiani et al., 2017). Therefore, we use data on corruption for both the observed cell and its neighbour to detect if there is a change in corruption when crossing the border. This allows us to generate a dataset that only includes observations where the nearest border is associated with a discontinuity in corruption.

TABLE 1 – DESCRIPTIVE SUMMARY OF OUR DATASET

	Non-corrupt	Corrupt	Total
Observations	178 162	199 748	377 910
Municipalities	78	60	138
Pairs	-	-	108

Notes: Number of observations, municipalities and pairs, depending on the level of corruption.

⁵ However, as the deforestation data measures forest cleared over the 12 months leading up to August of a given year, we are effectively looking at deforestation between September 2001 and August 2004.

5. EMPIRICAL STRATEGY

The goal for our analysis is to establish whether there is a causal effect of what we consider the treatment; institutional quality as proxied by corruption, on the outcome of interest; deforestation. Identifying causality would be a relatively simple task if corruption was assigned randomly among Brazilian municipalities. In other words, if on average, the municipalities were comparable on all other accounts than the treatment of interest. As randomisation is such an important concept when establishing causal relationships, the gold standard for inference is said to be experimental studies with randomised control groups (Banerjee and Duflo, 2011; Keele and Titiunik, 2016; Sundstöm, 2016). Unfortunately, creating an experiment where municipalities are randomly assigned to being corrupt is in reality inconceivable, and we are unable to construct perfect counterfactuals (Gibbons and Overman, 2012). The absence of randomisation entails challenges with regards to selection bias and endogeneity when establishing inference (Keele and Titiunik, 2014).

Different strategies have been developed to overcome such problems and serve as substitutes for the gold standard. As such, natural experiments and quasi-experimental methods are considered the best alternatives and are therefore frequently used to make causal inferences (Keele and Titiunik, 2014; Sekhon, 2009; Sekhon and Titiunik, 2012). Our study can be regarded as a natural experiment, where we observe existing levels of corruption that have been assigned to Brazilian municipalities by nature and not in a controlled, randomised manner. Although natural experiments are regarded as the best substitute for randomised experiments, they exhibit some distinctive features that give rise to inferential and conceptual challenges (Sekhon and Titiunik, 2012). To ensure validity of our results, it is therefore crucial that we employ a credible empirical strategy. Furthermore, Barbier and Burgess (2001) stress that novel and insightful models are needed to study the drivers of deforestation. Accordingly, we intend to generate more knowledge in this field by analysing data through a Regression Discontinuity (RD) design, which is regarded one of the most convincing and innovative econometric designs for natural experiments (Keele and Titiunik, 2016; Lee and Lemieux, 2010).

RD is considered a quasi-experimental design, and it allows us to test causal hypotheses even if randomisation is not conceivable (Lee and Lemieux, 2010). Furthermore, it is argued that causal inferences from this methodology is potentially more credible than those from *typical natural experiments* such as difference-in-differences and instrumental variables (Lee and Lemieux, 2010,

p. 282). According to Calonico et al. (2016), the RD design is considered one of the most credible econometric techniques because it uses relatively weak and easy-to-interpret identifying assumptions, which permits flexible and robust estimation and inference for treatment effects.

5.1 IDENTIFICATION

The geographic nature of our data enables us to utilise a special case of RD, namely a Geographic Regression Discontinuity (GRD) design (Black, 1999; Keele and Titiunik, 2014). We develop a GRD design that exploits the discontinuity in the quality of institutions as proxied by corruption at Brazilian municipal borders to identify its effect on deforestation. This design allows us to look at the average difference in deforestation rates when moving from control units to treated units.⁶

The most crucial aspect of our empirical design is that institutional quality should be the only feature that changes, whereas all other factors that may affect deforestation remain continuous when crossing a municipal border (Galiani et al., 2017; Keele and Titiunik, 2014; Sekhon and Titiunik, 2012). This is also referred to as the continuity assumption in GRD (Hahn et al., 2001). As we are researching the exploitation of a natural resource, a fundamental factor to consider with respect to the continuity assumption is the underlying natural geography (Anderson et al., 2016; Cust and Harding, 2014; Turner et al., 2014). For instance, it may be more attractive to clear out the forest in an area with high-quality soil for agriculture, or if the type of timber is of high economic value in the market. Furthermore, other drivers that have been identified in the literature to affect deforestation are population pressures (Cropper and Griffiths, 1994; Ehrhardt-Martinez, 1998; Laurance, 1999), local economic factors such as distances from forest areas to markets (Angelsen and Kaimowitz, 1999; Pfaff, 1999) and specific policies related to deforestation (Assunção et al., 2015; Hargrave and Kis-Katos, 2013). As the Legal Amazon is a vast and varied area, we cannot assume that these factors are comparable between all municipalities in the region, and especially not between those that lie far away from each other. Some even argue that it is infeasible to find a credible control group for a treatment group when the data is spatial, given that all locations are unique (Gibbons and Overman, 2012).

⁶ Our study relates to papers that investigate the role of institutions through GRD design, e.g.: Anderson et al. (2016), Cust and Harding (2014), Michalopoulos and Papaioannou (2013).

Although it may be impossible to find the perfect treatment and control groups in a geographic study like ours, the key is to find groups that are comparable along the most salient dimensions that influence the outcome. In our GRD design we can ensure this by analysing observations that are spatially adjacent (Keele and Titiunik, 2016). The central idea is that neighbouring municipalities are comparable on most unobservable factors, but that their institutional quality for some reason differs (Gibbons and Overman, 2012). Therefore, we use a matching technique and create pairs of neighbouring municipalities (Sundström, 2016) where there is a discontinuity in corruption at the border. By doing this, we are focusing on relatively small geographic areas, which enhances our confidence in the homogeneity and comparability of these areas in other factors than institutional quality. Furthermore, by analysing grid-cells that lie adjacent to the municipal border, the geographical area shrinks and further improves the credibility of our analysis (Keele and Titiunik, 2016). In this regard, Van Der Klaauw (2008) state that if it is reasonable to assume that units of observation just below and above the threshold are comparable, then we may view our GRD design as almost experimental near the threshold. Consequently, if the border only represents a discontinuity in institutional quality, any differences in deforestation at the border can then be interpreted as evidence of a causal relationship between these two observations (Imbens and Lemieux, 2008). Nonetheless, to ensure that the continuity assumption holds, we will now consider the most salient factors that could affect deforestation other than institutional quality.

GEOGRAPHY

By looking at grid-cells in a relatively small area around a border, we can assume that the natural geography is very similar on each side (Anderson et al., 2016). It is further ensured that neighbouring municipalities share environmental characteristics that are relatively stable over time, namely altitude, slope and soil composition, as they lie in similar *Homogeneous Response Units* (Cuaresma et al., 2017) (Appendix 1). Thus, we can assume that the borders are as-if randomly assigned with respect to the underlying geology (Cust and Harding, 2014). Furthermore, we account for the extent of initial forest cover by incorporating it into an alternative deforestation variable. This is done in the robustness tests section. Because there are areas without forest cover in the Legal Amazon, we also take this into account by including a control variable for land without forest in our main regression equation.

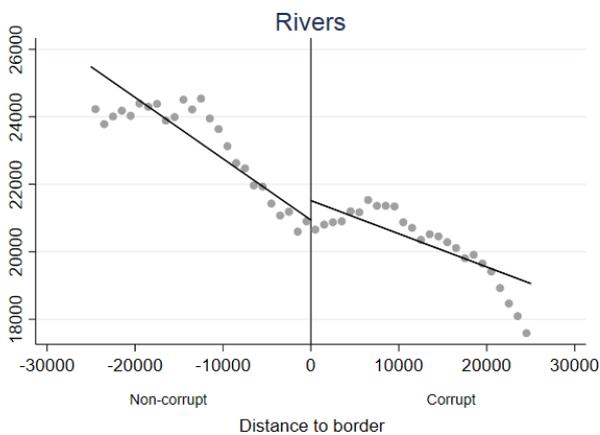
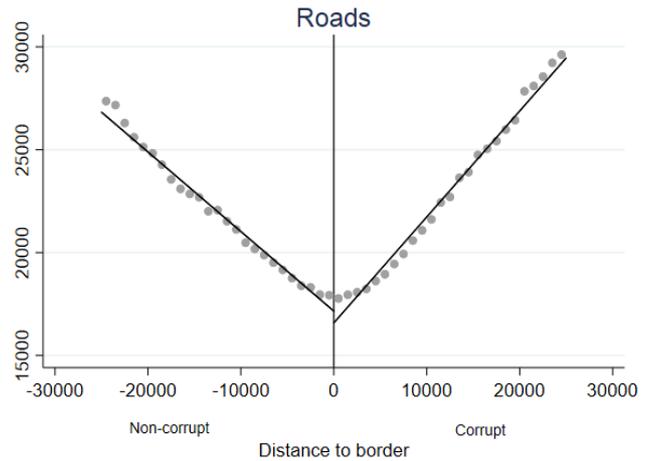
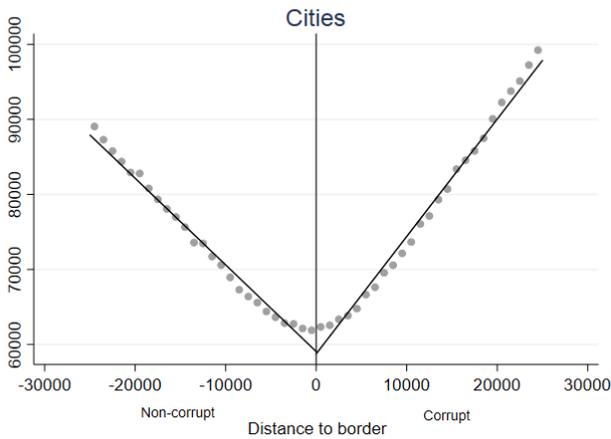
POPULATION GROWTH AND MUNICIPALITY SIZE

We have information about the size of the population for each municipality, obtained from Instituto Brasileiro de Geografia e Estatística (IBGE, 2016). These data are collected for the year 2004, as this is the last year of the mayoral term we are analysing. We control for the population size measure, as well as municipality size in the robustness test section.

LOCAL ECONOMIC FACTORS

Furthermore, we have information about several local economic factors, namely the distance from each grid-cell to cities, roads and rivers. Because we use a design where neighbouring observations are compared, these should as mentioned be continuous across the borders. To verify this, we plot the relationship between the distance to the nearest neighbour and the distance to each of the economic factors. The results are presented graphically in figures 1A, 1B and 1C.

FIGURES 1A, 1B & 1C – COVARIATES: DISTANCE TO CITIES, ROADS AND RIVERS



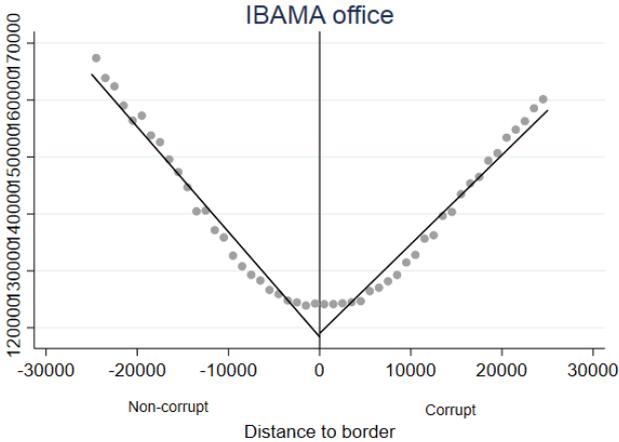
Notes: Distances to economic factors are plotted against the distance from each observation to nearest municipality border. 1km bins. Bandwidth of 25,000 meters.

According to Keele and Titiunik (2014), the credibility of the GRD design would be enhanced if balance improves as we approach the boundary. None of the graphs have a jump at the border, which implies that the factors are continuous across the boundary. Furthermore, the distance to cities and roads seems to be shorter for observations lying closer to a municipal border, independent on whether they are lying in a non-corrupt or corrupt municipality. Because there are more roads near cities, it makes sense that these variables are correlated in the sample. On the other hand, the graph plotting distance to rivers indicates a difference between control and treated units. It seems that, on average, observations lying on the corrupt side of the border are located closer to rivers. However, we are interested in making sure that the variables are continuous across the border, which all of these are.

POLICIES RELATED TO DEFORESTATION

We have information about policies directly concerning deforestation in the Legal Amazon, such as the presence of IBAMA and location of the conservation zones implemented by the federal government. Similarly to the local economic factors discussed above, we plot the distance to IBAMA offices against distance to the nearest neighbour in figure 2. These data have somewhat the same structure as distance to roads and cities, as the distance decreases when approaching the border. IBAMA offices are often located close to cities, and will therefore be correlated with the distance to cities variable. The graph shows no discontinuity at the border. In addition, to control for conservation zones, we add this as a variable in our main regression equation.

FIGURE 2 – COVARIATE: DISTANCE TO IBAMA OFFICES



Notes: Distance to IBAMA offices is plotted against the distance from each observation to nearest municipality border. 1km bins. Bandwidth of 25,000 meters.

OTHER FACTORS

In our analysis, we look at several municipal pairs that are located in different areas of the Legal Amazon. There may therefore be other additional unobservable factors influencing deforestation that differ from one pair to another. For instance, there could be differences in cultural traits between non-adjacent municipal pairs, and we take this into consideration by incorporating pair fixed effects in our analysis (Cust and Harding, 2014).

In essence, though there are unobserved aspects that influence deforestation and make each Brazilian municipality unique, we minimise the variation in this uniqueness by looking at neighbouring municipalities and controlling for other factors when necessary. While we cannot state that they are identical, we can make the assumption that a municipality and its neighbour are very similar along the most salient dimensions related to deforestation. Consequently, we can make causally informative comparisons (Gibbons and Overman, 2012).

5.2 SETUP AND NOTATION

TREATMENT, RUNNING VARIABLE & CUT-OFF

The GRD design relies on binary treatment assigned to units in the sample (Imbens and Lemieux, 2008), which in our case is the corruption level for each Brazilian municipality. Thus, we can divide all municipalities into treatment and control groups based on their level of corruption. In addition, as we are looking at geographic areas, treatment will change at physical borders. We therefore combine treatment status with geo-specific information for each observation, in order to analyse the ones that are lying close to a border. As a result, we generate an important variable; the running variable (Angrist and Pischke, 2014). This is created by assigning negative (positive) values of the distance to nearest border to observations located in non-corrupt (corrupt) municipalities. This places non-corrupt and corrupt observations on opposite sides of the third central feature of an RD design; the cut-off. Hence, at the cut-off there is an abrupt discontinuity in institutional quality. As we ensured that the most salient factors causing deforestation are continuous when crossing the border, units barely below the cut-off can be considered counterfactuals to the ones just above it. This is a key principle and a great advantage of the RD design (Cattaneo et al., 2017). If there is a difference in deforestation between these units, it can be attributed to the institutional quality.

REGRESSION EQUATION

We use a linear regression function where we regress deforestation on the binary corruption variable, the running variable and an interaction term. The binary corruption variable is our main explanatory variable, the running variable is considered a control variable and the interaction term is included in order to let the regression function differ on both sides of the cut-off (Lee and Lemieux, 2010). To establish a relationship between deforestation and institutional quality, we obtain estimates using the following equation:

$$Deforestation_i = \alpha + \beta Corruption_i + \gamma Running\ var_i + \delta Running\ var_i * Corruption_i + \eta_i \quad (1)$$

β is the coefficient of interest, γ is the coefficient of the running variable, δ is the coefficient of the interaction term, α is the intercept and η is the error term.

The central idea is to estimate whether there is a difference in the mean deforestation level at the cut-off, captured by a statistically significant β -coefficient. Its magnitude can be interpreted as the effect of crossing the border on the level of deforestation. The simple equation above is developed further by adding control variables and pair fixed effects, which we use in the main analysis. Control variables should not affect the coefficients in a correctly specified RD design (Cust and Harding, 2014), but they may reduce the sampling variability (Lee and Lemieux, 2010). We therefore investigate estimates when including a dummy variable for whether an observation is located in a conservation zone, and the amount of each grid-cell without forest cover.

PARAMETRIC VS. NONPARAMETRIC STRATEGY

We are faced with the choice between two types of strategies for correctly specifying the functional form; parametric and nonparametric. The difference between the two is that in a parametric strategy, one tries to find the right model to fit the dataset, whereas in the nonparametric strategy one tries to find right dataset to fit the model (Jacob et al., 2012). In the parametric strategy all available data is included, whereas in the nonparametric setting only observations lying in a smaller range around the cut-off is included (Cattaneo et al., 2017). When using a parametric strategy, one thus has to make assumptions about the correct functional form, which may be challenging.

As mentioned, focusing on grid-cells that lie adjacent to the municipal border is essential for the credibility of our analysis. Furthermore, excluding observations that lie far away from the threshold and use a smaller proportion of the data will make the nonparametric strategy perform better than the parametric strategy in terms of reducing bias. On the other hand, because of the smaller sample size, it may also lead to limited statistical power (Jacob et al., 2012). In general, however, the nonparametric approach provides a good compromise between flexibility and simplicity, and is recommended in practice (Cattaneo et al., 2017). We therefore use a nonparametric approach in our analysis.

MODEL SPECIFICATION

In the nonparametric strategy, the estimated treatment effect does not depend on the correct specification of the model (Angrist and Pischke, 2009), and we can interpret it as a local approximation (Cattaneo et al., 2017). In our main estimation we use a linear model with one interaction term as specified in equation (1). According to Cattaneo et al. (2017), higher order polynomials tend to produce unreliable results near boundary points. Furthermore, we do not want to overfit the data, and in a nonparametric setting, a linear model seems to provide a good tradeoff between simplicity, precision and stability (Cattaneo et al., 2017).

BANDWIDTH

Findings are often sensitive to the width on each side of the threshold, and the choice of bandwidth is therefore highly influential on estimation and interpretation (Cattaneo et al., 2017). The chosen functional form affects the optimal width, as the correctly specified model changes with respect to the structure of the data. Because the bandwidth influences the amount of observations included in the analysis, there is a bias-variance trade-off when choosing the optimal size. A variety of ad-hoc approaches for choosing bandwidth are used in practice, such as standard plug-in and cross-validation methods. These do not always yield optimal widths in practice, and data-driven methods are often recommended (Imbens and Kalyanaraman, 2012). A method proposed by Imbens and Kalyanaraman (2012) is to minimise the mean squared error (MSE) of the local polynomial RD point estimator (Cattaneo et al., 2017). As in most bandwidth selection methods, the MSE-optimal bandwidth seeks to optimise the variance-bias trade-off, and the estimator is fully data-driven. In terms of the mean squared error, when using an MSE-optimal point estimation approach, the estimator is optimal, which makes it desirable for empirical work (Cattaneo and Vazquez-Bare, 2016). We therefore rely on this approach to find the bandwidth,

which we use as a baseline measure. However, as the MSE-optimal bandwidth selector is a strictly data-driven method, important information that relates to the nature of the data may be lost. In the robustness tests section, we therefore check whether estimates and statistical significance change when applying different bandwidths.

CLUSTERING

Groups of observations in the sample are likely to be affected by the same phenomena, and we therefore have to account for the fact that all observations in the dataset are not unrelated. Units sharing observable characteristics are likely to share unobservable characteristics, and this would lead the regression disturbances to be correlated (Moulton, 1990). We therefore use clustering to account for the fact that groups of observations may be spatially correlated (Anderson et al., 2016). Failure to control for the within-cluster error correlation can lead to misleading standard errors, which typically become too small (Cameron and Miller, 2015). Because we use a design where municipalities rather than units are assigned to treatment, the treatment variable will be perfectly correlated within municipalities (Abadie et al., 2017). Therefore, we can assume that model errors for observations in the same municipality are correlated, while model errors for observations in different municipalities are uncorrelated (Cameron and Miller, 2015). This reasoning makes it necessary to cluster at the municipality level, but because this way of clustering is strict and causes large standard errors, we will in addition consider a different dependence model.

To overcome some of the strictness when clustering by municipalities, we attempt to apply a more complicated spatial model of dependence between observations. Conley (1999) presents such a model where economic distance between units is used to model the dependence between them. In our case, proximity to infrastructure and cities can be considered as economic distances, as they influence the amount of deforestation through accessibility to markets (Burgess, 1993) and transportation costs of forestry and agricultural products (Pfaff, 1999). Proximity to these factors will be correlated at physically adjacent locations (Conley, 1999). Additionally, economic distances can in our setting be related to climate and geographical conditions. As deforestation in locations that are spatially separated will be affected differently by factors such as weather and soil quality, there will be dependence between grid-cells lying in areas with similar conditions.

Based on these arguments, it seems reasonable to use physical distances to establish dependency between units. Because of technical issues when implementing *Conley's OLS spatial HAC*

estimator, we create a new and alternative variable for clustering. This variable groups together grid-cells lying in a 1km range, given that they belong to the same municipality. Based on this variable, we assume that observations lying in the same municipality, and in a narrow spatial distance from each other have correlated unobserved characteristics. Together with standard errors clustered by municipality and standard errors without clustering, we will report results with standard errors clustered by this alternative measure, which we will refer to as 1km bins.

5.3 THREATS TO IDENTIFICATION

For our empirical results to be internally valid and for causal interpretation to be made, certain conditions must be met. The continuity assumption discussed earlier is the most vital for our RD analysis, and we will in the following section assess additional threats to our identification strategy. Some of these are common for all RD designs, and some are unique for GRD designs.

COMPOUND TREATMENT

An important challenge for our empirical strategy is that of compound treatment, which typically arises in GRD designs. In addition to corruption, Brazilian municipality borders may represent several treatments that occur simultaneously, making it challenging to isolate its effect on deforestation (Keele and Titiunik, 2016). However, as argued in the literature review, we therefore use corruption as a proxy for the overall institutional quality. Hence, a border does not merely symbolise a change in corruption, but also other institutional characteristics, which resolves the complication of compound treatment (Keele and Titiunik, 2016). Furthermore, as we are only considering one mayoral term, we ensure there is no change in the municipal government, and hence institutional factors will remain relatively stable in the entire period. However, specific policies that are directly related to deforestation, such as conservation zones are included as a control variable in our estimations.

MANIPULATION OF THE SCORE

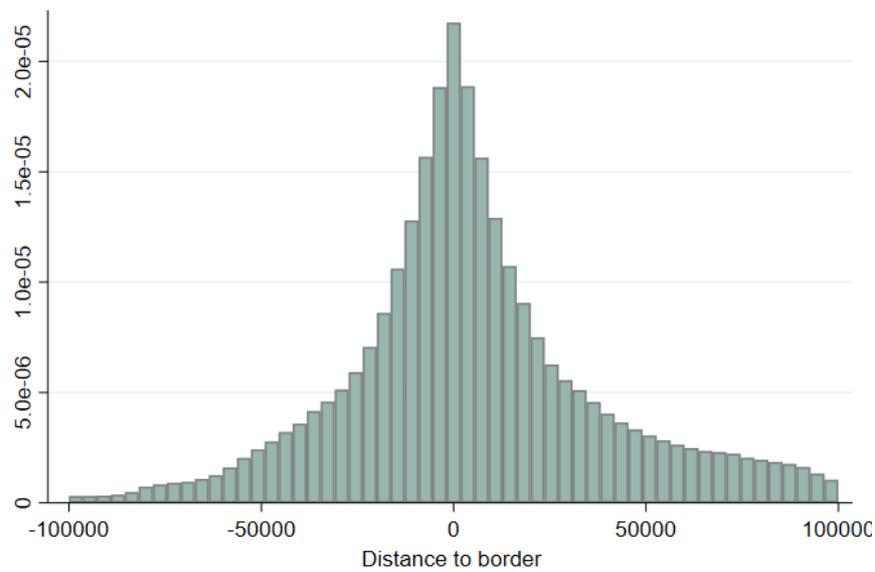
Another common concern in RD is manipulation of the score. If there is some benefit associated with the treatment and individuals can influence whether they receive it or not, we can expect the units on one side of the cut-off to be systematically different from units on the opposite side (Lee and Lemieux, 2010). However, the corruption score is based on past incidences (Brollo et al., 2013), which gives municipalities limited power to manipulate the evidence. Furthermore, as Brazilian municipalities are being audited quickly after they are selected in the fully randomised

national lottery (Ferraz and Finan, 2011), it gives them little ability to prepare or suddenly influence their corruption. The concern of manipulation in our setting can rather be related to the actual auditing process (Ferraz and Finan, 2008). If the auditors themselves are corrupt, and municipalities are able to bribe them, this would make our data unreliable. According to Ferraz and Finan (2008), these concerns are not likely to be warranted, as the auditors are well-trained, have implemented a detailed and fairly consistent procedure and earn highly competitive wages. They are also less prone to collusion with local governments as they come unexpectedly and are only there for a very short period of time. There are also no reported incidents where auditors have been offered a bribe, and as the recorded corruption levels are balanced, it can be deemed unlikely that bribes have taken place (Ferraz and Finan, 2008).

SORTING AROUND THE CUT-OFF

A special case of manipulation in a GRD setting is the ability for units to sort precisely around the border (Keele and Titiunik, 2014). This is a significant concern when dealing with mobile units of observation, such as humans, and if being on one side of a border is preferable to certain individuals. Then, units may sort around the border and discontinuities in variables other than the treatment will arise. However, when using grid-cells as observations where land and climate is naturally affecting the characteristics of the units, we assume that there is no possibility to sort around the border. The only way to achieve this would be to relocate the borders. However, the municipal borders were created before the mayoral term of 2001-2004, and borders are therefore predetermined. Nevertheless, to verify that there is neither manipulation nor sorting around the cut-off, we inspect the distribution of the data. The density of the running variable should be the same on each side of the threshold. In figure 3, the amount of observations is illustrated against the running variable using a histogram, and we do not observe any difference in density between the two sides of the cut-off.

FIGURE 3 – DENSITY OF THE RUNNING VARIABLE



Notes: Histogram illustrating the number of observations against the running variable. Observations lying outside a 100,000 meter bandwidth are excluded. Width of the bins are approximately 3,500 meters.

SENSITIVITY TO PERIPHERAL OBSERVATIONS

Units located very close to municipal borders, peripheral observations, may be subject to special issues due to their proximity to the border. Firstly, as few borders in the Legal Amazon follow the lines of the cell-grid that our dataset is based on, there will be many observations where the border cuts through a cell. Hence, though each cell is allocated to one municipality, some will in fact be partly located in the neighbouring municipality. Furthermore, borders in forest areas are not easy to observe as there are no fences or the like that follow and visualise the borders. Therefore, deforestation occurring close to a border may affect areas in the neighbouring municipality, intentionally or unintentionally. Furthermore, effects stemming from institutions may vary systematically with distance from the border (Cust and Harding, 2014; Turner et al., 2014). In addition to being affected by the institutional quality of its own municipality, peripheral units are likely to be exposed to the institutional quality of the neighbouring municipality. Cells further inside a given municipality, interior observations, are relatively more affected by the institutional quality of their own municipality. Hence, cross-border spillover effects may limit the impact of institutional quality on peripheral observations, and it is not fully realised until we are at some distance inside the municipality (Turner et al., 2014).

Therefore, to avoid disturbances close to the cut-off, we exclude peripheral units. This is referred to as a *donut hole* around the discontinuity (Cattaneo et al., 2017; Dahl et al., 2014), or *thicker borders* in specific relation to GRD studies (Cust and Harding, 2014). With regards to determining the thickness of the border, theory does not provide any precise guidance, but it is suggested that external effects are expected to decline by the mile (1.6 km) (Turner et al., 2014). We will exclude observations lying within 500 meters on each side of the cut-off in our main estimations, but increase the border thickness to check for disturbances close to the cut-off in the robustness tests section. Furthermore, we will also look at differences between peripheral and interior observations inside the same municipality.

NAIVE APPROACH TO DISTANCE

Another concern in the GRD design relates to the value of the running variable assigned to each observation. When this is defined as distance to the closest border, it is not possible to know the exact geographic location of the observation. This means that observations lying in different locations, but that have the same distance to the border will be treated in the same way, and important heterogeneity may be masked (Keele and Titiunik, 2014). Furthermore, as the running variable only represents the distance to a border and not the position along it, the problem will be exacerbated when the border is longer. The measure of distance will in this case ignore the spatial nature of geographic locations, something Keele and Titiunik (2014) refer to as a *naive* approach to distance.

By studying municipalities, we are focusing on Brazilian jurisdictions at the lowermost level, and hence looking at bordered areas of the smallest size possible. Furthermore, as the Legal Amazon is divided into around 600 municipalities (Cisneros et al., 2013), this entails that there are many relatively small-sized municipalities in the region. Furthermore, we concentrate our analysis on a limited share of any given municipality's border, namely where there is a counterfactual on the other side. Thus, it can be assumed that the borders in our analysis are not excessively long, compared to what it could have been if this was a cross-country study. Albeit it cannot be completely ruled out, it can be argued that the *naive* approach to distance can be used without having too much heterogeneity along the border.

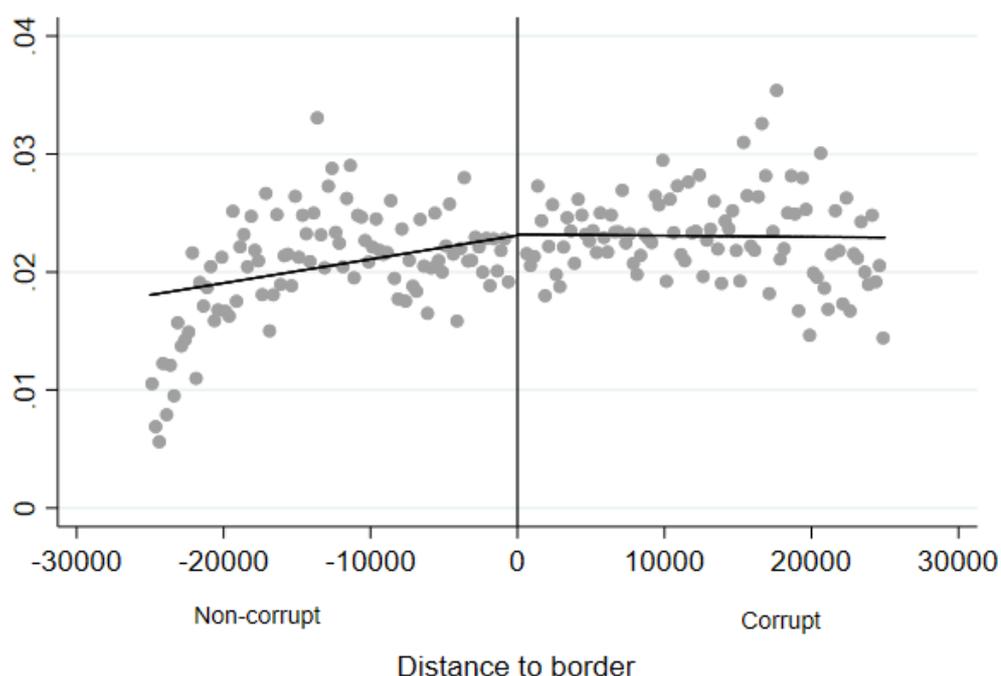
6. FINDINGS AND DISCUSSION: MAIN SAMPLE

SAMPLE

6.1 GRAPHICAL PRESENTATION

Before turning to the econometric estimates, we present our data graphically by plotting deforestation against the running variable. An advantage of the RD design is that it offers visual illustration of whether there is a discontinuity in deforestation at the municipal border (Cust and Harding, 2014). Furthermore, from the graphical presentation we are provided with an impression of the functional form of the relationship between deforestation and institutional quality. By inspecting the plot, we can also ensure that there are no unexpected discontinuities at other points than at the border (Lee and Lemieux, 2010).

FIGURE 4 – AMOUNT OF DEFORESTATION: MAIN SAMPLE



Notes: Average deforestation plotted against the running variable. Bin size of 250 meters. Bandwidth of 25,000 meters. Observations within 500 meters to the border are excluded. Linear regression functions.

Figure 4 is created by dividing the running variable into 100 evenly spaced bins on each side of the cut-off. Each bin contains observations lying within a range of 250 meters, and the average

amount of deforestation within each bin is calculated and plotted against its middle-point. We make sure that no bin is overlapping the border by excluding observations lying closer than 500 meters to the cut-off. Furthermore, we use all observations lying within 25,000 meters of the threshold, on either side. We use a linear regression model to fit the data, and by creating a separate regression line for each side of the cut-off we allow for different slopes and intercepts.

Figure 4 illustrates that there is no clear discontinuity when moving from the control group to the treatment group. We cannot detect any systematic difference in deforestation between non-corrupt and corrupt municipalities, as the dots on both sides of the cut-off are lying in the same range. This means that from inspecting the graph, it does not seem likely that lower institutional quality leads to increased deforestation.⁷ Furthermore, the dots do not reveal any specific nonlinearities, and a linear regression model therefore seems to be appropriate when estimating the relationship. Note that the regression line on the left hand side has a slightly steeper slope than the one to the right, as the observations on the left endpoint have on average lower values of deforestation. This may be an indication of lower deforestation when moving away from a corrupt border, but to establish causality we have to apply econometric estimation.

6.2 ECONOMETRIC ESTIMATION

From the graphical illustration above we do not observe a discontinuity at the threshold. However, to verify this formally, we estimate the relationship using a regression equation. The dependent variable is measured as the proportion of a grid-cell deforested, accumulated over the period 2002-2004, and this is what we use in all regressions unless otherwise specified. Estimates are presented through four econometric models, all based on equation (1), but containing different covariates; conservation zone, no forest and pair fixed effects. We present p-values based on standard errors clustered by municipality, 1km bins and without clustering. Because we choose to implement a nonparametric RD approach using a linear regression, we have to find a bandwidth that fits this model specification. We use the MSE-optimal bandwidth selector proposed by Imbens and Kalyanaraman (2012), and with this method we get a bandwidth of approximately 12,000 meters.

⁷ In appendix 2 we provide a satellite image that zooms in on a specific border between two municipalities in Pará. The municipality to the left, Altamira, is corrupt while Sao Felix, to the right, is non-corrupt. We cannot detect a difference in forest-cover between the two areas, which illustrates that not all borders separating low from high institutional quality indicates a change in the amount deforested.

We use thick borders (Cust and Harding, 2014) and exclude observations lying within 500 meters on each side of the cut-off. Our main RD estimation results are presented in table 2.

TABLE 2 – MAIN RD REGRESSION RESULTS

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.00115 (0.845) ^I (0.766) ^{II} (0.289) ^{III}	0.000692 (0.903) ^I (0.852) ^{II} (0.522) ^{III}	0.000984 (0.639) ^I (0.565) ^{II} (0.342) ^{III}	0.000834 (0.690) ^I (0.627) ^{II} (0.421) ^{III}
Running var.	-0.000000143 (0.662) ^I	3.21e-08 (0.919) ^I	-0.000000239 (0.368) ^I	-8.68e-08 (0.743) ^I
Running var. × corruption	0.000000337 (0.478) ^I	4.41e-08 (0.922) ^I	0.000000479 (0.236) ^I	0.000000267 (0.500) ^I
Conservation zone		-0.0252 (0.000) ^I		-0.00443 (0.025) ^I
No forest		-0.0280 (0.000) ^I		-0.0251 (0.000) ^I
Observations	141,145	141,145	141,145	141,145
Bandwidth (meters)	12,000	12,000	12,000	12,000
Border thickness (meters)	500	500	500	500
R-squared	0.000	0.014	0.095	0.097
Pair fixed effects	No	No	Yes	Yes

Notes: RD estimates. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. In column (1) deforestation is regressed on corruption, the running variable and the interaction term. Column (2) includes covariates for conservation zone and a measure of area without forest. (3) and (4) uses pair fixed effects, where column (3) in addition includes the two covariates. A linear regression model is applied.

The first row in table 2 presents the corruption coefficient, which is defined as the change in average deforestation when crossing a border, and this is what we are interested in. The coefficient is positive across all regressions, which is an indication of higher average deforestation levels in municipalities with poor institutional quality. Specifically, when considering all covariates, the regression results imply that when moving from an observation with good institutional quality to one with worse, the proportion of a grid-cell deforested increases on average with 0.0834%.

However, none of the estimates are statistically significant, even when standard errors are not clustered, and therefore we cannot establish inference.

The conservation zone coefficient is negative and statistically significant at the 1%-level in regression (2), and at the 5%-level in (4), when standard errors are clustered by municipality. This means that if a cell lies in a conservation zone, everything else equal, the amount of deforestation is on average lower. In addition, when the amount of area without forest increases in a cell, deforestation on average decreases. The estimate for this effect is statistically significant at the 1%-level in regression (2) and (4). Furthermore, when controlling for pair fixed effects, the corruption coefficient becomes smaller compared to when regressing without any controls. As pairs are located in different parts of the Legal Amazon, unobserved characteristics related to specific neighbourhoods may affect the estimate, which is important to correct for.

Furthermore, as the corruption coefficient decreases when adding the covariates, the estimate found in the regression results presented in column (1) is affected by the exclusion of these controls. The coefficient of interest in this estimation can therefore not be wholly attributed to differences in institutional quality. Thus, we assume that the results in column (4), where all covariates and pair fixed effects are included, is the most correct estimation, and we will proceed with robustness tests for this model.

6.3 ROBUSTNESS TESTS

We test the robustness of our findings by running several sensitivity checks. By testing assumptions and assessing whether important metrics change in magnitude when adjusting certain factors, we get a better understanding of how robust our findings are. As mentioned, we will be comparing the results in this section to the estimates obtained in column (4) in table 2.

SENSITIVITY TO BANDWIDTH CHOICE

Choice of bandwidth is crucial when using a nonparametric RD design, and we use a data-driven method to find optimal bandwidth. In our main regression results, this provided us with a recommended bandwidth of 12,000 meters. However, excluding or adding data at the end points of our sample gives knowledge about how sensitive the estimates are to observations lying far away from the threshold. Estimates for four different bandwidths are presented in table 3.

TABLE 3 – DIFFERENT BANDWIDTHS

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.000198 (0.958) ^I (0.939) ^{II} (0.896) ^{III}	0.00103 (0.712) ^I (0.458) ^{II} (0.218) ^{III}	0.00176 (0.487) ^I (0.149) ^{II} (0.012) ^{III}	0.00284 (0.249) ^I (0.012) ^{II} (0.000) ^{III}
Observations	77,967	199,624	246,284	279,564
Bandwidth (meters)	6,000	20,000	30,000	40,000
Border thickness (meters)	500	500	500	500
R-squared	0.102	0.097	0.093	0.089
Pair fixed effects	Yes	Yes	Yes	Yes

Notes: RD estimates for different bandwidths. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. The bandwidth is 6,000 meters in column (1), 20,000 meters in column (2), 30,000 meters in (3) and 40,000 meters in (4). We control for pair fixed effects, conservation zone and a measure of area without forest in all regressions. A linear regression model is applied.

In table 3 the coefficient of interest remains positive in all regressions. With a narrow bandwidth of 6,000 meters in column (1), bias is reduced compared to the MSE-optimal width. The corruption coefficient decreases to 0.0198%, compared to 0.0834% found in the main regression (table 2). For the larger bandwidths in columns (3), (4) and (5) the estimates increase compared to that of the main regression. We do not obtain statistically significant results when standard errors are clustered by municipality. However, when the bandwidth is 40,000 meters, the results are statistically significant at the 5%-level when clustering by 1km bins, and at the 1%-level without clustering. With a bandwidth of 30,000 meters and without clustering, we get statistically significant results. In general, p-values decrease when the bandwidth expands, and this is true for both clustering approaches and without clustering.

The main take-away from this robustness test is that the estimates increase and p-values decrease with higher bandwidths. These findings imply that there is a greater difference in the amount of deforestation between units located further away from the border. A more technical explanation for this variation is that the linear model is not a good fit when bandwidth increases, as inclusion of more observations leads to a different structure of the data. Another functional form may be more appropriate when increasing the bandwidth. Furthermore, as the r-squared shrinks with more observations, it seems like the specified functional form explains the data better when the width is small. To investigate whether different models influence statistical measures, we run

regressions with higher order polynomials later in this section. Nevertheless, as stated, all estimates do remain positive and are only slightly diverging from what we obtained in the main regression. In addition, as most of the p-values stay insignificant, we can conclude that the results are robust to a change in bandwidth.

SENSITIVITY TO THICKNESS OF BORDERS

In our main estimation, we have excluded observations lying within 500 meters on either side of a municipality border. However, as theory does not provide guidance with regards to the correct thickness choice, we will experiment with several widths. By estimating the treatment effect when applying gradually thicker borders, we get an indication of whether the results are sensitive to observations near the cut-off. In table 4, we present estimates using wider borders by excluding observations lying within 1,000, 1,500, 2,000 and 2,500 meters on either side of the cut-off.

TABLE 4 – DIFFERENT THICKNESS OF BORDER

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.00129 (0.555) ^I (0.498) ^{II} (0.263) ^{III}	0.000866 (0.737) ^I (0.650) ^{II} (0.498) ^{III}	0.00123 (0.674) ^I (0.553) ^{II} (0.390) ^{III}	0.000737 (0.836) ^I (0.753) ^{II} (0.645) ^{III}
Observations	133,103	125,263	117,673	110,262
Bandwidth (meters)	12,000	12,000	12,000	12,000
Border thickness (meters)	1,000	1,500	2,000	2,500
R-squared	0.097	0.097	0.098	0.098
Pair fixed effects	Yes	Yes	Yes	Yes

Notes: RD estimates using different thickness of border. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. Thickness of the border is 1,000 meters in column (1), 1,500meters in column (2), 2,000 meters in (3) and 2,500 meters in (4). We control for pair fixed effects, conservation zone and a measure of area without forest in all regressions. A linear regression model is applied.

We note that the corruption estimates in table 4 stay positive across all regressions, and that the results are statistically insignificant. The estimates change only slightly compared to the main estimate, and in both directions when increasing the thickness of the border. Thus, there is no indication of sensitivity to units located close to a border.

DIFFERENCES WITHIN MUNICIPALITIES

As we find no significant change in the coefficient of interest when applying different thickness of border, it is unlikely that there is a great difference in deforestation between units lying closer to the threshold and interior observations. However, we want to investigate this further, and check whether cross-border spillover effects are present by comparing peripheral and interior observations. We include all grid-cells located within a 1,000 or 2,000 meter range from the boundary. Observations not yet considered will therefore be accounted for, as we now use the grid-cells that was excluded in the main regression.⁸ A cut-off is implemented at the midpoint of the range of either 1,000 or 2,000 meters, in order to compare the observations that may be affected by the neighbouring municipality with the interior ones. This means that we create a placebo threshold inside each municipality, where peripheral observations receive a value of zero, while interior observations receive a value of one. To distinguish between the effect in non-corrupt and corrupt municipalities, we run separate regressions for each of these.

TABLE 5 – PERIPHERAL AND INTERIOR OBSERVATIONS

	Deforestation			
	Non-corrupt municipalities		Corrupt municipalities	
	(1)	(2)	(3)	(4)
Location	-0.00853 (0.076) ^I (0.076) ^{II} (0.045) ^{III}	-0.00159 (0.729) ^I (0.806) ^{II} (0.604) ^{III}	0.00354 (0.459) ^I (0.459) ^{II} (0.389) ^{III}	0.00323 (0.187) ^I (0.646) ^{II} (0.292) ^{III}
Observations	8,186	15,899	8,212	15,929
Bandwidth (meters)	500	1,000	500	1,000
Border thickness (meters)	0	0	0	0
R-squared	0.019	0.017	0.019	0.017
Pair fixed effects	No	No	No	No

Notes: Estimates comparing peripheral and interior observations inside a municipality. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. Bandwidth in columns (1) and (3) is 500 meters, and in (2) and (4) the bandwidth is 1,000 meters. Location is a dummy taking the value of 0 if the observation is peripheral and 1 if it is interior. Columns (1)-(2) include non-corrupt municipalities only, and (3)-(4) include only corrupt ones. We control for conservation zone and a measure of area without forest in all regressions, but not pair fixed effects. A linear regression model is applied.

⁸ The thickness of border in the main regression is 500 meters on each side of the cut-off.

The coefficients of the location variable in table 5 describe the change in average deforestation when going from a peripheral to an interior observation in a municipality. Columns (1) and (2) indicate that deforestation in our sample decrease on average in a non-corrupt municipality when moving away from a municipal border. In corrupt municipalities the effect is positive, which means that corrupt observations located further away from a border have on average higher amounts of deforestation than the peripheral ones. These results indicate a cross-border spillover effect when we assume that corrupt municipalities deforest more on average. However, none of these estimates are statistically significant, and we cannot say that a change in deforestation when moving away from a border is caused by cross-border spillover effects.

ORDER OF POLYNOMIAL

Until now we have applied a linear model to fit the data in a nonparametric strategy. However, we want to test whether the estimates are sensitive to inclusion of higher order polynomials in the running variable. The MSE-optimal bandwidth expands in this case, which results in a greater sample size and increased efficiency. Although the two approaches are different, they should lead to the same statistical results in practice (Lee and Lemieux, 2010), and this is what we want to test. We estimate regressions applying both quadratic and cubic functions, presented in table 6.

TABLE 6 – HIGHER ORDER POLYNOMIALS

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.00155 (0.537) ^I (0.336) ^{II} (0.099) ^{III}	-0.000452 (0.894) ^I (0.860) ^{II} (0.762) ^{III}	0.00129 (0.708) ^I (0.451) ^{II} (0.198) ^{III}	-0.000374 (0.918) ^I (0.888) ^{II} (0.816) ^{III}
Observations	165,885	165,885	225,601	225,601
Bandwidth (meters)	15,000	15,000	25,000	25,000
Border thickness (meters)	500	500	500	500
R-squared	0.097	0.097	0.096	0.096
Order of polynomial	2	2	3	3
Interaction terms	No	Yes	No	Yes
Pair fixed effects	Yes	Yes	Yes	Yes

Notes: RD estimates using different order polynomials of the running variable. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. Column (1) includes a quadratic running variable term and column (2) adds interaction terms. Column (3) includes cubic polynomials and column (4) adds interaction terms. Bandwidths are found by using the MSE-optimal selector. We control for pair fixed effects, conservation zone and a measure of area without forest in all regressions.

When using second or third order polynomials without interaction terms, in columns (1) and (3), the estimates stay similar to what we find in our main regression. However, when including interaction terms in columns (2) and (4), the estimates become negative. The interaction terms allow for a different slope on each side of the cut-off, and this causes the negative coefficient. As these estimates are highly insignificant with p-values above 0.7 we cannot draw any conclusions from these results. From the graphical presentation we cannot identify patterns that necessarily entail different slopes on each side of the cut-off. The fact that r-squared stays the same when including interaction terms, indicates that one model is not better than the other with regards to explaining the data. The negative estimates only prove that the results are sensitive to order of polynomial, which strengthens the finding of no causal effect of corruption on deforestation.

RELATIVE MEASURE OF DEFORESTATION

We will now test whether the coefficient of interest and statistical significance change when using an alternative dependent variable. This is created by dividing accumulated deforestation on the amount of forest observed in a cell at the beginning of the period of analysis. By using this variable, we look at relative variations in deforestation and control for potential heterogeneity in initial forest cover (Assunção et al., 2015). We regress relative deforestation on all the independent variables in table 7.

TABLE 7 – RELATIVE MEASURE OF DEFORESTATION

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.000868 (0.942) ^I (0.902) ^{II} (0.623) ^{III}	0.000328 (0.978) ^I (0.962) ^{II} (0.852) ^{III}	0.000382 (0.930) ^I (0.891) ^{II} (0.820) ^{III}	0.000282 (0.947) ^I (0.919) ^{II} (0.867) ^{III}
Observations	158,024	158,024	158,024	158,024
Bandwidth (meters)	14,000	14,000	14,000	14,000
Border thickness (meters)	500	500	500	500
R-squared	0.000	0.011	0.098	0.099
Pair-specific effects	No	No	Yes	Yes

Notes: RD estimates using relative deforestation. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. In column (1) deforestation is regressed on corruption, the running variable and the interaction term only. Column (2) includes covariates for conservation zone and a measure of area without forest. (3) and (4) use pair fixed effects, where column (3) in addition includes the two covariates. A linear regression model is applied.

Table 7 presents the same regression equations as in table 2, except that the dependent variable now is the relative measure of deforestation. In these results, the coefficient of interest stays positive, but shrinks compared to the main regressions. The estimate in column (4) indicates that the share of initial forest that is deforested increases on average with 0.028% when moving from a non-corrupt to a corrupt municipality. The p-values are high, and as before, we conclude that there is not a causal relationship between deforestation and institutional quality.

POPULATION AND MUNICIPALITY SIZE AS COVARIATES

Population growth has been frequently cited as a prominent determinant of deforestation, especially in the early literature on deforestation drivers. However, more recent literature shows that this relationship has been overemphasised, and that other factors related to the state of the economy are closer linked to deforestation. In appendix 3, column (1) and (2), we control for population size⁹, and show that that it does not affect our corruption coefficient, and that the variable is insignificant in the regressions. This implies that our results are in line with the newer literature stating that alternative factors have a greater influence on deforestation, and that population size is not a main driver.

In columns (3) and (4) in appendix 3, we control for the area size, approximated by the number of observations, of each municipality. The reason for taking this into consideration relates to the benefits associated with economies of scale. When a municipal area is large, there may be greater benefits of deforestation than when the area is small. In addition, the length of a border could be correlated with the size of a municipality (Cust and Harding, 2014), and controlling for this may cope with the problem of *naïve* distance. When including a covariate for municipality size and using pair fixed effects, results remain statistically insignificant, and the estimate decreases only slightly. This implies that municipality size is not an influential factor in our model.

PLACEBO CUT-OFF

The regression function should be continuous across the threshold in absence of treatment (Cattaneo et al., 2017). It is difficult to test for this assumption at the cut-off, but by looking at points further away, we can investigate whether the continuity assumption holds at different levels

⁹ The population variable used in our regression is based on number of people living in a municipality in 2004.

of the running variable. To verify that there is no discontinuity at other points in the sample, we therefore implement various placebo cut-offs (Cattaneo et al., 2017). This is done by analysing observations on each side of the threshold separately, where an artificial cut-off is implemented by subtracting the desired cut-off value from the running variable. Placebo treatment will then be assigned to those observations above the new threshold, even though all units analysed have the same level of corruption. We choose to implement placebo cut-offs where the running variable is -3,000, -6,000, 3,000 and 6,000 meters. When interpreting the regression results with placebo cut-off (appendix 4), our main interest is the statistical significance. The p-values are high, which is an indication of continuity in the regression function at the specified cut-offs. Thus, we can assume that the regression model is continuous at points further away from the threshold. This gives more credibility to our RD design, verifying that there is no jump in the level of deforestation at any other points than at the border.

6.4 INTERPRETATION

From our econometric estimations, we find indications of higher average deforestation levels in municipalities with poor institutional quality. However, none of our coefficients are statistically significant, which means that we cannot establish inference. When inspecting the graphical evidence, there is also no sign of a clear discontinuity. Furthermore, the robustness tests verify this. Hence, we do not have causal evidence that deforestation is an outcome of poor institutional quality.

Our findings are inconsistent with the literature stating that institutions are central drivers of deforestation (Asher, 1999; Assunção et al., 2015; Bhattarai and Hammig, 2004; Barbier, 2004; Barbier and Burgess, 2001; Burgess et al., 2012; Geist and Lambin, 2002; Gibson et al., 2000; Hargrave and Kis-Katos, 2013; Umemiya et al., 2010). This is an unexpected finding, considering the high levels of deforestation and corruption in Brazilian municipalities, combined with the fact that the Legal Amazon forest is controlled by local agents, and the many mechanisms through which poor institutional quality likely leads to increased deforestation. Indeed, according to Kolstad and Søreide (2009), as 80% of the deforestation in Brazil is illegal, it seems like this would be impossible without substantial corruption taking place.

However, much of the literature is focusing on the theoretical notion of institutional quality having strong effects on deforestation outcomes. On the other hand, empirical literature on the impact of corruption and other institutional factors is in its infancy and important questions regarding the causal relations remain unanswered (Angelsen, 2009; Bhattarai and Hammig, 2004; Cisneros et al., 2013; Mendes and Porto Jr., 2012; Pailler, 2016). Furthermore, most existing empirical evidence is based on cross-country analysis, utilising inadequate corruption data.¹⁰ Barrett et al. (2006) put forward firm warnings against establishing inference based on such studies. There are strong arguments for redirecting focus to within-country analysis, but only a few studies have done so with regards to the Legal Amazon. To our knowledge, there are three studies looking at this issue, but all with different angles (Cisneros et al., 2013; Mendes and Porto Jr., 2012; Pailler, 2016). Hence, there are very few studies that are directly comparable to ours. Moreover, as we account for more unobserved factors by conducting a within-country analysis, utilising an objective corruption measure and a credible econometric strategy, we have ensured that our findings are internally valid.

THE LIMITED ROLE OF INSTITUTIONS

According to our findings, deforestation occurs regardless of the level of institutional quality in Brazilian municipalities. Hence, the correlation found in the literature can actually be a reflection of something else, and deforestation is probably more driven by other unobserved factors. This is an argument for downplaying the role of institutional quality (Michalopoulos and Papaioannou, 2013) as a driver of deforestation, and instead focus on the importance of alternative factors. As mentioned in the identification section, geographical features and economic aspects are likely to be influential causes of deforestation. Furthermore, these factors may affect municipalities regardless of their institutional quality, as they are not necessarily impeded by municipality boundaries, but affect larger areas.

LAX REGULATIONS AND ACCEPTANCE FOR DEFORESTATION

A particular explanation for the average non-effect of institutional quality on deforestation could be the lack of strict regulations in the period of our study. As we are using data from the early 2000s, we are analysing a period with some of the highest deforestation rates seen throughout

¹⁰ See the literature review for limitations of the existing literature.

history, which led to a subsequent intensification of government actions (Castelo, 2015). Although Brazil had complex forest laws before the 2000s, and modern forest legislation began as early as in 1965, these were modest and not always effective (Bauch et al., 2009). In 2004 however, a major policy; the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon, was implemented. Assunção et al. (2012) describe 2004 and 2008 as turning points in the history of conservation policies aimed at preventing deforestation in the Amazon. They show that the implementation of forest policies in both these years were successful in terms of reducing deforestation rates. In appendix 5 we present a graph illustrating the turning points, and that these coincide with a decrease in the deforestation rates. Based on the history of forest laws in the Legal Amazon, we can assume that deforestation was not very discountenanced in the early 2000s compared to the years following 2004. As there were less strict policies in place during the years of our study, corruption may have been less necessary in order to engage in deforestation. Our results may therefore reflect the lack of need to act illegally, and that deforestation between 2002 and 2004 was more accepted. Being corrupt would therefore not necessarily lead to higher deforestation rates, as forest clearings were to a less extent illegal and detested actions.

CROSS-BORDER SPILLOVER EFFECTS

Furthermore, an important consideration is the possibility of spillover effects across municipal borders, and especially whether areas located near borders are affected by the institutional quality in the neighbouring municipality (Turner et al., 2014). Compared to national borders, municipal borders may not provide an as strict divide between the institutional reach of the neighbouring jurisdictions. For example, consider that a corrupt mayor facilitates substantial deforestation activity in its municipality and that this brings agricultural activity, infrastructure and economic growth to the area. Agents in the neighbouring, non-corrupt municipality will be incentivised to engage in deforestation in order to benefit from the economic activity occurring in the other municipality. In this case, low institutional quality in a given municipality may lead to deforestation in a neighbouring municipality with better institutional quality. However, an agent located further inside the non-corrupt municipality may not be as affected by the deforestation activities occurring in the neighbouring municipality, which means that institutional quality may not have the same effect on all areas lying within its borders (Turner et al., 2014).

Consider another context, where a sawmill is lying close to a border, and that this sawmill is conveniently located to all loggers within a given radius from it. A logger that happens to live on

the other side of the border and wants to do business and sell timber to the sawmill, will not necessarily be constrained by the fact that there is a border between them. Hence, there may be an increase in deforestation in one municipality caused by something happening in the neighbouring municipality. Note that this had nothing to do with the institutional quality in either of the municipalities, as this is an outcome caused purely by economic factors.

In the robustness tests section, we find indications of different levels of deforestation between peripheral and interior observations within the same municipality. This means that there could be important cross-border spillover effects in Brazilian municipalities affecting deforestation. As the estimates are not statistically significant, however, we do not have evidence for this mechanism.

REVERSE CAUSALITY

Though we cannot assert that there is a causal relation between institutional quality and deforestation, high corruption and deforestation levels tend to occur simultaneously. A potential reason for this is that instead of causing deforestation, poor institutional quality may be an outcome of deforestation activities. It is argued that institutional quality is an equilibrium outcome (Acemoglu et al., 2005), likely to be affected by natural resources (Cust and Harding, 2014). This relates to the notion of the natural resource curse, where the logic is that natural resources foster rent-seeking and deteriorating institutional quality (Kolstad and Søreide, 2009; Ross, 1999). As mentioned, forests possess certain characteristics that make them a potential breeding ground for illegal and corrupt activities (FAO, 2001). For instance, forest activities often revolve around large and remote areas, far from public scrutiny and government agencies, where local politicians are granted broad discretionary power to make decisions on highly subjective matters. Thus, the prospects of public officers to enhance their private wealth by supporting the overharvesting of forests are great (Barrett et al., 2006; Søreide, 2007). For example, public officials may realise that they can acquire substantial amounts of rents by requiring compensation for facilitating illegal or unsustainable logging (Ascher, 1999; Fearnside, 2003). Hence, a mayor that is not necessarily *ex ante* corrupt, may become involved in corrupt activities because of the deforestation opportunities in the municipality. As such, there is a strong possibility for reverse causality in our study. Additionally, there is a high probability of joint endogeneity, with feedback channels going in both directions, leading to simultaneous outcomes with regards to institutional quality and deforestation.

7. FINDINGS AND DISCUSSION: EXTENDED SAMPLE

We have argued that our empirical strategy relies on the assumption that control and treatment units are similar in characteristics other than institutional quality, which makes them comparable and provides a basis for identifying a causal relationship. We have done so by constructing a sample consisting of observations lying geographically close to each other, and where the border between municipalities only indicates a change in institutional quality.

However, the sample analysed until now is extracted from a larger dataset that also contains municipal pairs where both municipalities have the same corruption level. Hence, these additional pairs do not exhibit any discontinuity in institutional quality at their shared border. Furthermore, as corrupt and non-corrupt municipalities in this case are not necessarily spatially adjacent, the continuity assumption mentioned in the identification strategy no longer holds for this sample. Hence, we do not have as ideal counterfactuals as in the main sample. Nevertheless, we will investigate graphical presentations, econometric estimates and descriptive data for these municipalities, as it may provide valuable insight into the tendencies of the relationship between institutional quality and deforestation in the Brazilian Amazon. Before turning to the graphical presentation, we present a descriptive summary and the density of the running variable for the extended sample (appendix 6).

TABLE 8 – DESCRIPTIVE SUMMARY OF EXTENDED SAMPLE

	Extended sample			Total
	Non-corrupt Non-corrupt	Corrupt Corrupt	Corrupt Non-corrupt	
Observations	333 714	73 169	377 910	784 793
Pairs	83	35	108	226

Notes: Number of observations and municipal pairs in the extended sample. The first column counts all pairs where both municipalities in the pair are non-corrupt, the second column counts pairs where both are corrupt, and the third column counts pairs where there is a discontinuity in corruption at the border (this is the main sample).

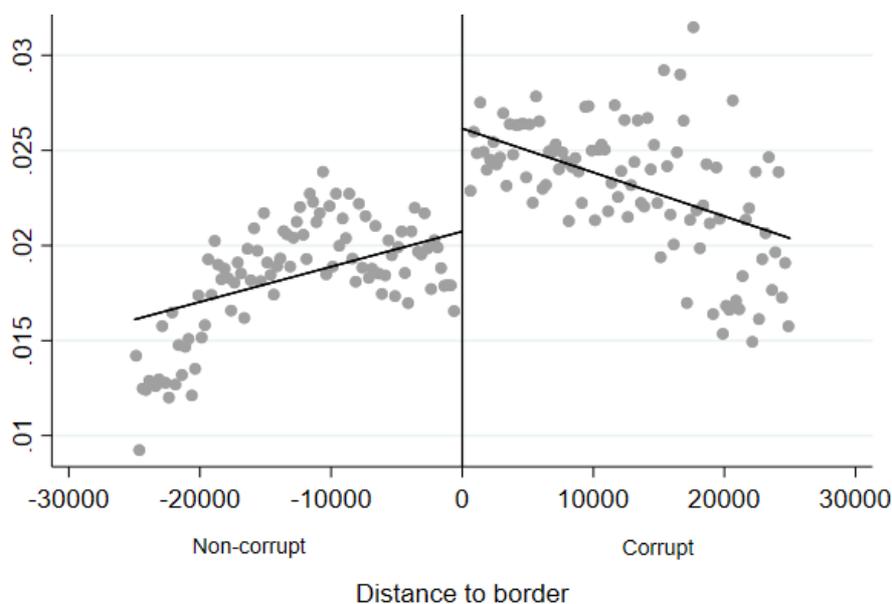
In the extended sample, there is a total of 784,793 observations, making it approximately twice as large as the main sample, which improves the statistical power of the econometric estimations. Furthermore, when looking at municipal pairs without a discontinuity in corruption, there is a substantial imbalance in the number of observations of non-corrupt and corrupt pairs, which contain 333,714 and 73,169 observations respectively. However, it can be analysed as the observations are chosen exogenously.

DENSITY OF THE RUNNING VARIABLE

From the graph presented in appendix 6, we observe a greater share of data just to the left of the cut-off. According to Cattaneo et al. (2017), if there is an unexplained abrupt change in the number of observations when crossing the cut-off, the RD design will be less credible. However, this argument is related to manipulation and sorting around the threshold. We assume that municipalities were not able to influence their measure of corruption, and therefore the density will not be related to this type of sorting around the cut-off. Any change in the density when crossing the border should therefore be due to the fact that we exclude observations outside the Legal Amazon. Hence, a change in density does not make the design less valid.

7.1 GRAPHICAL PRESENTATION

FIGURE 5 – AMOUNT OF DEFORESTATION: EXTENDED SAMPLE



Notes: Average deforestation plotted against the running variable. Bin size of 250 meters. Bandwidth of 25,000 meters. Observations within 500 meters to the border are excluded. Linear regression functions.

Because a visualisation of the data provides us with a better understanding of how the sample is structured, and whether there is a discontinuity at the specified cut-off, we present the extended sample graphically in figure 5. Average amount of deforestation in evenly spaced bins is plotted against the running variable. The bandwidth, bin size and thickness of border are similar to that of figure 4, where the main sample is illustrated.

Figure 5 exhibits a jump at the border, where observations with poor institutional quality seem to have higher average levels of deforestation than the ones with better institutional quality. Comparing these results to the graph visualising average outcomes using the main sample (figure 4), the additional observations included in the extended sample seem to have larger differences in average deforestation between non-corrupt and corrupt units. However, as the continuity assumption does not hold for this sample, the clear difference between treatment and control groups cannot be directly related to institutional quality. Nevertheless, we get a good visualisation of the overall tendencies in the data, which we will explore further with econometric estimations.

7.2 ECONOMETRIC ESTIMATION

We estimate the relationship between deforestation and institutional quality using the extended sample in order to analyse the size of the jump observed in the graph. We implement regressions using three different functional forms, apply several bandwidths and test with thicker borders, all presented in table 9. Furthermore, as we are now comparing municipalities that are not spatially adjacent, the economic aspects of being located close to a city may affect the results. Therefore, in addition to the control variables included earlier, we introduce distance to cities as a new covariate in the regression displayed in column (2). As before, we use the MSE-optimal bandwidth proposed by Imbens and Kalyanaraman (2012) to find a baseline measure of the bandwidth, which in this case is 10,000 meters on each side of the threshold.

TABLE 9 – RD REGRESSIONS: EXTENDED SAMPLE

	Deforestation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Corruption	0.00257 (0.173) ^I (0.072) ^{II} (0.002) ^{III}	0.00268 (0.155) (0.061) ^{II} (0.001) ^{III}	0.00178 (0.366) ^I (0.128) ^{II} (0.004) ^{III}	0.00207 (0.289) ^I (0.047) ^{II} (0.000) ^{III}	0.00287 (0.177) ^I (0.077) ^{II} (0.002) ^{III}	0.00315 (0.247) ^I (0.083) ^{II} (0.002) ^{III}	0.00344 (0.223) ^I (0.056) ^{II} (0.001) ^{III}
Running var.	-0.000000144 (0.409) ^I	-0.000000140 (0.418)	-1.63e-08 (0.867) ^I	2.65e-08 (0.735) ^I	-0.000000121 (0.510) ^I	-0.000000345 (0.368) ^I	-0.000000587 (0.155) ^I
Running var. × corruption	5.28e-08 (0.892) ^I	5.59e-08 (0.886)	-1.38e-08 (0.942) ^I	-0.000000146 (0.284) ^I	2.68e-09 (0.995) ^I	0.000000104 (0.907) ^I	0.000000528 (0.638) ^I
Conservation zone	-0.00476 (0.013) ^I	-0.00397 (0.034)	-0.00375 (0.018) ^I	-0.00345 (0.032) ^I	-0.00447 (0.017) ^I	-0.00433 (0.012) ^I	-0.00349 (0.031) ^I
No forest	-0.0249 (0.000) ^I	-0.0252 (0.000)	-0.0250 (0.000) ^I	-0.0243 (0.000) ^I	-0.0252 (0.000) ^I	-0.0252 (0.000) ^I	-0.0246 (0.000) ^I
Distance to cities		-6.39e-08 (0.022)					
Observations	281,550	281,550	455,951	555,828	262,805	380,749	530,857
Bandwidth (meters)	10,000	10,000	20,000	30,000	10,000	15,000	27,000
Border thickness (meters)	500	500	500	500	1,000	500	500
R-squared	0.120	0.121	0.118	0.115	0.120	0.119	0.116
Order of polynomial	1	1	1	1	1	2	3
Pair-specific effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: RD estimates using the extended sample. p-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. Columns (1)-(5) present linear regressions. Column (1) exhibits the main regression. Column (2) adds an additional covariate; distance to cities. Columns (3) and (4) test different bandwidths. In column (5) thickness of border is increased compared to the other columns. Column (6) presents a quadratic regression function, while column (7) presents a cubic regression function (both with interaction terms and MSE-optimal bandwidth). We control for pair fixed effects, conservation zone and a measure of area without forest in all regressions.

As in our main regression, the corruption coefficient is positive, meaning that municipalities with poor institutional quality have, on average, more deforestation. The corruption coefficient in these regressions is however greater than what we obtained when analysing the main sample, which coincides with the indications in figure 5. Specifically, in column (1), the estimate increases from 0.0834% in the main sample to 0.257% in the extended sample. P-values in the regressions above are also on average lower than those in the main sample, but all results are still statistically insignificant when standard errors are clustered by municipality. When including higher order polynomials, the estimate increases somewhat, but seems more stable than when applying quadratic and cubic terms using the main sample. Furthermore, the results are also stable when changing the bandwidth, increasing the thickness of the borders and when including the control variable for distance to nearest city. However, we can still not conclude that there is a causal relationship between institutional quality and deforestation based on the extended sample.

Because there is more unobserved heterogeneity across the observations in the extended sample, we are hesitant to draw any conclusions from the results in table 9. When comparing the average amount of deforestation to the left with those to the right of the cut-off, we are not necessarily comparing observations that are adjacent. However, the estimates give some knowledge about differences in deforestation levels across the municipalities in the sample. Moreover, we can argue that the pair fixed effects control for some of the heterogeneity between pairs that are spatially separated. However, as it is crucial to only analyse observations close to a border with a discontinuity in treatment when using a GRD design (Galiani et al., 2017), we do not consider the extended sample valid when implementing this empirical strategy.

7.3 LOCATION OF MUNICIPALITIES

Due to the spatial separation of a significant share of treatment and control groups in the extended sample, we are concerned that there are unobserved factors related to their location. Therefore, we will now proceed with analysing the location of the municipalities that are not in pairs where there is a discontinuity in institutional quality at the shared border. If we observe a correlation between institutional quality, deforestation levels and location of municipalities, this might be essential information when analysing whether institutional quality is a driver of deforestation.

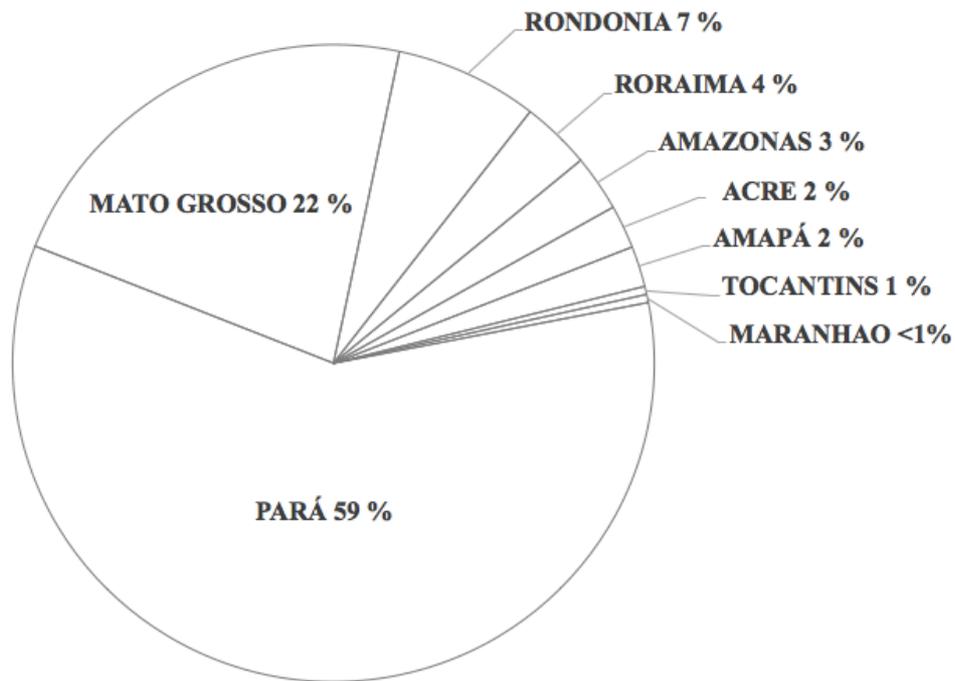
We know that some states in the Legal Amazon are known for excessively higher deforestation rates than others. In the period 2000-2005, Pará and Mato Grosso presented some of the highest deforestation levels ever seen (Mendes and Porto Jr., 2012). Furthermore, clearance is known to be concentrated in the arc of deforestation, which is located on the southern and eastern margins of the Legal Amazon, where cattle and soybean agricultural areas are constantly expanding (Malhi et al., 2008). However, what is less known is the location of Brazilian municipalities in relation to their institutional quality. In table 10 we present the location, on state level, of observations belonging to municipal pairs without a discontinuity in corruption at the shared border. Furthermore, in appendix 7, we provide maps exhibiting the state borders in the Legal Amazon and the location of the arc of deforestation.

TABLE 10 – LOCATION OF MUNICIPALITIES

State	Total	Non-corrupt		Corrupt			
	Observations	Muni- cipalities	Observations	Share of observations	Muni- cipalities	Observations	Share of observations
Acre	8 420	6	8 420	100 %	-	-	0 %
Amapá	80 533	8	76 372	95 %	3	4 161	5 %
Amazonas	74 246	13	69 464	94 %	5	4 782	6 %
Maranhao	4 264	12	2 864	67 %	7	1 400	33 %
Mato Grosso	32 454	8	14 731	45 %	12	17 723	55 %
Rondonia	17 453	5	4 972	28 %	8	12 481	72 %
Roraima	48 826	4	36 737	75 %	2	12 089	25 %
Pará	133 269	22	114 318	86 %	15	18 951	14 %
Tocantins	7 418	14	5 836	79 %	2	1 582	21 %
Total	406 883	92	333 714	-	54	73 169	-

Notes: Descriptive summary of the number of observations and municipalities residing in each of the nine Legal Amazon states, according to their corruption level. Sample used is where municipal pairs do not have a discontinuity in corruption at the shared border. Share of observations is found by dividing the number of non-corrupt or corrupt observations on the total amount of observations for each state.

FIGURE 6 – DEFORESTATION CONTRIBUTION PER STATE



Notes: Deforestation contribution per state in the period 2002-2004, based on the extended sample.

In table 10 we observe that the states of Acre, Amapá and Amazonas have virtually all of their observations classified as non-corrupt. Furthermore, by consulting figure 6, we can see that these states together are responsible for only 7% of the deforestation that occurred in the Legal Amazon in our period of analysis. Additionally, these three states are also some of the most remote, located far away from the arc of deforestation. This indicates that there could be a correlation between having high institutional quality, low levels of deforestation and being remotely located.

The same arguments can be made for Roraima, a remotely located state with 75% of its observations classified as non-corrupt and responsible for 4% of deforestation. On the other hand, Tocantins and Maranhão are contradictory examples of this logic, as they have high shares of non-corrupt observations, relatively low levels of deforestation, but are located very close to the arc of deforestation. However, this can be explained by the fact that these are states with minimal initial forest cover (appendix 8), which means that there effectively is very little to deforest.

The two states with the highest shares of corrupt observations are Mato Grosso and Rondonia, which are both located in the midst of the arc of deforestation. Mato Grosso has exhibited some

of the most excessive levels of deforestation in history (Mendes and Porto Jr., 2012), and Rondonia has been referred to as the new deforestation frontier in the Brazilian Amazon (Pedlowski et al., 2005). Furthermore, in our sample, these states contributed to almost 30% of total deforestation. This supports the notion of an inverse relationship between institutional quality and deforestation, potentially also correlated with location, and more specifically the proximity to the arc of deforestation.

On the other hand, Pará, is an example significantly contradicting this logic. Pará is by far the largest contributor to deforestation, responsible for 59% of the deforestation occurring throughout our period of analysis. Remarkably, it has 86% of its observations classified as non-corrupt. However, Pará is a large-sized state, and although parts of it lie in the arc of deforestation, a substantial share of its forest areas is still remote and relatively untouched. To detect whether there are any systematic trends between the three factors corruption, deforestation and location, we therefore deem it necessary to conduct a within-state analysis for Pará.

7.4 WITHIN-STATE ANALYSIS: PARÁ

In addition to detecting potentially valuable sub-state trends in the relationship between corruption, deforestation and location for the special case of Pará, a within-state analysis is beneficial because many important unobservable factors are naturally controlled for. Factors such as timber and meat prices as well as culture and legal issues are more stable within a state than across the entire Legal Amazon (Mendes and Porto Jr., 2012). Hence, although the continuity assumption does not hold when we include the municipal pairs without a discontinuity in corruption at the shared border, we are able to keep salient unobservable factors constant by analysing an individual state. Hence, we are now more capable of isolating the effects of corruption on deforestation. However, for comparability, we will also perform regressions using the sample where the continuity assumption holds.

In table 11, we present econometric estimations of the relationship between deforestation and institutional quality in Pará, using three different samples. In column (1) we only include pairs where the municipalities have the same level of corruption, hence there is no discontinuity at the shared border. In column (2) we only include the pairs where the municipalities have different corruption levels, hence there is always a discontinuity at the border and the continuity assumption holds. In column (3) we combine both of the samples used in column (1) and (2).

TABLE 11 – RD ESTIMATES FOR PARÁ

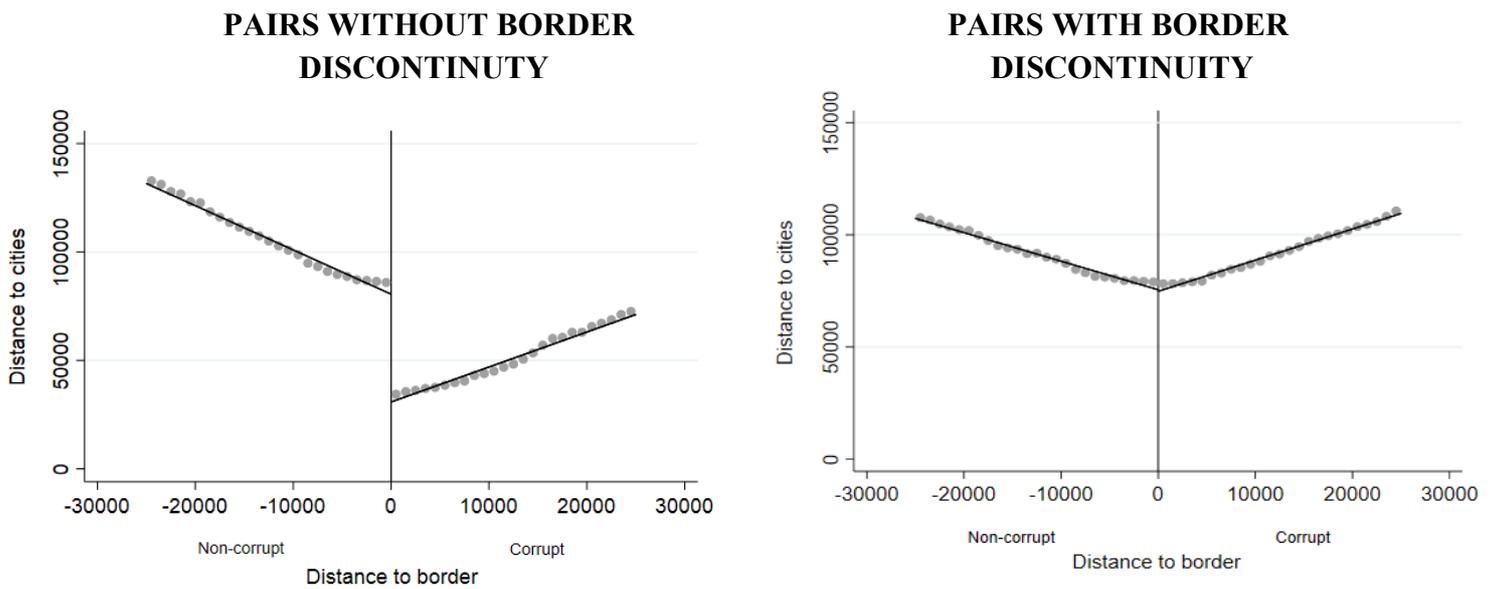
	Deforestation		
	(1) Pairs without discontinuity at shared border	(2) Main sample	(3) Extended sample
Corruption	0.0609 (0.000) ^I (0.000) ^{II} (0.000) ^{III}	-0.00166 (0.714) ^I (0.666) ^{II} (0.450) ^{III}	0.00261 (0.598) ^I (0.431) ^{II} (0.162) ^{III}
Observations	44,497	46,814	83,455
Bandwidth	10,000	10,000	9,000
Border	500	500	500
R-squared	0.102	0.094	0.100
Pair fixed effects	Yes	Yes	Yes

Notes: RD estimates for Pará. *p*-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. We control for pair fixed effects, conservation zone and a measure of area without forest in all regressions. A linear regression model is applied.

The results from the main and extended samples within Pará show, similarly to our econometric estimations based on the entire Legal Amazon, no effect of institutional quality on deforestation. However, when analysing municipal pairs without a discontinuity in institutional quality at the shared border, we find highly statistically significant results, where the corrupt observations have on average 6,1% higher levels of deforestation than the non-corrupt ones. The results remain statistically significant at the 1%-level for both types of clustering and without clustering.

Nevertheless, as the treated and control units are likely to be spatially separated in the sample only including pairs without a discontinuity at the border, we will now proceed with analysing the location of these municipalities compared to those in the main sample. In figure 7 we present the difference in distance to cities for the non-corrupt and corrupt observations within Pará, for each of the two samples.

FIGURE 7 – DISTANCE TO CITIES WITHIN PARÁ



Notes: Distance to cities are plotted against the distance from each observation to nearest municipality border. The sample including only municipal pairs without a discontinuity in corruption is presented in the graph on the left-hand side, and municipal pairs with a discontinuity in corruption is presented on the right-hand side. All observations are located in the state of Pará. 1km bins. Bandwidth of 25,000 meters.

By investigating the distance from observations to cities within Pará, we find that there is a significant discontinuity between non-corrupt and corrupt units in the sample containing municipal pairs without a discontinuity in corruption at the border. Hence, we can conclude that the non-corrupt municipalities are in more remote locations, and that the corrupt ones are lying closer to economic hubs. We also expect that these economic hubs are relatively adjacent to the arc of deforestation. Hence, similarly to what we observed for the entire Legal Amazon, there are strong indications of a correlation between deforestation, corruption and location for the municipalities within Pará.

7.5 INTERPRETATION

When analysing the extended sample for the entire Legal Amazon, we observe tendencies where municipalities with poor institutional quality have, on average, higher levels of deforestation. However, when inspecting the statistical significance, and because the continuity assumption does not hold in the extended sample, we cannot say that there is a causal relationship between deforestation and institutional quality. However, there seems to be a systematic pattern with

regards to the proximity to the arc of deforestation for the states that show an inverse relationship between deforestation and institutional quality. The major exception to this logic is Pará. However, from the within-state analysis of Pará we find statistically significant estimates of such an inverse relationship. This coincides with the study of Mendes and Porto Jr. (2012), who found no causal relationships when studying the entire Legal Amazon, but only when conducting analyses on specific states, namely Mato Grosso and Pará. Within Pará, we also find indications of corrupt municipalities being located considerably closer to economic hubs than non-corrupt municipalities. Hence, location seems to be a vital factor that is correlated with both the level of institutional quality and deforestation in Pará, as well as the in the entire Legal Amazon. There could be a simultaneous interaction between these three variables, but we cannot say for sure in which direction these relationships work. We know that deforestation is an outcome and location is a given, but not exactly how the mechanisms between these factors and institutional quality function. A possible scenario is that corruption increases as a response to acknowledging opportunities of engaging in deforestation based on location factors.

The conclusion that location is a key determinant affecting deforestation coincides with Alves (2002) who argue that deforestation is concentrated in more developed regions in Brazil, due to better infrastructure and market access. Pfaff et al. (2007) support this theory by asserting that a country with great road density is associated with higher levels of deforestation, and that neighbouring areas will be affected in a similar way. Furthermore, it has been found that deforestation is spatially autocorrelated, which means that once an area has become heavily deforested, such as the arc of deforestation, a continuous progression of forest clearings in such areas is likely to follow.

Moreover, the notion that location and distances matter is in line with the theory of economic geography put forward by Krugman (1991), who argued that the effects of location are much greater than what mainstream economics recognises. A central location is beneficial due to lower transportation costs, increased access to markets and economies of scale due to linkages and agglomeration effects, all of which are factors that influence the profitability of deforestation activities. Furthermore, if the factor in question is mobile, which forestry and agricultural products are, such effects of location are even stronger.

8. CONCLUSION

In this study, we investigate the role of institutional quality at the municipal level and its influence on deforestation in the Legal Amazon. The importance of institutions as a driver of deforestation has received little attention in the empirical literature, and we add knowledge to the topic by investigating this relationship through a geographical regression discontinuity analysis. We do so by taking advantage of high-resolution spatial data on deforestation and an objective measure of corruption as a proxy for institutional quality to create a dataset for the period 2002-2004. These high-quality datasets allow us to analyse the difference in deforestation rates between neighbouring municipalities that differ only in institutional quality. In addition, as we are concerned with unobserved characteristics that may influence our outcome, we investigate the role of location in relation to deforestation and institutional quality using a greater sample where the compared observations are not necessarily spatially adjacent.

The main finding in our analysis is inconsistent with the literature stating that institutions are key drivers of deforestation, as we find no such causal relationship. There is no significant difference in deforestation rates among municipalities with high institutional quality compared to the ones with low institutional quality in the Brazilian Amazon in the period we study. We present several robustness tests that validate these results. Although our findings do not rule out any correlation between the two factors in general, they suggest that other unobserved factors are more important in explaining deforestation, which are likely masked in the existing empirical literature. It is plausible that the regulatory environment, local economic factors and spatial interdependence of adjacent areas are important features affecting the interplay between institutional quality and deforestation.

The second important finding of our study strengthens the credibility of these presumptions. We observe that the location of municipalities plays a crucial role in determining the interaction between deforestation rates and institutional quality. We find indications of higher deforestation rates in municipalities with poor institutional quality that are located in states near the arc of deforestation. The state of Pará is the only anomaly to this notion, as it presents high levels of deforestation and seemingly good institutional quality. However, by exploiting a within-state analysis, we find highly statistically significant results of corruption leading to increased deforestation. Moreover, we find that corrupt municipalities in Pará on average are located closer

to economic hubs. Hence, for the entire Legal Amazon, it seems plausible that location is a vital underlying factor in the interplay between deforestation and institutional quality. This is reasonable, as location affects economic factors such as access to markets, transportation costs, economies of scale and agglomeration effects (Angelsen and Kaimowitz, 1999; Alves, 2002; Burgess, 1993; Hargrave and Kis-Katos, 2013; Krugman, 1991; Pfaff, 2007). We cannot, however, say anything about the direction in which the possible relationship between the three variables, deforestation, institutional quality and location, works. The one thing we can argue is that it is not as simple as institutional quality being the sole explanatory variable causing deforestation in the Legal Amazon.

Our finding of an average non-effect of institutional quality on deforestation in the Brazilian Amazon should be regarded as positive news. It means that unlike what FAO (2001) contends, issues with institutional quality do not have to be resolved before conservation efforts can have an impact. This is not to say that it would not be beneficial to tackle both problems simultaneously. Indeed, we find no evidence that poor institutional quality leads to less deforestation. Nonetheless, it means that it is worth implementing policies to conserve forests, even if the problem of institutional quality is not resolved *ex ante*.

Nevertheless, as we are analysing a sample of Brazilian municipalities in a relatively short and distinct period in the history of deforestation in the Legal Amazon, we are cautious with generalising the results of our study. There is no guarantee that the estimated response is representative of other areas or time periods. However, one can argue that it is better to identify and estimate effects, or non-effects in our case, for a specific group of the population, than to find ambiguous or even false estimates through a large cross-country study. Furthermore, as the empirical literature on this relationship is in its infancy and definite evidence is scarce, our study is adding novel insight to this insufficiently understood topic. Nevertheless, supplementary research is much needed before any grand conclusions should be drawn.

9. RECOMMENDATIONS FOR FUTURE RESEARCH

We have identified some key areas that would benefit from improvement and increase the validity of our estimations, which we recommend to future researchers who contemplate using a similar empirical strategy or the same dataset.

9.1 IMPROVED CONTROL OF UNOBSERVED FACTORS

LOCATION AND ECONOMIC FACTORS

The effect of location, and everything that entails, especially with regards to economic distances, seems to be a vital determinant that is correlated with both the level of deforestation and institutional quality. To isolate the effect of institutional quality on the level of deforestation, it is therefore necessary to properly account for characteristics associated with the areas where observations are located. We do this in section 7.4 by looking further into the state Pará, where we assume that observations within this state are more similar as many important unobservable factors are naturally controlled for. This type of within-state analysis should be replicated for the remaining Legal Amazon states.

Another approach to cope with the issue of locations, is to obtain information about characteristics of specific areas, and include them as control variables. We would in this case suggest using variables accounting for spatial autocorrelation, as units located in a neighbourhood with large amounts of deforestation seem more likely to deforest themselves, due to factors common to spatially adjacent areas such as access to markets and transportation costs. It could therefore also be beneficial to include specific distance measures, where we suggest proximity to the arc of deforestation as an additional control variable.

GEOGRAPHY

Furthermore, as we are researching the exploitation of a natural resource, an important underlying factor to consider is the natural geography (Anderson et al., 2016; Cust and Harding, 2014; Turner et al., 2014). For instance, it may be more attractive to clear out forest areas with high-quality soil for agriculture, or if the type of timber is of high economic value in the market. If these factors interact with the corruption variable, for example if illegal logging typically appears in areas

where the soil quality is high, geographical features must be accounted for. We argued that by looking at relatively narrow areas, we can assume that observations lying in different municipalities are still located in the same homogeneous response units (HRUs) and that the natural geography is thus very similar (Cuaresma et al., 2017) (Appendix 1). However, an approach that would strengthen this argument is to focus on municipal borders that are relatively straight. The idea is that nothing in nature is straight (Anderson et al., 2016). A potential discontinuity in deforestation across a straight border is more likely to be related to institutional quality than across a non-straight border, as this is probably more determined by the local natural geography, and we could have an omitted variable problem (Anderson et al., 2016). We could not obtain data on straight borders for this study, but such data can be generated through calculation, and should be incorporated in future research utilising a similar empirical strategy.

GENERAL APPROACH TO CONTROL FOR UNOBSERVED FACTORS

Instead of comparing different municipalities, where we are concerned with heterogeneity in other factors influencing deforestation, one can rather study the same municipalities over time. By observing the same unit across several periods, we can be more certain that the abovementioned and additional unobserved factors are likely to stay constant. Then, if the level of corruption changes between two periods, and there is a systematically corresponding difference in the deforestation outcome, this is likely to be a consequence of the change in institutional quality. Assuming that all other irrelevant unobserved treatments occur in two periods, these treatments from the first period can under appropriate assumptions be subtracted from the overall effect in the second period, and we have isolated the effect of the treatment of interest on the outcome Keele and Titunik (2014). However, this strategy will require conditions that specify the ways in which the institutional quality and the other unobserved factors affect deforestation over time (Keele and Titunik, 2014). When carefully considering period-specific circumstances that may affect deforestation rates, such a strategy will be fruitful and may add to the knowledge about the impact of institutional quality on deforestation.

9.2 INCREASED SAMPLE

LONGER TIME-FRAME

In relation to the point made above, analysing variation in deforestation and institutional quality over a longer time-frame is desirable also on a more general level (Umemiya et al., 2010). Both the deforestation and corruption datasets stretch back to around the year of 2000 and new data is

accumulating on a continuous basis to the present. We chose to study one mayoral term, and the period of 2002-2004 due to the amount of observations in the corruption dataset produced by Brollo et al. (2013). However, we strongly suggest future researchers to create a substantially larger corruption dataset by coding more of the CGU audits and merging this with the deforestation data, and thus obtain data for a 16-years or so period. This would improve the statistical power of the results and give a more comprehensive picture of the relationship between institutional quality and deforestation in the Legal Amazon.

Furthermore, by adding periods after the year 2004 to the analysis, we can look at whether there is a change in the results when stricter policies are in place. As mentioned, when interpreting the results from the main sample, the non-effect found in this study may be a consequence of modest forest laws during the period of analysis. Some of the major government programs for the Legal Amazon were implemented in 2004 and 2008, and it would be interesting to investigate whether the effect of poor institutional quality changes when deforestation is more regulated. In this regard, we recommend future research to take later periods into account when estimating the relationship between deforestation and institutional quality.

ADDITIONAL COUNTERFACTUALS

Another approach that would make the sample size larger, is to use an alternative method when identifying municipalities to be compared. The corruption dataset we use is based on a binary corruption score. Hence, if we observe two neighbouring municipalities with a similar score, we cannot know which one has worse institutional quality in reality. Because we are dependent on spatially adjacent municipalities with a jump in the treatment variable at the border, a binary measure of corruption leads to a smaller sample where municipal pairs without a discontinuity in corruption are excluded. However, with access to data with a more nuanced measure of institutional quality, we could place a municipality to the left or the right hand side of the cut-off based on its institutional quality relative to its neighbour (Cust and Harding, 2014). With a measure making this process possible, the dataset used for analysis would contain more pairs with an ideal counterfactual, which in turn would improve statistical power.

9.3 PREDETERMINED INSTITUTIONS

As mentioned in the discussion section, a central issue is potential feedback channels from the deforestation processes to institutional quality. To avoid such reverse causality, several studies

look at predetermined institutional quality by ensuring a time-lag between the institutional quality and deforestation observations, which means that institutions are measured before the outcome of interest (Cust and Harding, 2014; Michalopoulos and Papaioannou, 2013). In our case, when looking at the mayoral term 2001-2004, deforestation rates in the years following 2004 can be used to avoid the issue of reverse causality. However, as political terms are based on four-year periods, many municipalities changed their mayor in 2004. If one decides to focus on deforestation as an outcome of institutional quality, one must therefore account for the fact that the political administration has not remained constant in the different terms analysed.

9.4 QUALITY CHECKING THE DATA

The potential error and uncertainty of data is always critical (Grainger, 2009). This is especially important in our study as different datasets have been merged and analysed. Therefore, controlling for the quality of data and thus improve reliability of this type of analysis is highly recommended.

DEFORESTATION DATA

The data used to observe deforestation levels in the Brazilian Amazon is highly objective as it is based on satellite images of actual changes in the amount of forest. However, two things are worth noting with respect to limitations of these data. Originally, the deforestation rates were calculated manually based on visual interpretation of the satellite images (Cisneros et al., 2013), and from 2003 the interpretation became partially automatic. This means that some of our data is based on the manual approach, which may introduce more errors than if the process was fully automatic. Cloud coverage is another factor that could imply difficulties in terms of having reliable data. Because clouds can interrupt the ability of the satellite to detect land cover patterns (Assunção et al., 2015), there may be errors in the data relating to this issue. As we do not have data for cloud coverage, we have not been able to control for this, something that may be included in further research. Nevertheless, the data used on deforestation rates are the best available at the present, and should not impose significant errors in our study.

CORRUPTION DATA

Although there are several arguments for the corruption data being of high quality, stated in the empirical analysis section (Ferraz and Finan, 2008), we cannot completely eliminate the chances that the resulting dataset is an inaccurate representation of reality. Hence, finding ways to control the quality of data is extremely important.

One threat is that the lotteries are in fact not random, as audit probabilities are equal within states, but may vary across states (Zamboni and Litschig, 2013). Nevertheless, there is no evidence that the Brazilian government has intentionally influenced the probabilities of being selected, and the lotteries can be considered as-if random (Timmons and Garfias, 2015). Another concern is regarding whether the auditors discover actual levels of corruption. As such actions are illegal and detested by society, agents undertaking them are likely to take substantial steps to cover them up. Hence, there is inevitably a concern relating to the ability of local government officials to influence the content of the audit reports either by misrepresenting information or bribing the auditors (Cisneros et al., 2013). Hence, the challenge with audit data is that it represents a combination of both actual corruption and the inability to hide illegal actions from auditors (Olken and Pande, 2012). Future research utilising corruption data based on the CGU reports could benefit from a more thorough assessment of these concerns.

Furthermore, there are potential limitations relating to the coding process of the CGU reports. As the coding of the reports were done on a manual basis by Ferraz and Finan (2008; 2011), they are prone to measurement error. This potential mis-coding of the reports has been criticised as contaminating the quality of the corruption data (Cisneros et al., 2013). Furthermore, Ferraz and Finan are researchers specialising in the topic of corruption in Brazil. Although it is a significant advantage that they are experts in the field, their engagement and involvement could somehow also make them biased when coding the reports. Consequently, Timmons and Garfias (2015) argue that multiple independent coders, that will not take part in assessing the outcome data would have been preferable.

Furthermore, as Brollo et al. (2006) manually code the dataset generated by Ferraz and Finan (2008; 2011), there may be additional measurement errors affecting the quality of the data. Moreover, by assigning binary corruption scores to Brazilian municipalities, we are looking at a very crude measure where a lot of potentially crucial information about the nature and magnitude of the corruption incidents is lost along the way. Hence, it can be argued that a more nuanced corruption score, by including monetary values or increasing the scale from binary to multi-numbered, would improve the accuracy of the measure. Furthermore, it would allow for additional counterfactuals, as discussed earlier.

9.5 OTHER ISSUES TO CONSIDER

ECONOMETRIC DESIGN

Even though the RD design is arguably one of the most credible economic techniques (Calonico et al., 2016), it is not yet fully developed, and constantly evolving (Cook and Wong, 2008). Current theory concerning bandwidth choice and other important features in a RD design is not very specific, and several methods are proposed for implementation in practice. Most approaches for choosing the optimal width is based on an econometric trade-off between bias and variance, and we cannot know for certain that the optimal width calculated with the MSE-optimal approach is superior when using geographical data. It would thus be beneficial to test different econometric designs using the same data in order to check whether the results change.

Another concern related to the econometric design is the choice of clustering approach. We argue that there will be dependence between spatially adjacent observations with regards to economic distance, and present results where standard errors are clustered on both municipalities and 1km bins. However, an improvement in terms of a more complicated spatial dependence structure would be beneficial in our analysis. For future research based on geographical observations, we recommend implementing *Conley's OLS spatial HAC estimator* for optimal clustering purposes.

MUNICIPAL VS. NATIONAL BORDERS AS DISCONTINUITY IN RD

We contend that a within-study analysis of the relationship between institutional quality and deforestation is preferable to that of a cross-country study. As we stated in the literature review, there are several critical pitfalls related to omitted variables when empirically analysing the impact of institutional quality on deforestation using cross-country data (Angelsen, 2009; Barrett et al., 2006; Mendes and Porto Jr., 2012). However, when it comes to the RD design, the arguments for choosing a within-country analysis over a cross-country study becomes more questionable. Compared to national borders, municipal borders may not provide an as strict divide between the level of institutional quality of the neighbouring jurisdictions. Consequently, the issue of cross-border spillover effects may be substantial, and larger than in cross-country studies (Turner et al., 2014). As we found subtle indications of areas located near municipal borders being affected by the institutional quality in the neighbouring municipality, this issue should be further scrutinised and controlled for. However, as unobservable characteristics are more likely to be similar among units within a country, a discontinuity in the outcome at municipality borders are more likely to reflect the treatment effect than when analysing across countries.

RELIABILITY OF PROXY VARIABLES

Institutional quality is a concept which is difficult to measure, and it is not a simple task to change it into a numerical value. We use corruption as a proxy for institutional quality, as we have access to objective measures of corruption in Brazilian municipalities. However, there might be limitations related to the fact that we are using a proxy to represent the main explanatory variable. Weaknesses in this regard relates to the question of whether the corruption measure actually is a decent proxy for institutional quality, and if the corruptness of a municipality represents the overall quality of their institutions. To make sure that the quality of institutions analysed is captured when implementing a proxy, a recommendation for future research is to adopt other measures in addition, such as rule of law and efficiency of bureaucracy. These alternative variables can be added to the corruption measure, or they could be tested separately.

NAIVE DISTANCE

We have argued that the measure of distance to nearest municipality border used in this study can be considered *naive*, as it does not recognise where an observation is located along the border (Keele and Titiunik, 2014). Because we focus on Brazilian jurisdictions at the lowermost level, and assume that the borders used in the analysis are relatively short, we ignore the fact that some heterogeneity may be masked. Observations might be compared to those on the other side of the border that are slightly different in unobserved characteristics, which could cause errors in the estimates. By establishing a more specific geographical measure that takes both latitude and longitude into consideration when calculating the running variable, the issue of *naive* distance will be reduced. This is beyond the scope of our study, but we recommend implementation of a more specific geographical measure in future GRD analysis.

REFERENCES

- Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* (No. w24003). National Bureau of Economic Research.
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth, 1*, 385-472.
- Alston, L. J., Libecap, G. D., & Mueller, B. (2000). Land reform policies, the sources of violent conflict, and implications for deforestation in the Brazilian Amazon. *Journal of environmental economics and management, 39*(2), 162-188.
- Alves, D. S. (2002). Space-time dynamics of deforestation in Brazilian Amazonia. *International Journal of Remote Sensing, 23*(14), 2903-2908.
- Amazon Fund (2017). *Purposes and Management*. Retrieved from: http://www.amazonfund.gov.br/FundoAmazonia/fam/site_en/Esquerdo/Fundo/
- Andersen, L. E. (1996). The causes of deforestation in the Brazilian Amazon. *The Journal of Environment & Development, 5*(3), 309-328.
- Anderson, L. O., De Martino, S., Harding, T., Kuralbayeva, K., & Lima, A. (2016). *The effects of land use regulation on deforestation: Evidence from the Brazilian amazon* (No. 172). Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.
- Angelsen, A. (2009). *Realising REDD+: National strategy and policy options*. Cifor.
- Angelsen, A., & Kaimowitz, D. (1999). Rethinking the causes of deforestation: lessons from economic models. *The world bank research observer, 14*(1), 73-98.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton Univ Pr.
- Arima, E. Y., Barreto, P., Araújo, E., & Soares-Filho, B. (2014). Public policies can reduce tropical deforestation: Lessons and challenges from Brazil. *Land use policy, 41*, 465-473.
- Ascher, W. (1999). *Why governments waste natural resources: policy failures in developing countries*. JHU Press.
- Assunção, J., Gandour, C., & Rocha, R. (2015). Deforestation slowdown in the Brazilian Amazon/ prices or policies?. *Environment and Development Economics, 20*(6), 697-722.
- Avis, E., Ferraz, C., & Finan, F. (2016). *Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians* (No. w22443). National Bureau of Economic Research.
- Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.
- Barbier, E. B. (2004). Explaining agricultural land expansion and deforestation in developing countries. *American Journal of Agricultural Economics, 86*(5), 1347-1353.
- Barbier, E. B., Damania, R., & Léonard, D. (2005). Corruption, trade and resource conversion. *Journal of Environmental Economics and Management, 50*(2), 276-299.

Barbier, E.B., and J. C. Burgess (2001). The Economics of Tropical Deforestation, *Journal of Economic Surveys*, 15(3): 413-433.

Barrett, C. B., Gibson, C. C., Hoffman, B., & McCubbins, M. D. (2006). The complex links between governance and biodiversity. *Conservation biology*, 20(5), 1358-1366.

Bauch, S., Sills, E., Rodriguez, L. C. E., McGinley, K., & Cubbage, F. (2009). Forest policy reform in Brazil. *Journal of Forestry*, 107(3), 132-138.

den Besten, J. W., Arts, B., & Verkooijen, P. (2014). The evolution of REDD+: An analysis of discursive-institutional dynamics. *Environmental Science & Policy*, 35, 40-48.

Bhattarai, M., & Hammig, M. (2001). Institutions and the environmental Kuznets curve for deforestation: a crosscountry analysis for Latin America, Africa and Asia. *World development*, 29(6), 995-1010.

Bhattarai, M., & Hammig, M. (2004). Governance, economic policy, and the environmental Kuznets curve for natural tropical forests. *Environment and Development Economics*, 9(3), 367-382.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2), 577-599.

Brollo, F., Nannicini, T., Perotti, R., & Tabellini, G. (2013). The political resource curse. *The American Economic Review*, 103(5), 1759-1796.

Bryant, D., Nielsen, D., & Tangley, L. (1997). *The Last Frontier Forests—Ecosystems & Economies on the Edge*. Washington, DC: World Resources Institute (WRI).

Bulte, E. H., Damania, R., & Lopez, R. (2007). On the gains of committing to inefficiency: corruption, deforestation and low land productivity in Latin America. *Journal of Environmental Economics and Management*, 54(3), 277-295.

Burgess, J. C. (1993). Timber production, timber trade and tropical deforestation. *Ambio*, 136-143.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., & Sieber, S. (2012). The political economy of deforestation in the tropics. *The Quarterly Journal of Economics*, 127(4), 1707-1754.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2016). rdrobust: Software for regression discontinuity designs. *Unpublished*. Retrieved from: http://faculty.chicagobooth.edu/max.farrell/research/Calonico-Cattaneo-Farrell-Titiunik_2016_Stata.pdf.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317-372.

Castelo, T. B. (2015). Brazilian forestry legislation and to combat deforestation government policies in the amazon (Brazilian amazon). *Ambiente & Sociedade*, 18(4), 221-242.

Cattaneo, M. D., & Vazquez-Bare, G. (2016). The choice of neighborhood in regression discontinuity designs. *Observational Studies*, 2, 134-146.

Cisneros, E., Hargrave, J., & Kis-Katos, K. (2013). Unintended consequences of anti-corruption strategies/ public fiscal audits and deforestation in the Brazilian Amazon. *Unpublished working paper*.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1), 1-45.

- Cook, T. D., & Wong, V. C. (2008). Empirical tests of the validity of the regression discontinuity design. *Annales d'Economie et de Statistique*, 127-150.
- Cropper, M., & Griffiths, C. (1994). The interaction of population growth and environmental quality. *The American Economic Review*, 84(2), 250-254.
- Cuaresma, J. C., Danylo, O., Fritz, S., McCallum, I., Obersteiner, M., See, L., & Walsh, B. (2017). Economic development and forest cover: evidence from satellite data. *Scientific reports*, 7, 40678.
- Cust, J., & Harding, T. (2014). Institutions and the location of oil exploration. *OxCarre Research Paper*, 127.
- Dahl, G. B., Løken, K. V., & Mogstad, M. (2014). Peer effects in program participation. *The American Economic Review*, 104(7), 2049-2074.
- Driffield, N., Mickiewicz, T., & Temouri, Y. (2016). Ownership control of foreign affiliates: A property rights theory perspective. *Journal of World Business*, 51(6), 965-976.
- Ehrhardt-Martinez, K. (1998). Social determinants of deforestation in developing countries: a cross-national study. *Social Forces*, 77(2), 567-586.
- Eliasch, J. (2008). Climate change: financing global forests. *The Eliasch review*. Earthscan.
- Esiyok, B., & Ugur, M. (2017). A spatial regression approach to FDI in Vietnam: province-level evidence. *The Singapore Economic Review*, 62(02), 459-481.
- European Commission (2014). *Reporting greenhouse gas emissions from deforestation and forest degradation with Pan-tropical biomass maps*. Retrieved from: <https://ec.europa.eu/jrc/en/science-update/reporting-greenhouse-gas-emissions-deforestation-and-forest-degradation-pan-tropical-biomass-maps>
- FAO (2001). State of the World's Forests. Rome, Italy. *Food and Agriculture Organization of the United Nations*.
- FAO (2006). Global Forest Resources Assessment 2005. Progress towards sustainable forest management. *Food and Agriculture Organization of the United Nations*, FAO forestry paper.
- FAO (2010). Global Forest Resources Assessment 2010. Main Report. *Food and Agriculture Organization of the United Nations*, FAO forestry paper.
- FAO (2016). State of the World's Forests. Rome, Italy. *Food and Agriculture Organization of the United Nations*.
- Fearnside, P. M. (2003). Conservation policy in Brazilian Amazonia: understanding the dilemmas. *World development*, 31(5), 757-779.
- Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians- the effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2), 703-745.
- Ferraz, C., & Finan, F. (2011). Electoral accountability and corruption/ Evidence from the audits of local governments. *The American Economic Review*, 101(4), 1274-1311.
- Foley, J. A., Asner, G. P., Costa, M. H., Coe, M. T., DeFries, R., Gibbs, H. K., ... & Snyder, P. (2007). Amazonia revealed: forest degradation and loss of ecosystem goods and services in the Amazon Basin. *Frontiers in Ecology and the Environment*, 5(1), 25-32.

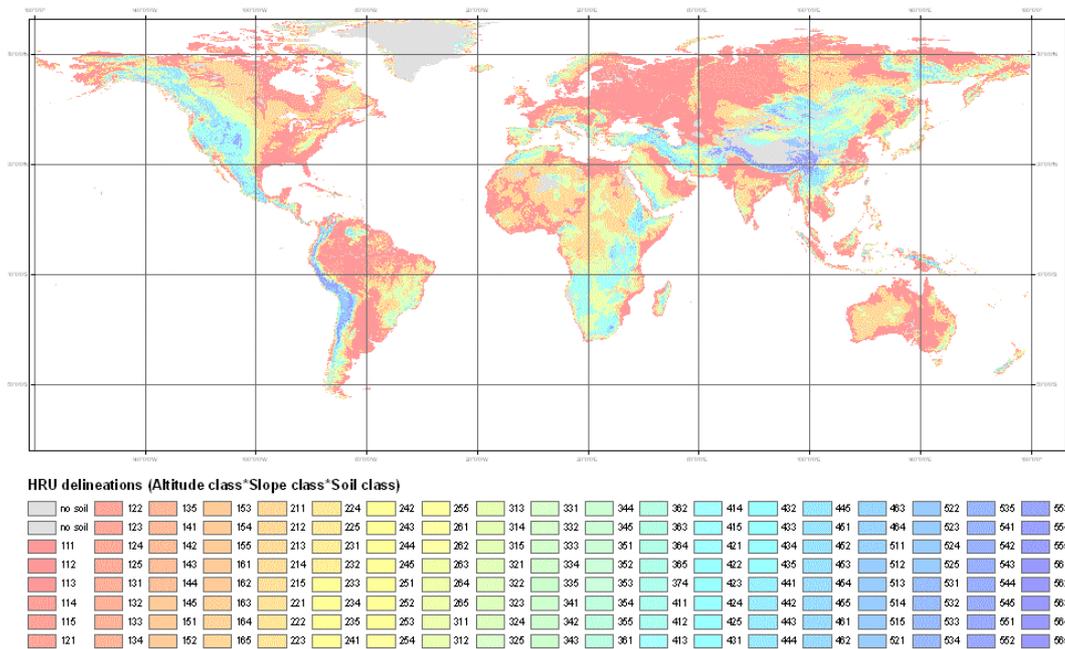
- Foster, A. D., & Rosenzweig, M. R. (2003). Economic growth and the rise of forests. *The Quarterly Journal of Economics*, 118(2), 601-637.
- Galiani, S., McEwan, P. J., & Quistorff, B. (2017). External and internal validity of a geographic quasi-experiment embedded in a cluster-randomized experiment. *Regression discontinuity designs: Theory and applications*, 195-236. Emerald Publishing Limited.
- GEOBENE (2006). *Delineation of Homogenous Response Units (HRU) at global scale*. Retrieved from: <http://www.geo-bene.eu/node/1754.html>
- Geist, H. J., & Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation: Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *BioScience*, 52(2), 143-150.
- Gibbons, S., & Overman, H. G. (2012). Mostly pointless spatial econometrics?. *Journal of Regional Science*, 52(2), 172-191.
- Gibson, C. C., Ostrom, E., & McKean, M. A. (2000). Forests, people, and governance: Some initial theoretical lessons. *People and Forests: communities, institutions, and governance*, 227-242.
- Grainger, A. (2009). Towards a new global forest science. *International Forestry Review*, 11(1), 126-133.
- Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201-209.
- Hargrave, J., & Kis-Katos, K. (2013). Economic causes of deforestation in the Brazilian Amazon: a panel data analysis for the 2000s. *Environmental and Resource Economics*, 1-24.
- IBGE (2016). *Statistics: Estimativas de Populacao 2004*. Retrieved from: https://downloads.ibge.gov.br/downloads_estatisticas.htm?caminho=/Estimativas_de_Populacao/Estimativas_2004/
- Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3), 933-959. ISO 690.
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs/ A guide to practice. *Journal of econometrics*, 142(2), 615-635.
- INPE/PRODES (2017). *Accumulated Deforestation 1988-2012*. Retrieved from: <http://terrabrasilis.info/composer/PRODES>
- Jacob, R., Zhu, P., Somers, M. A., & Bloom, H. (2012). A Practical Guide to Regression Discontinuity. *MDRC*.
- Keele, L. J., & Titiunik, R. (2014). Geographic boundaries as regression discontinuities. *Political Analysis*, 23(1), 127-155.
- Keele, L., & Titiunik, R. (2016). Natural experiments based on geography. *Political Science Research and Methods*, 4(1), 65-95.
- Kirby, K. R., Laurance, W. F., Albernaz, A. K., Schroth, G., Fearnside, P. M., Bergen, S., ... & Da Costa, C. (2006). The future of deforestation in the Brazilian Amazon. *Futures*, 38(4), 432-453.
- Kolstad, I., & Søreide, T. (2009). Corruption in natural resource management: Implications for policy makers. *Resources Policy*, 34(4), 214-226.

- Koyuncu, C., & Yilmaz, R. (2009). The impact of corruption on deforestation- a cross-country evidence. *The Journal of Developing Areas*, 42(2), 213-222.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483-499.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2), 281-355.
- Lee, D. S., Moretti, E., & Butler, M. J. (2004). Do voters affect or elect policies? Evidence from the US House. *The Quarterly Journal of Economics*, 119(3), 807-859.
- Laurance, W. F. (1999). Reflections on the tropical deforestation crisis. *Biological conservation*, 91(2), 109-117.
- Lipscomb, M., & Mobarak, A. M. (2016). Decentralization and pollution spillovers/ evidence from the re-drawing of county borders in Brazil. *The Review of Economic Studies*, 84(1), 464-502.
- Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., & Nobre, C. A. (2008). Climate change, deforestation, and the fate of the Amazon. *Science*, 319(5860), 169-172.
- Mendes, C. M., & Porto Jr., S. (2012). Deforestation, economic growth and corruption: a nonparametric analysis on the case of Amazon forest. *Applied Economics Letters*, 19(13), 1285-1291.
- Meyer, A. L., Van Kooten, G. C., & Wang, S. (2003). Institutional, social and economic roots of deforestation: a cross-country comparison. *International Forestry Review*, 5(1), 29-37.
- Michalopoulos, S., & Papaioannou, E. (2013). National institutions and subnational development in Africa. *The Quarterly Journal of Economics*, 129(1), 151-213.
- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, 334-338.
- Nepstad, D., Soares-Filho, B. S., Merry, F., Lima, A., Moutinho, P., Carter, J., ... & McGrath, D. G. (2009). The end of deforestation in the Brazilian Amazon. *Science*, 326(5958), 1350-1351.
- North, D. C. (1994). Economic performance through time. *The American Economic Review*, 84(3), 359-368.
- Olken, B. A., & Pande, R. (2012). Corruption in developing countries. *Annu. Rev. Econ.*, 4(1), 479-509.
- Pailler, S. (2016). *Local politics and deforestation in the Brazilian Amazon*. Clark University.
- Pedlowski, M. A., Matricardi, E. A., Skole, D., Cameron, S. R., Chomentowski, W., Fernandes, C., & Lisboa, A. (2005). Conservation units: A new deforestation frontier in the Amazonian state of Rondônia, Brazil. *Environmental Conservation*, 32(2), 149-155.
- Pfaff, A. S. (1999). What drives deforestation in the Brazilian Amazon- evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management*, 37(1), 26-43.
- Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., ... & Kirby, K. (2007). Road investments, spatial spillovers, and deforestation in the Brazilian Amazon. *Journal of Regional Science*, 47(1), 109-123.
- Ross, M. L. (1999). The political economy of the resource curse. *World politics*, 51(02), 297-322.

- Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, 12, 487-508.
- Sekhon, J. S., & Titiunik, R. (2012). When natural experiments are neither natural nor experiments. *American Political Science Review*, 106(1), 35-57.
- Smith, R. J., Muir, R. D., Walpole, M. J., Balmford, A., & Leader-Williams, N. (2003). Governance and the loss of biodiversity. *Nature*, 426(6962), 67-70.
- Sundström, A. (2016). Understanding illegality and corruption in forest governance. *Journal of environmental management*, 181, 779-790.
- Søreide, T. (2007). *Forest Concessions and Corruption*. Food and Agriculture Organization of the United Nations.
- Timmons, J. F., & Garfias, F. (2015). Revealed corruption, taxation, and fiscal accountability - Evidence from Brazil. *World Development*, 70, 13-27.
- Transparency International (2017). *What is Corruption?*. Retrieved from: <https://www.transparency.org/what-is-corruption>
- Turner, M. A., Haughwout, A., & Van Der Klaauw, W. (2014). Land use regulation and welfare. *Econometrica*, 82(4), 1341-1403
- Umemiya, C., Rametsteiner, E., & Kraxner, F. (2010). Quantifying the impacts of the quality of governance on deforestation. *Environmental Science & Policy*, 13(8), 695-701.
- Van der Klaauw, W. (2008). Regression–discontinuity analysis: a survey of recent developments in economics. *Labour*, 22(2), 219-245.
- Vanclay, J. K. (1993). Saving the tropical forest: needs and prognosis. *Ambio*, 22, 225.
- Wibowo, D. H., & Byron, R. N. (1999). Deforestation mechanisms: A survey. *International Journal of Social Economics*, 26(1/2/3), 455-474.
- Williams, A., & Siddique, A. (2008). The use (and abuse) of governance indicators in economics: a review. *Economics of Governance*, 9(2), 131-175.
- Wilson, E. O. (1988). The current state of biological diversity. *Biodiversity*, 521(1), 3-18.
- World Bank (1992). *Governance and development*. Retrieved from: <http://documents.worldbank.org/curated/en/604951468739447676/Governance-and-development>
- Wright, S. J., Sanchez-Azofeifa, G. A., Portillo-Quintero, C., & Davies, D. (2007). Poverty and corruption compromise tropical forest reserves. *Ecological Applications*, 17(5), 1259-1266.
- Zamboni, Y., & Litschig, S. (2013). Audit risk and rent extraction: Evidence from a randomized evaluation in Brazil. *Universitat Pompeu Fabra*.

APPENDICES

APPENDIX 1 – GLOBAL HOMOGENEOUS RESPONSE UNIT LAYERS



Notes: Global Homogeneous Response Unit Layers (GEOBENE: Global Earth Observation, 2006).

APPENDIX 2 – BORDER BETWEEN ALTAMIRA AND SÃO FÉLIX DO XINGU



Notes: Satellite image of the border between the two municipalities Altamira and São Félix do Xingu in the state of Pará (Google Earth, 2017).

APPENDIX 3 – POPULATION SIZE AND MUNICIPALITY SIZE

	Deforestation			
	(1)	(2)	(3)	(4)
Corruption	0.00111 (0.845) ^I (0.763) ^{II} (0.306) ^{III}	0.000710 (0.737) ^I (0.682) ^{II} (0.496) ^{III}	0.00160 (0.764) ^I (0.658) ^{II} (0.139) ^{III}	0.000181 (0.932) ^I (0.917) ^{II} (0.862) ^{III}
Population	-3.20e-08 (0.445) ^I	9.21e-09 (0.526) ^I		
Municipality size			-0.000000222 (0.019) ^I	0.000000148 (0.068) ^I
Observations	141,145	141,145	141,145	141,145
Bandwidth (meters)	12,000	12,000	12,000	12,000
Border thickness (meters)	500	500	500	500
R-squared	0.014	0.097	0.016	0.097
Pair fixed effects	No	Yes	No	Yes

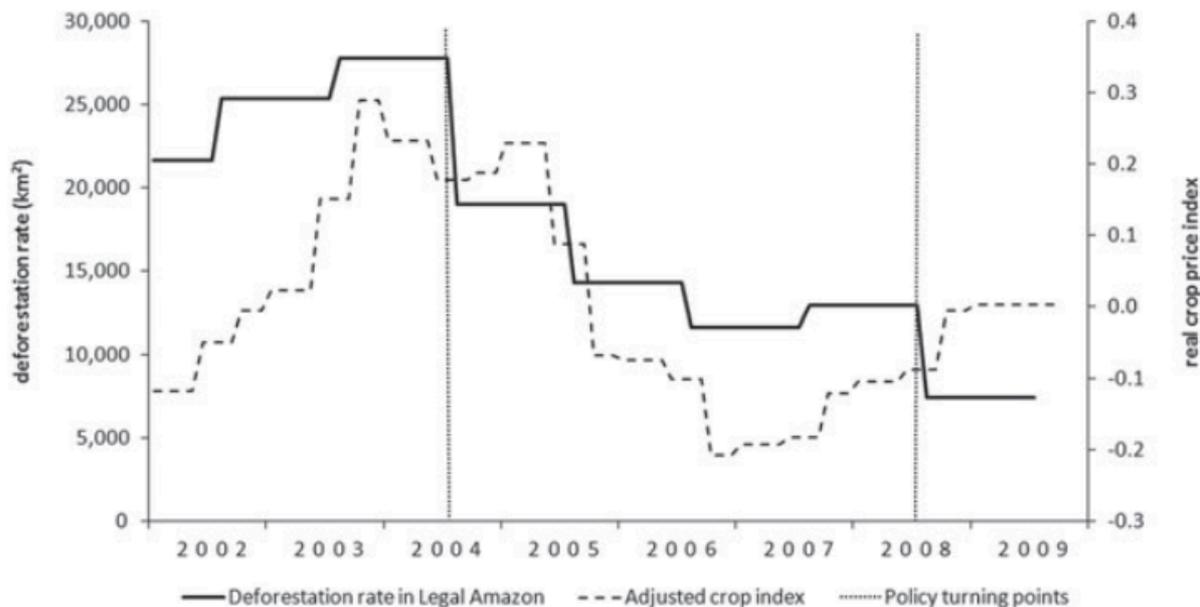
Notes: RD estimates including additional covariates. *p*-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. We control for border fixed effects in column (2) and (4), and for conservation zone and amount of no-forest in all regressions. A linear regression model is applied.

APPENDIX 4 – PLACEBO CUT-OFFS

	Deforestation			
	(1)	(2)	(3)	(4)
Cut-off	-3,000	-6,000	3,000	6,000
Placebo dummy	0.00202 (0.603) ^I (0.752) ^{II} (0.540) ^{III}	0.00128 (0.384) ^I (0.866) ^{II} (0.644) ^{III}	0.00194 (0.490) ^I (0.826) ^{II} (0.501) ^{III}	-0.000568 (0.785) ^I (0.954) ^{II} (0.852) ^{III}
Observations	14,280	19,618	20,088	18,464
Bandwidth (meters)	1,000	1,600	1,400	1,500
Border thickness (meters)	0	0	0	0
R-squared	0.016	0.014	0.015	0.015
Pair fixed effects	No	No	No	No

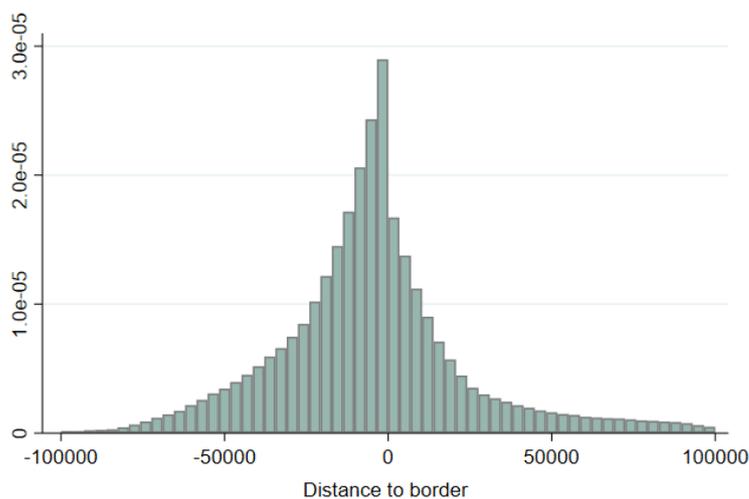
Notes: RD estimates using different cut-offs in the running variable. *p*-values in parentheses: ^Iclustered by municipality, ^{II}clustered by bins, ^{III}without clustering. The cut-offs used are 3,000 and 6,000 meters from the border, on each side of the original threshold (0). We control for conservation zone and amount of no-forest in all regressions. A linear regression model is applied.

APPENDIX 5 – AMAZON DEFORESTATION RATE, AGRICULTURAL PRICE TRENDS, AND POLICY TURNING POINTS BETWEEN 2002 AND 2009



Notes: Annual deforestation rates. The adjusted crop price index is real prices for soybean, cassava, rice, corn and sugarcane. The policy turning points mark the timing of relevant changes in the direction of Brazilian environmental policy. Sources: PRODES/INPE (deforestation) and SEAB-PR (commodity prices) (Assunção et al., 2015, p. 699).

APPENDIX 6 – DENSITY OF THE RUNNING VARIABLE: EXTENDED SAMPLE

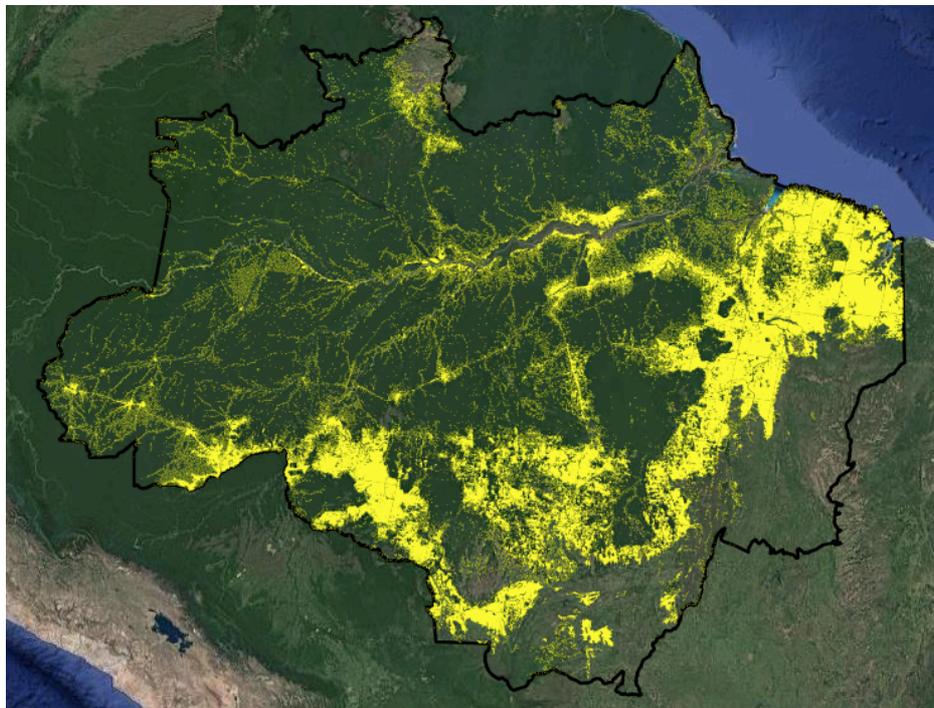


Notes: Histogram illustrating the number of observations against the running variable. Observations lying outside a 100,000 meter bandwidth are excluded. Width of the bins are approximately 3,500 meters.

APPENDIX 7 – MAPS OF THE LEGAL AMAZON



Notes: Map of the Legal Amazon states, with the arc of deforestation indicated by the red stippled line.



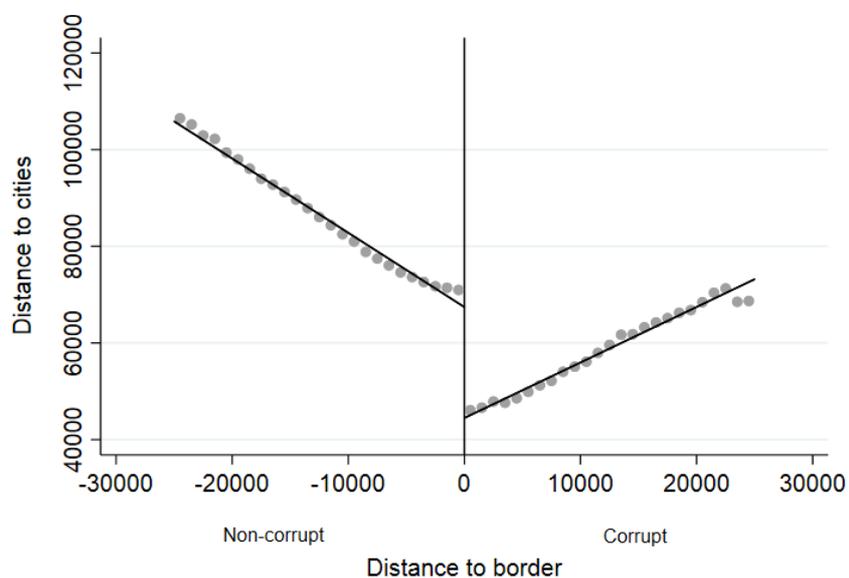
Notes: Satellite image of the Legal Amazon, with accumulated deforestation between 1988 and 2012 indicated by the yellow area (INPE/PRODES, 2017).

APPENDIX 8 – INITIAL FOREST COVER PER STATE IN 2000

State	Initial forest cover in 2000
Amazonas	83,2%
Acre	83,0%
Pará	77,3%
Amapá	60,4%
Roraima	57,9%
Mato Grosso	53,8%
Rondonia	51,4%
Maranhao	9,7%
Tocantins	4,5%
Average across states	53,5%

Notes: Overview of initial forest cover as share of total area for all Legal Amazon states, in the year 2000.

APPENDIX 9 – DISTANCE TO CITIES FOR OBSERVATIONS IN MUNICIPAL PAIRS WITHOUT DISCONTINUITY IN CORRUPTION AT SHARED BORDER



Notes: Distance to cities are plotted against the distance from each observation to nearest municipality border, using only observations in pairs without discontinuity in corruption, in the state of Pará. 1km bins. Bandwidth of 25,000 meters.