STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2021–2022

Criminal Politicians and Deforestation: Causal Evidence from India

Edvin Johannesson (41880) and Sandra Ågren (41900)

Abstract

This paper examines linkages between the criminalization of politics and deforestation on the local level in India, with the ambition of contributing to the emerging corruption-environment nexus in development research. We combine satellite data on forest cover with state assembly election outcomes and criminal records of candidates into a panel of Indian constituencies spanning from 2004 to 2014. Implementing a novel close-election regression discontinuity design, our study provides the first causal evidence of environmental costs due to a criminal politician coming to power. We find that barely electing a criminal politician results in about 3.1 – 3.8 percentage points lower annual forest cover growth over the course of the election term. Albeit limited to competitive elections, results suggest that this effect accumulates and translates to deforestation over time. Together with tendencies of more accentuated negative estimates in highly corrupt and least developed states, this indicates that the effect operates through institutional channels.

Keywords: Criminal Politicians, Deforestation, India, Regression Discontinuity

JEL: D72, D73, Q15, Q23, Q56

Supervisor: Jonathan Lehne
Date submitted: December 6, 2021
Date examined: December 16, 2021
Discussants: Maakwe Cumanzala and Zhaoqin Zhu
Examiner: Magnus Johannesson

Acknowledgements

We want to extend our warmest thanks to our supervisor Jonathan Lehne, for all the invaluable advice and guidance you have given us throughout this journey, and to Toby Lunt at the Development Data Lab for answering all our questions. On that note, we owe our gratitude to the entire DDL team for making this project possible - we admire your amazing work.

Contents

1 Introduction

In the last few decades, the criminalization of Indian politics has been accompanied by a growing number of political scandals, ranging from the embezzlement of public funds to bribery and electoral fraud. In spite of numerous attempts to combat these worrying trends, the share of national legislators with criminal records in India has nearly doubled since 2004 (Times of India 2020a). Recently, the Association for Democratic Reform (2018) reported that one out of three elected representatives to state parliaments were accused of committing a crime. Even though entrusted lawmakers being apparent lawbreakers is problematic from both a democratic and an economic perspective in most settings, recent research suggests that consequences of this are particularly dire in India.[1] Moreover, the abuse of power for private gains is seemingly widespread across the country – a Transparency International (2020) survey estimates that about 50 % of Indians believe most local politicians to be corrupt.

While the economic and social costs of political corruption have been widely studied, evidence on its impact on the environment is not as extensive. Over the past few years, however, the interest for this corruption-environment nexus in academia has grown larger. This trend is not only driven by the fact that countries rich in natural resources tend to be plagued with corruption, but also by unprecedented rates of environmental deterioration worldwide. In this context, deforestation has detrimental impacts on the environment – forest loss is one of the main causes of climate change and poses a great threat to biodiversity and sustainable land use (FAO and UNEP 2020). With the tenth largest forest cover in the world, India has tried to address deforestation issues with a number of acts and commitments including the Forest Conservation Act (1980), Forest Rights Act (2006), the Bonn Challenge and REDD+. Despite this, many regions still suffer from forest loss, and reports of corruption in the forestry sector remain frequent (Kumari, 2019; GFW 2021).[2] Motivated by these trends and with an overarching ambition to contribute to the emerging corruption-environment nexus in the literature, this paper seeks to examine the linkages between deforestation and the criminalization of local politics in India.

We present the first causal evidence of the impact of electing a criminal politician on forest cover. Our analysis builds on data on the criminal records of candidates competing in elections to the State Legislative Assemblies of India. This is based on sworn affidavits required to be filed by all candidates striving to represent their constituency in the State Legislative Assembly as per a 2003 Supreme Court decision, marking one of the first major attempts to combat criminality and corruption in Indian politics. We then combine these criminal records with election outcomes and satellite-based forest cover data to construct a balanced panel of Indian constituencies spanning from 2004 to 2014. Employing a close-election regression discontinuity (RD) design, we estimate a causal and statistically significant negative effect of a criminal politician coming to power.

This finding is our main contribution to the literature: barely electing a criminal politician results in about 3.1 – 3.8 percentage points lower annual forest cover growth over the course of the election term, compared to constituencies where a criminal candidate lost by a small margin. Albeit limited

[1] See for instance Chemin (2012), Banerjee et al. (2013), Fisman et al. (2014), Nanda and Pareek (2016), Vaishnav (2017), Cheng and Urpelainen (2019), Prakash et al. (2019) and Asher and Novosad (2021).

[2] Examples of recent cases of corruption in the forestry sector can be found in the Times of India (2020b; 2020c; 2021a; 2021b; 2021c).

to competitive elections, our results suggest that this effect translates to deforestation over time. Our results thus add an environmental perspective to an emerging body of empirical research examining the effects of electing criminally charged politicians in India on various development outcomes, including Chemin (2012), Nanda and Pareek (2016), Cheng and Urpelainen (2019), and Prakash et al. (2019).

On this note, our second contribution is to illustrate the environmental cost of the criminalization of India's democratic system by presenting constituency-level evidence of local political influence on Indian forests. Although various actions have been taken to reduce deforestation on a national level, our results suggest that these measures are not sufficient in constituencies with criminal incumbents. By considering different definitions of criminal politicians, we show that this environmental cost may increase with the crime record of the representative and that results are generally robust to the severity and number of crimes. Although we cannot draw any conclusive parallels between our estimated causal effects and corrupt behaviour in India based on our data, this paper highlights the possibility of using the criminality of candidates as a proxy for corruption - especially in environmental research.

The third contribution of this paper lies in our identification strategy. Econometric analyses of linkages between corrupt behaviour (or determinants thereof) and deforestation on the regional and local levels are scarce, with exceptions including Smith et al. (2003b), Burgess et al. (2012), Eldeeb et al. (2015) and Pailler (2018). By implementing a state-of-the-art RD design as suggested by Cattaneo et al. (2019), we estimate the causal effect of barely electing a criminal politician over a non-criminal candidate on forest growth. We adopt both robust and optimal local polynomial inference procedures to analyse our estimate and, in a series of extensive validity tests, justify its causal interpretation. In general, we also find that results are robust to various alternatives to choices made throughout the analysis. With this said, we argue that the empirical strategy of this paper also contributes to an ongoing debate over the validity of assumptions invoked in close-election RD designs (Caughey and Sekhon 2011; Eggers et al. 2015; de la Cuesta and Imai 2016; Hyyvinen et al. 2018).[3]

Finally, a modest contribution of our paper relates to potential mechanisms underlying the estimated negative effect of electing a criminal politician on forest cover. Attempting to pin down these, we investigate how estimates vary over the election term, by institutional quality and by level of development. We find a distinct latent effect only in the final year of term and that the impact is accentuated in states considered to have weaker institutions and lower levels of development. As such, we interpret this as indications of the effect on forest growth accumulates over time and operates through institutional channels. Although this interpretation warrants caution, it may provide insights for future research - especially on the role of institutions - within the corruption-environment nexus.

The remainder of this paper is organized as follows. Section 2 provides a background on the Indian electoral system, the role of Members of the Legislative Assembly (MLAs) and forest issues in India. In Section 3, we outline the recent literature relevant for the purpose of this study, followed

[3]Hyyvinen et al. (2018) compare experimental results from tied elections resolved through randomization with the point estimation and inference procedures outlined in Cattaneo et al. (2019) and used in this paper. They find that the results are aligned, and that "state-of-the-art implementation of RD designs can meet the replication standard in the context of close elections".

by a description of our data and the baseline sample in Section 4. Section 5 is devoted to our identification strategy and describes inference procedures in the continuity-based close-election RD design. We present the results of this paper in Section 6 and discuss our findings in Section 7. Section 8 concludes.

2 Background

2.1 Forest Cover and Deforestation in India

Forest cover loss may result from natural causes, such as forest fires, storm damage or forest diseases, or be a consequence of human activities like deforestation or other forestry practices (WRI 2021). Deforestation is often defined as the clearing of forested areas to obtain, for instance, production input, fuel or land for other purposes such as agriculture, mining activities, animal grazing and urban development (Kumari et al. 2019). India has the tenth largest forest cover in the world, with diverse forest landscapes ranging from alpine and scrub to tropical, temperate and subtropical montane forests (FAO and UNEP 2020; Global Forest Watch 2021). In 1980, India enforced the Forest (Conservation) Act (FCA) which included an important ban on cutting trees without permission from government committees to specifically address deforestation (FCA 1980, Aggarwal et al. 2009). Since its implementation, India has made several additional commitments to preserve forests, such as joining the Bonn Challenge and the United Nations REDD+ programme, and involved local communities in forest management.[4]

Although India has begun to tackle deforestation issues in a series of efforts, there are still regions suffering from forest loss (Kumari et al. 2019; Global Forest Watch 2021). Aggarwal et al. (2009) list a number of direct and indirect plausible drivers of deforestation and environmental degradation in India. The direct causes include shifting cultivation, forest fires, diversion of forests for various developmental purposes activities and enduring gaps between demand and supply of key forest products. Increased overexploitation due to population growth and poverty, since many impoverished households depend on forests, are some of the indirect causes of deforestation in the Indian context.

However, different measures and sources paint varying pictures of forest cover change in India. For example, estimates from the Forest Survey of India (2019) indicate that total forest cover increased during the last decades, while the Global Forest Watch (2021) reports a 5 % tree cover loss between 2001 and 2020.[5]

While not going into details about the various techniques used to measure forest cover over time, it suffices to state that not all measures are comparable since they capture different things (Hansen

[4]See for instance the Forest Rights Act of 2006, Borah et al. (2018), Aggarwal et al. (2009) and the Indian National REDD+ Strategy by the Ministry of Environment, Forest and Climate Change (2018) for more details regarding these efforts.

[5]The automated algorithms used by the Forest Survey of India to analyze satellite imagery of forests do not distinguish between plantations and natural forests (Sasaki and Putz 2009; Puyravaud et al. 2010; Davidar et al. 2010; Sarmah 2020; Ashutosh and Roy 2021). Puyravaud et al. (2010) subtracted plantations from the total forest cover in India and found that native forest cover actually declined by 1.5 % - 2.7 % per year. The satellite-based forest cover measure used in this study (see Section 4.3) manages to partially, albeit not perfectly, distinguish between plantations and natural forest cover through thermal signatures (Asher et al. 2020b).

et al. 2010). Thus, there is an important distinction to be made between forest gains, losses and net forest change - since a region may experience the former two simultaneously.[6] In this study, we focus on the net annual forest cover growth and hence capture both forest losses and forest gains (see Section 4.3 for more details).

2.2 Electoral System

India is a federal state with partially self-governing states, holding significant control over their territories (Election Commission of India 2018). While both the national and state level parliamentary legislature consists of an Upper and a Lower House, this paper focuses on elections to the state-level Lower House called the Legislative Assembly (or Vidhan Sabha). Each constituency elects one representative to the Legislative Assembly based on the plurality rule of first-past-the-post (FPTP) voting. That is, electors choose one candidate to cast their vote on and the winner is the candidate who received the most votes in the constituency. As Members of the Legislative Assembly (MLAs), these representatives are elected for a term of 5 years and may be re-elected (Ravishankar 2009; Election Commission of India 2018). Holding both executive, financial and legislative power within their territories, the Indian Constitution grants State Legislative Assemblies some influence over factors that have direct or indirect bearing on deforestation.[7] As per the Indian REDD+ strategy by MoEFCC (2018), this is stipulated in the fifth Directive Principle of State Policy of Article 48: "The State shall endeavour to protect and improve the environment and to safeguard the forests and wildlife of the country".

2.2.1 Delimitation of Constituency Boundaries

In India, constituency boundaries are regularly readjusted to accommodate for disproportionate population changes across constituencies (Election Commission of India 2020). In our sample, these boundaries were fixed until 2008, when they were redrawn based on the Delimitation Act of 2002. Hence, all Legislative Assembly elections held in 2008 or later rest on the new constituency boundaries, while elections prior to this define constituencies as per the old delimitation. Since we observe forest growth rates within both old and new boundaries for the entire sample period, the delimitation of 2008 is not a major problem for our empirical strategy even though it renders constituencies before and after the changes non-comparable (see Section 4.3). It does, however, make it difficult to compare certain predetermined characteristics for elections held just after the redrawing of boundaries – simply because electoral data for previous terms is unavailable for post-delimitation elections.

[6]This could for instance happen either when some subregions experience deforestation while others are reforested, or when two different measures of forest change are used where one includes non-natural plantations while the other does not.

[7]For example, Schedule VII of the Constitution gives assemblies legislative power pertaining to forests and local infrastructure such as roads (see MoEFCC 2018)

2.3 The Role and Power of Members of the Legislative Assembly

Although individual MLAs have limited legislative power, previous research on Indian politicians accentuates their informal power and role as intermediaries between residents and the state. For example, they have access to development funds and can secure resources for their constituencies (Jensenius 2015; Asher and Novosad 2017, Cheng and Urpalainen 2019). They may also appoint and reassign local public officers (Jensenius 2015; Vaishnav 2017). Furthermore, MLAs exert considerable informal influence over many bureaucratic processes, including environmental clearances and surface rights to government-owned land (Asher and Novosad 2021). These informal activities are summarized by Kopas et al. (2021) as a role of a mediating agent between project developers, regional institutions and local residents in land acquisition consultations. Moreover, anecdotal evidence suggests that such networks exist in the context of Indian forest management. In an interview-based study of corruption in the forestry sector of Rajasthan, Robbins (2000) highlights the importance of social status and in establishing "strong bonds of trust for extra-legal exchange" between local officials and foresters from similar backgrounds. Although this study was carried out two decades ago, corruption cases in the forestry sector are still cited in Indian news today.[8] As such, we hypothesize that MLAs with local elite status can exert their formal and informal power over forest development in their constituency.

Relatedly, corruption is widespread in Indian politics and politicians with criminal records are a common feature of both national- and state-level legislatures (Prakash et al. 2019). As a step towards more transparency and accountability, all candidates to national and state parliaments are required to file affidavits with personal information following a Supreme Court ruling in 2003 (Vaishnav 2017). These affidavits include details about the assets, liabilities, education, professional background and, most importantly, criminal records of a candidate. The latter is restricted to criminal charges filed more than six months before the election and prior convictions (Prakash et al. 2019). Candidate affidavits provide a valuable source for understanding how criminal politicians impact governance issues in India.

3 Literature Overview

This section outlines the recent literature relevant for the purpose of this study, which is done in three steps. We begin with a general description of the corruption-environment nexus and then limit the scope to research on corruption and deforestation. The third subsection focuses on empirical evidence of the impact of corrupt politicians in India.

3.1 Corruption and Environmental Degradation

There are a number of theoretical explanations for how corruption may affect the deterioration of natural resources in the literature. One commonly proposed mechanism is that potential monetary gains, such as bribes, incentivize local officials to allow illegal overexploitation of resources (Smith and Walpole 2005; Tacconi et al. 2009). Other theories argue that corruption and embezzlement

[8] Several news articles in Times of India describe this issue (see e.g. 2020b, 2020c, 2021a, 2021b and 2021c).

reduce the power of environmental regulations through for instance weakened law enforcement and financial resources (Damania et al. 2003; Fredriksson et al. 2004; Haseeb and Azam 2021).

In this context, an influential framework frequently adopted when analysing the relation between economic development and environmental indicators is the Environmental Kuznets Curve (EKC). The EKC is a hypothesized inverted U-shaped relationship between environmental degradation and income, which is positive up to a certain point of economic development and then becomes negative as environmental quality starts improving (Cole et al. 1997; Dinda 2004). It has been argued that the prevalence of corruption, especially in developing countries, shifts this turning point eastward to higher levels of income (Lopez and Mitra 2000). Although most studies applying the EKC framework examine environmental degradation in the form of pollution, there is also some literature focusing on deforestation.[9]

Building on the EKC hypothesis, Welsch (2004) proposes and empirically examines two different channels through which corruption may affect environmental quality in the form of air pollution. The first is the direct effect, in which corruption contributes to environmental deterioration by eroding the stringency and enforceability of environmental regulations at given levels of economic development. This channel has been widely studied, not least when it comes to the role of institutional quality in reducing the direct effect (Lippe 1999; Culas 2007; Pailler 2018). The second channel investigated by Welsch (2004) is the indirect effect, where corruption reduces environmental quality through its negative impact on economic development – arguably mediated through the EKC. As exemplified by Welsch (2004), most studies within the field of corruption and environment concentrate on the direct and indirect effects on air pollution and emissions (Cole 2007; Leitão 2010; Biswas et al. 2012; Halkos and Tzeremes 2013; Greenstone and Jack 2015; Sinha et al. 2019). In one of the few experimental analyses of these linkages, Duflo et al. (2013) finds that widespread corruption undermines industrial pollution monitoring in the Indian state of Gujarat and that incentivizing more accurate reporting leads to reduced emissions. In a more recent study, Zhou et al. (2020) combines exogenous variation from an anti-corruption campaign in China with legal data on corruption cases among bureaucrats to show that the policy resulted in reduced air pollution in several cities.

3.2 Corruption and Deforestation

Despite the predominant focus on air pollution, a growing body of research within the corruption-environment nexus attempts to analyse the linkages between corruption and deforestation (see for instance Palmer (2001), Barbier (2004), Smith et al. (2006), World Bank (2006), Laurance et al. (2011), Urrunaga et al. (2012), Galinato and Galinato (2013), Koyuncu and Yilmaz (2013), Sundström (2016), Meehan and Tacconi (2017), Abman (2018) and Cozma et al. (2021)). In general, however, the estimated impact of corruption on forests is ambiguous, much depending on the regional context and the type of measurements used.

One strand of research in this field adopts cross-country analyses to investigate the relationship between corruption and deforestation. This includes Smith et al. (2003a), who analyze the links

[9]See for instance Ehrardt-Martinez et al. (2002), the analysis by Culas (2007) of institutional factors impacting this relationship, and Choumert et al. (2013) for a meta-analysis of 69 studies on the EKC for deforestation.

between governance scores given by the Corruption Perception Index and forest cover in 126 countries. They find a positive correlation between mean governance scores and changes in total forest cover, but not when only considering natural forests. These findings were later questioned in a follow-up study by Barrett et al. (2006) which employed several robustness checks and found no statistically significant correlation. On the other hand, Koyuncu and Yilmaz (2009) do find statistically significant positive correlation between corruption and deforestation, using three different types of corruption indices for a sample of 100 countries in different time periods.

There are also several studies centred on developing theoretical frameworks for forest management and testing their predictions for different levels of corruption. Barbier et al. (2005) develop an open-economy model of resource conservation, which they test empirically in middle and low income countries. They find that higher levels of corruption will, depending on terms of trade, affect forest control in countries where the government is corrupt. Another theoretical approach is the lobbying model proposed by Bulte et al. (2007). Applied on nine countries in Latin America, the model predicts that corruption indirectly contributes to deforestation from land expansion through its effect on fiscal instruments. Amacher et al. (2012) develop a framework which compares concession policy design in scenarios with and without corruption, and show that the central government adjusts concession policies in the presence of corruption introduced as bribery in local forest management.

Many studies of corruption and deforestation on the regional level use qualitative and descriptive methods, with examples from Indonesia (McCarthy 2002), Tanzania (Milledge et al. 2007) and Ghana (Teye 2013). Robbins (2000) gathered anecdotal evidence through 162 interviews with local forestry actors in the Indian state of Rajasthan, which indicated that bribing forestry officials to enable illegal logging was common in 1998. However, econometric analyses of the corruption-deforestation relationship on a regional level remain scarce, with exceptions including Smith et al. (2003b), Burgess et al. (2012) and Eldeeb et al. (2015). In a recent study, Pailler (2018) combines satellite-derived deforestation data with electoral statistics from 2002-2012 to construct a municipal-level panel data set and examine whether deforestation rates increase when an incumbent mayor runs for re-election. She shows, inter alia, that municipalities where highly corrupt incumbents run for re-election experience 45-60 % more deforestation during the election cycle than electorates without an incumbent mayor running for re-election.[10] In another recent paper, Cisneros et al. (2021) analyse political incentives and deforestation in Indonesia using a district-level panel data set ranging 2001 to 2016. Their results indicate that deforestation rates are on average 4 % higher in the year prior to local mayoral elections, which may reflect politicians engaging in more corrupt activities to raise funds ahead of elections.[11]

Some of the more specific channels through which corruption can affect deforestation are, as outlined in Tacconi et al. (2009), by impairing land allocation processes and undermining the enforcement and control of land use plans (Damania et al. 2003). Several studies have also focused on the aforementioned bribery mechanism, which may manifest in the deforestation context as bribes to local politicians or forestry officials. These bribes could be aimed at allowing for timber harvesting without a legal permit (Smith et al. 2003b), facilitating transportation of illegally logged timber

[10] Pailler (2018) uses a corruption measure which is based on irregularities in fiscal audits of randomly selected municipalities, initiated during Brazil's anti-corruption campaign in 2003. Originally constructed and reviewed by Ferraz and Finan (2011) and Brollo et al. (2013), this measure has become influential in the corruption literature.

[11] This is in line with previous findings in for instance Khemani (2004) and Aidt et al. (2020).

(Southgate et al. 2000) or reducing monitoring of concessions (Barnett 1990).

3.3 Criminal Politicians and Political Corruption in India

Following the 2003 Supreme Court ruling and the subsequent availability of affidavit data on criminal records, an emerging line of corruption research has centred on examining the effect of criminally charged politicians in India – including Banerjee et al. (2014) and Fisman et al. (2014). Chemin (2012) was among the first to use this unique data for the 2004 elections to State and National Legislative Assemblies, employing a regression discontinuity (RD) design to examine the impact of electing a criminal candidate on corruption, consumption and crime rates. He finds a statistically significant negative effect on consumption among vulnerable groups and provides suggestive evidence that this, in turn, might be driven by the impact on crime rates and corruption. Although results also indicate that electing a criminal politician significantly reduces corruption, as proxied by the value of gifts received by local officials, Chemin (2012) concludes that this interpretation warrants caution since his proxy is likely to be insufficient and underestimate other forms of corruption.

In a recent paper, Prakash et al. (2019) implements a similar RD design to study whether constituencies electing a criminally charged representative to the State Legislative Assembly experience different economic growth rates, using night-time light intensity as a proxy. They find a causal effect amounting to roughly 2.4 percentage points lower GDP growth compared to constituencies barely electing a non-criminal candidate. Another contribution using data on criminal politicians in a close-election RD setting is Cheng and Urpelainen (2019), which considers various village-level outcomes related to electricity connection, road accessibility, poverty and human capital. Their findings shed light on the importance of differentiating between policy outcomes, as their estimated effect of electing a criminal politician is negative for both literacy rates and household electrification but inconclusive for local infrastructure projects. Similarly, Nanda and Pareek (2016) find reductions in total firm investment following the election of criminal candidates in an analysis resting on the same affidavit records and method.

Beyond efforts using the criminality accusation status of candidates as a proxy for corruption, other close-election RD designs on the regional level in India focus on different indicators. By matching names of winning politicians and road construction contractors, Lehne et al. (2018) provide evidence of political corruption in local infrastructure projects. Both Asher and Novosad (2017) and Kopas et al. (2021) look at partisan alignment with the ruling party, where the former study finds generally positive effects of political favouritism on economic outcomes and the latter show that local partisan alignment leads to more environmental clearance applications. Likewise, the analysis of Mahadevan (2019) indicates a discrepancy between billed and actual electricity consumption in constituencies aligned with the governing party. Her estimates suggest that this is a consequence of politicians allowing for systematic manipulation of electricity bills.

To the best of our knowledge, the causal effect of corruption on natural resources in India has not been empirically examined to this date. However, the reverse relationship has recently been studied by Asher and Novosad (2021). Exploiting global price shocks affecting local mineral rents, they investigate how rent-seeking behaviour impacts the election and behaviour of politicians. They find that increased rents cause the election of criminal politicians and that, if the value of local minerals continues to rise during their election term, these officials tend to face more criminal charges and become richer over time.

4 Data

To identify the effect of electing a criminal politician on forest growth in India, we need three different types of information: constituency-level forest cover data, electoral outcomes and criminal records of candidates. This data is collected from the Socioeconomic High-resolution Rural-Urban Geographic Data Platform for India (SHRUG), an open access repository maintained by the Development Data Lab (DDL), and then combined into the single panel ranging from 2004 to 2014 used in the analysis.

4.1 Electoral Data

Data on results from elections to State Legislative Assemblies originates from information published by the Election Commission of India (ECI). This data reports specific information about candidates and the elections they compete in. On the candidate level, this includes party affiliation, the number of votes received and gender, while constituency-level electoral data includes the total number of votes in a specific election and the size of the electorate. Furthermore, the electoral dataset specifies whether constituency seats are reserved for minority groups in India: Scheduled Tribes (ST) and Scheduled Castes (SC) (Prakash et al. 2019). Representatives to State Legislative Assemblies in India are elected every fifth year, but electoral cycles vary across states.

Our main analysis considers elections between 2004 and 2010, limited by the timeframe of the candidate affidavit and forest cover data. As described in Section 2.2.1, elections held between 2004 and 2007 are based on old constituency boundaries while elections from 2008 and onwards define constituencies by the new delimitation order. Our electoral data was originally processed by Jensenius and Verniers (2017) and published by the Trivedi Center for Political Analysis at Ashoka University.[12]

4.2 Candidate Affidavit Data

All candidates running for election in India are required to file a sworn affidavit with details about their assets, liabilities, education and professional background. Crucially, these affidavits also contain information about the criminal records of candidates in the form of prior convictions and charges filed more than six months before the election. The raw affidavits provided by ECI were originally digitized by the Association for Democratic Reform (ADR), and later validated and cleaned by Prakash et al. (2019) who also extended the data before making it available in the SHRUG. Our affidavit data consists primarily of the Prakash et al. (2019) entries for the pre-

[12]For more information about the election data, we refer to the SHRUG resources compiled by Asher et al. (2020a; 2021) and Jensenius and Verniers (2017)

delimitation period of 2004-2007, and ADR data for candidates running for elections between 2008 and 2018. Prior to publishing this data in the SHRUG, Asher et al. (2021) also added information about years of punishment for the most serious criminal charge faced by a candidate.[13] The data we collect from the DDL only contains affidavit information for the two candidates gaining the most votes in each election.

We define candidates as criminal if they reported at least one criminal charge ahead of the election (as in Chemin (2012), Banerjee et al. (2014), Fisman et al. (2014), Asher and Novosad (2021), and Prakash et al. (2019)). Thus, we obtain our treatment indicator - a binary variable taking the value 1 if the candidate had at least one criminal charge six months prior to the election and 0 otherwise. It is important to note that this definition includes candidates who have been accused of crimes but not convicted. This may be worrying since we risk assigning a criminality status to candidates who have, for instance, been charged for political activism such as democratic protesting or subjected to fabricated charges by their political opponents (Jaffrelot and Verniers 2014; Prakash et al. 2019). However, the latter is likely to be a minor concern as only charges filed six months before the election are accounted for. Nonetheless, we conduct our analysis with more conservative definitions of a criminal candidate in Section 6.3.1 to assess this potential noise, which yields similar conclusions.

4.3 Forest Cover Data

Our outcome of interest is based on satellite imagery from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) and several higher resolution satellites, refined through a machine learning model by DiMiceli et al. (2015) into the Vegetation Continuous Fields (VCF) data used in the analysis. This data describes yearly forest cover as the predicted percentage of each 250x250 m pixel covered by forest (Townshend et al. 2011). Polygon shapefiles were then used by Asher et al. (2021) to accurately link roughly 90 % of SHRUG locations in India to this raw VCF data, while about 10 % of the locations were approximated with Thiessen polygons extrapolated from a single pixel.[14]

The VCF data we collect from the SHRUG is aggregated by the DDL to the constituency level and is available from 2000 to 2014.[15] This was done separately for pre- and post-delimitation constituency boundaries over the *entire* period. Consequently, we observe forest cover data within our timeframe for constituencies defined under both the old and the new delimitation orders. Hence, the redrawing of constituency boundaries in 2008, discussed in Section 2.2.1, is not a problem for our analysis.

We use the variable average forest cover, defined as the sum of all pixel values divided by the number

[13]Asher et al. (2021) matched candidates in the affidavit data to the corresponding election in the electoral data, resulting in unique candidate- and election-level identifiers for about 25 % of all elections available in the SHRUG. These are the elections we utilize when constructing our sample. Further details regarding the affidavit data are available in Prakash et al. (2019) and Asher et al. (2020a; 2021).

[14]Some of the locations approximated with Thiessen polygons lie in densely forested areas, primarily in the north-east of India (Asher et al. 2021). We thus investigate whether our results are robust to excluding states in which more than 50 % of the locations were Thiessen polygons. Six states (Arunachal Pradesh, Manipur, Meghalaya, Mizoram and Nagarand) fulfil this criterion, but only 11 elections are excluded. Table A1 of the appendix confirms the robustness of our results to excluding these locations.

[15]For detailed discussions of the VCF data, see Townshend et al. (2011), Hansen et al. (2013), DiMiceli et al. (2015), Asher et al. (2020b) and the SHRUG resources of Asher et al. (2020a; 2021).

of pixels in a constituency for a given year, to construct our key dependent variable – annual growth in average forest cover over the election term.[16] Figure A1 in the appendix illustrates the average forest cover across Indian districts in 2014. This figure shows that, on average, the most densely forested areas in India are found in the north-east, south-west and in the northern parts of the country.

Our measure of annual forest growth captures both forest cover gains (growth) and losses (degrowth) - forest cover change is commonly used as a proxy for deforestation in the literature (see for instance Koyuncu and Yilmaz (2009; 2013), Cisneros (2021), and Cozma (2021)). As discussed in Section 2.1, forest loss may result from both natural causes and intentional or illicit deforestation. While it is important to emphasize that our measure do not allow us to fully distinguish between these different causes, the VCF data is based on broad-spectrum techniques which discriminate between thermal signatures of for instance plantations and naturally occurring forests (Asher et al. 2020a). This feature is generally not embedded in other frequently used forest data sources, such as the Global Forest Watch (GFW) database. Moreover, sources like the GFW data do not reflect net changes in forest cover since they tend to aggregate only over a binary indicator of forest cover loss, thus not accounting for forest gains (Asher et al. 2020b; WRI 2021).[17] Since most assessments, including our own, suggest that total forest cover in India marginally increased during our sample period, other sources do not seem to give an accurate picture of the context of this study (see Asher et al. (2020a) and Table 1 below). In light of these considerations, we argue that the VCF data is the appropriate choice for empirically examining causes of deforestation in India.

In our analysis, we focus on constituencies with forested areas. That is, we want to avoid including constituencies with negligible forest cover in the estimations, such as those located in densely populated areas, deserts and other barren landscapes. Motivated by this, we follow Hargrave and Kis-Katos (2013) and Pailler (2018) and limit our sample to constituencies with at least 10 % average forest cover at the time of the election. Even though this restriction might appear rather extreme, especially since it reduces the sample size by over 75 %, we deem it imperative to eliminate the otherwise highly plausible bias induced by disproportionate vegetation growth rates in these constituencies.[18]

As with all satellite data, the VCF is sensitive to atmospheric phenomena, such as aerosols, clouds and surface reflections, as well as the angle from which the satellite records imagery at a given time (DiMiceli et al. 2015; Brown 2021; Santos et al. 2021). More specific examples of such disturbances include monsoons, storms or smog spill-overs from neighbouring cities. Although the algorithms used to process the VCF data partially filter out this noise, some measurement error

[16]We calculate this as the difference in the natural logarithm of average forest cover for a constituency between the current and previous year, multiplied by 100. Results are robust to growth rates computed by the standard discrete-time growth formula. Defining forest cover change as the level difference in average forest cover also yields similar conclusions. These results are available upon request.

[17]More specifically, the GFW data is based on binary indicators, defining a pixel as reforested only if its forest cover percentage goes from zero in 2000 to a positive value in 2012 and deforested if more than 90 % of the initial 2000 forest cover was lost by a given year (Hansen et al. 2013; Asher et al. 2020b).

[18]In Table A2, we conduct our main analysis on constituencies with at least 5 % forest cover in the election year instead. While this gives roughly twice as many effective observations, it is also likely to introduce more bias to the estimates. Despite this, our preferred estimations in columns (3) and (4) of Panel A yield similar conclusions. Reassuringly, analogous estimations on the full sample (Panel B) and using only constituencies with less than 10 % forest cover (Panel C), respectively, yield insignificant results – although the coefficients of interest remain negative in the former case.

is likely to remain. Arguably, this contributes to the occasionally large annual fluctuations in our forest growth measure – which are reflected by the comparatively large standard deviation of 37.3 in the full sample (see Table 1). We take several steps to control for this unrepresentative variation. The first is the sample limitation detailed above. Second, we exclude constituencies who experience seemingly abnormal forest cover changes during our sample period, to make sure extreme outliers are not driving our results.[19] Finally, Section 6.1 includes state-year fixed effects in the analysis.[20]

4.4 Baseline Sample and Descriptive Statistics

We obtain our full sample by merging the electoral, affidavit and forest cover data using unique identifiers provided by the DDL. This yields a balanced panel of 4232 unique constituencies and thus elections for our sample period of 2004-2014.[21] Since the close-election RD design requires an appropriate control group for comparison with constituencies where a criminal candidate won the election, we restrict the full sample to constituencies where only one of the top two candidates was charged with a crime. In other words, we only include constituencies where either a criminal candidate won against a non-criminal runner-up or where a non-criminal candidate won against a criminal runner-up (as defined by our treatment indicator). This reduces the full sample to 2373 unique constituencies and elections.

As necessitated above, excluding extreme outliers and constituencies with less than 10 % average forest cover in the election year leaves us with our final baseline sample used in the RD design. This balanced panel consists of 533 unique constituencies and thus elections.[22] For each constituency, we observe annual forest growth rates for the four years following the election year – totalling 2132 observations in our baseline RD sample.[23] Summary statistics and descriptions for all relevant variables included in our analysis are presented, separately for the full and the baseline RD sample, in Table 1.

[19]We define outliers as observations which lie outside the outer fence (that is, three times the interquartile range) of our baseline sample, leading to the exclusion of six constituencies. Table A3 shows that our main results are robust to this.

[20]Beyond nearly halving the standard deviation of our measure compared to the full sample, the inclusion of state-year fixed effects raises the R^2 of a least squares estimation of equation (6) in Section 5.1 from about 0.003 to approximately 0.30.

[21]Although appending the panel by including elections up until 2014 is possible, we do not observe annual forest growth rates for the entirety of these election terms (due to the timeframe of the VCF data). Extending the dataset would thus give us roughly 20-25 % more effective observations but come at the cost of rendering our panel unbalanced. Since the latter years of term for these additional elections are non-randomly and systematically "missing", an unbalanced panel would by construction bias our estimates towards the earlier years of the election term. We hypothesize that it takes time for an elected politician to change previous standards and norms, which we provide suggestive evidence for in Section 6.3.2. Hence, we use the balanced panel for our analysis. Table A4 shows that our main results are robust to using an unbalanced panel and, given the comparatively smaller magnitude of our preferred estimates in columns (3) and (4), arguably further strengthens the hypothesis of a latent effect.

[22]Our baseline sample consists of constituencies across 25 out of the 28 states in India. The following states are included (with the number of constituencies in parentheses): Andhra Pradesh (28), Arunachal Pradesh (5), Assam (16), Bihar (33), Chhattisgarh (1), Dehli (3), Goa (6), Haryana (3), Himachal Pradesh (30), Jammu Kashmir (11), Jharkhand (12), Karnataka (12), Kerala (78), Madhya Pradesh (3), Maharashtra (79), Manipur (2), Mizoram (4), Nagaland (1), Orissa/Odisha (46), Punjab (14), Sikkim (9), Tamil Nadu (36), Uttar Pradesh (24), Uttarakhand (18) and West Bengal (59).

[23]As per Prakash et al. (2019), we exclude the annual forest growth rate for the election year (see Section 5.1 for details).

Table 1: Descriptive Statistics

	Full Sample						Baseline RD Sample					
	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max
Forest Growth	16928	1.22	1.89	37.3	-181.9	319.2	2132	-0.17	-0.90	29.4	-97.4	100.5
Forest Growth (t-1)	4232	-1.11	-2.93	33.1	-234.2	139.2	533	-2.12	-1.99	24.7	-84.3	66.7
Forest Growth (t-2)	4232	-2.42	-0.11	41.4	-135.3	503.4	533	-0.57	-4.47	29.2	-92.0	83.6
Forest Growth (t-3)	4232	8.61	8.74	33.6	-554.1	161.5	533	6.67	8.06	27.2	-77.4	126.2
Log Forest Cover (t)	4232	2.15	2.21	0.85	-2.98	4.37	533	2.80	2.92	0.48	2.30	4.29
Criminal Winner	4232	0	0.31	0.46	0	1	533	1	0.52	0.50	0	1
Log Const. Area	4232	524.0	800.7	1947.7	0.25	79983.3	533	418.9	680.3	810.5	0.25	9513.9
Const. SC Reserved	4232	0	0.14	0.35	0	1	533	0	0.11	0.31	0	1
Const. ST Reserved	4232	0	0.12	0.32	0	1	533	0	0.094	0.29	0	1
Log Electors	4232	12.1	11.9	0.74	8.23	14.3	533	12.0	11.9	0.67	8.65	13.7
Log Valid Votes	4232	11.7	11.5	0.68	7.45	13.7	533	11.7	11.5	0.64	8.54	13.2
Winner Gender	4232	1	0.92	0.27	0	1	533	1	0.94	0.24	0	1
Runner-up Gender	4232	1	0.92	0.27	0	1	533	1	0.93	0.26	0	1
Winner Age	4228	49	49.0	9.96	25	88	533	49	49.5	10.1	26	82
Runner-up Age	4224	49	49.2	10.5	25	86	533	49	49.5	11.0	25	86
Winner Education	4020	12	11.8	2.41	0	14	513	12	11.9	1.98	2	14
Runner-up Education	3993	12	11.8	2.40	0	14	501	12	11.8	2.35	0	14
Winner Log Assets	4199	15.4	15.4	1.72	7.30	21.5	533	15.0	15.1	1.76	8.70	19.8
Runner-up Log Assets	4193	15.3	15.2	1.79	5.64	22.0	529	15.1	14.9	1.78	7.50	20.7
Winner Type (ST/SC)	4232	0	0.28	0.45	0	1	533	0	0.22	0.41	0	1
Runner-up Type (ST/SC)	4232	0	0.28	0.45	0	1	533	0	0.22	0.41	0	1

Variable descriptions: *Forest Growth* is the annual average forest cover growth during the four years after the election year; *Forest Growth (t-x)* specifies the annual average forest cover growth x years ahead of the election; *Log Forest Cover (t)* is log average forest cover at the year of the election (t); *Criminal Winner* is the treatment indicator taking the value 1 if the winning candidate was criminal and 0 if the winning candidate was non-criminal; *Log Const. Area* is the log constituency area (in km2); *Const. SC/ST Reserved* is equal to 1 if a constituency seat is reserved for Scheduled Castes (SC) or Scheduled Tribes (ST) and 0 otherwise; *Winner/Runner-up Gender* is equal to 1 if winner is male and 0 if winner is female; *Winner/Runner-up Age* is the age of a candidate; *Winner/Runner-up Education* is the number of years of education of a candidate; *Winner/Runner-up Log Assets* is the log value of assets possessed by a candidate; *Winner/Runner-up Type (ST/SC)* is equal to 1 if a candidate belongs to a Scheduled Caste (SC) or a Scheduled Tribe (ST) and 0 otherwise.

5 Identification Strategy

The goal of this paper is to investigate whether constituencies electing criminal politicians experience different rates of forest cover growth compared to those represented by non-criminal members of the Indian State Legislative Assembly, and to identify whether this effect is causal in nature. While a simple comparison of average outcomes between these two groups might suffice to achieve the former, it would yield a biased estimate for which causal identification would be invalid. This could arise due to unobservable heterogeneity across constituencies or, similarly, the highly plausible endogeneity of our treatment – electing a criminal MLA. In other words, the presence of unobservable constituency- or candidate-level characteristics, which may explain the candidacy or victory of criminal politicians and also correlate with deforestation, poses a major challenge for the analysis. To tackle this, we employ a regression discontinuity (RD) design in an attempt to identify the causal effect of electing a criminal politician on annual forest cover growth.

First introduced by Thistlehwaite and Campbell (1960), later pioneered in the contributions of Angrist and Lavy (1999) and Hahn et al. (2001), and further formalized by Imbens and Lemieux (2008) as well as Lee and Lemieux (2010), to name a few, the RD design has grown increasingly popular within many fields of economics. In what is commonly referred to as the "sharp" RD design, binary treatment assignment is a deterministic function of a continuous covariate, often called the running variable. Since individuals with a value greater or equal to some predetermined threshold of the running variable are placed in the treatment group, the probability of treatment jumps discontinuously from zero to unity at this cut-off (Cunningham 2021; Lee and Lemieux 2010). Under the crucial assumption that nothing else than treatment assignment varies discontinuously at this threshold, which will be investigated in Section 6.2, one can retrieve a causal estimate of the relationship of interest in the neighbourhood of the cut-off. With this in mind, elections provide a promising setting for the sharp RD design since candidates win an election only if they receive a vote share above a certain threshold (Chemin 2012; Eggers et al. 2015). Following the seminal paper of Lee (2008), which applies an RD design to close races between the top two candidates running for the United States House of Representatives, the close-election RD design has been widely used to investigate the effect of certain candidate- or constituency level characteristics on economic outcomes.[24]

5.1 Main Specification

In our context, the close-election sharp RD design translates to comparing annual average forest cover growth rates in constituencies where a criminal runner-up barely lost to a non-criminal winner (the control group) with constituencies where a criminal winner barely won over a non-criminal runner-up (the treatment group). In order to achieve this, we first combine electoral data and criminal records from candidate affidavits to generate our running variable and treatment indicator. As opposed to electoral systems with a majority rule, like the extensively studied US House elections (Caughey and Sekhon 2011), our electoral setting of Indian Legislative Assembly elections with FPTP plurality voting (as discussed in Section 2.2) necessitates a relative threshold based on the difference in vote shares between the top two candidates. Thus, we define our running variable margin in constituency i in election year t as:

$$margin_{i,t} = \frac{votes(criminal)_{i,t} - votes(non\text{-}criminal)_{i,t}}{votes(total)_{i,t}} * 100 \tag{1}$$

This variable is, by construction, negative if a criminally accused candidate lost the election (in favour of a non-criminal candidate) and positive if the winner was criminal (where the runner-up was not). In other words, $margin_{i,t}$ can be interpreted as the margin of victory of a criminal candidate. Our treatment indicator, whether the winner in constituency i in election year t was criminally accused or not, is hence given by:

$$criminalwinner_{i,t} = \begin{cases} 0 & \text{if } margin_{i,t} < 0 \\ 1 & \text{if } margin_{i,t} \geq 0 \end{cases} \tag{2}$$

As evident from equations (1) and (2), the criminality status of the winning candidate changes discontinuously from non-criminal to criminal at the threshold value of $margin_{i,t} = 0$. Following Cattaneo et al. (2019) and switching to the notation of the Rubin (2005) causal model, the conditional expectation of our outcome variable given the running variable can be illustrated as

$$E[\,Y_{i,t} \mid margin_{i,t}] = \begin{cases} E[\,Y_{i,t}(0) \mid margin_{i,t}] & \text{if } margin_{i,t} < 0 \\ E[\,Y_{i,t}(1) \mid margin_{i,t}] & \text{if } margin_{i,t} \geq 0 \end{cases} \tag{3}$$

where $Y_{i,t}$ is defined more precisely below. Note that we only observe the average outcomes given by

[24] Section 3.3 describes several examples of close-election RD designs in the Indian setting.

equation (3) – counterfactual forest cover changes for constituencies electing a criminal candidate ($E[\,Y_{i,t}(0)\,|\,margin_{i,t}]$ if $margin_{i,t} \geq 0$) or constituencies choosing a non-criminal representative ($[\,Y_{i,t}(1)\,|\,margin_{i,t}]$ if $margin_{i,t} < 0$) are unknown. This highlights the fundamental problem of causal inference in our RD setting, embodied in the fact that no control and treatment constituency share the same value of $margin_{i,t}$ (Cattaneo et al. 2019). More importantly, this also means that valid estimation of the theoretical sharp RD treatment effect

$$\phi = E[\,Y_{i,t}(1) - \,Y_{i,t}(0)\,|\,margin_{i,t} = 0] \tag{4}$$

rests on the pivotal assumption that average potential outcomes $E[Y_{i,t}(0)|margin_{i,t} = x]$ and $E[Y_{i,t}(1)|margin_{i,t} = x]$ are continuous in the running variable at $margin_{i,t} = 0$, rendering control constituencies just below the threshold comparable to treatment constituencies marginally above the cut-off (Lee 2008). Again, the importance of this assumption calls for careful consideration of its validity, which is the subject of Section 6.2. If this holds, then we have:

$$\phi = E[\,Y_{i,t}(1) - \,Y_{i,t}(0)\,|\,margin_{i,t} = 0] = \lim_{x\downarrow 0} E[\,Y_{i,t}\,|\,margin_{i,t} = x] - \lim_{x\uparrow 0} E[\,Y_{i,t}\,|\,margin_{i,t} = x] \tag{5}$$

That is, if control and treatment constituencies are comparable close to the cut-off, the causal effect of electing a criminal candidate on forest cover growth at $margin_{i,t} = 0$, ϕ, is identified as the difference between the limits of the observed average outcomes for each group as the victory margin converges to the threshold (Hahn et al. 2001; Cattaneo et al. 2019).[25]

In our main specification, we estimate ϕ for our sample nonparametrically by examining the following local linear approximation:

$$Y_{i,t+\mathbf{j}} = \beta_1 + \beta_2 criminalwinner_{i,t} + \beta_3 margin_{i,t} + \beta_4\, criminalwinner_{i,t} * margin_{i,t} + \Omega_{i,t+\mathbf{j}} + \mu_{i,t+\mathbf{j}} \tag{6}$$

$$\forall\, \mathbf{j} \in [1,4] \text{ and } i \text{ where } margin_{i,t} \in [-h, h]$$

In equation (6), $Y_{i,t+\mathbf{j}}$ is the annual growth rate of average forest cover in constituency i from time $t + j - 1$ to time $t + j$ and the other variables are defined as per equation (1) and (2). Note the time subscripts: we let t be the year of election (ranging from 2000 to 2010) and j be the number of years after the election within the five-year term (ranging from 1 to 4), to properly reflect the panel structure of our baseline sample (ranging from 2004 to 2014). Motivated by the occasionally large annual variation observed in the VCF data discussed in Section 4.3, we control for state-year fixed effects with the term $\Omega_{t+\mathbf{j}}$.[26] We also extend the main specification by including a vector $\Gamma_{i,t}$ of available constituency- and candidate level controls to reduce variation and improve the efficiency of our estimates, even though this is not required for identification (Cattaneo et al.

[25] As evident from equation (5), this parameter should be interpreted as a local average treatment effect (or LATE) (Imbens and Angrist 1994; Lee and Lemieux 2010). We discuss this further in Section 7.2.

[26] In general, constituencies within a state follow the same election cycle. It follows that, beyond controlling for state- or year-specific shocks such as noise induced by the satellite-based VCF measure (e.g. monsoons, typhoons or smog spill-overs from neighbouring cities), the inclusion of year-state fixed effects should absorb shocks specific to a certain year of term.

2019).[27] β_2 corresponds to the population parameter ϕ of equation (5) and is our focal coefficient. It estimates the causal effect of electing a criminal candidate on forest cover growth as the difference in intercepts at $margin_{i,t} = 0$ of linear approximations of the regression functions in equation (3), fitted separately on each side of the threshold. This emphasizes an essential future of the sharp RD design – it relies heavily on extrapolation and thus estimates depend on the functional form chosen to approximate the fit of the data (Cunningham 2021). Thus we extend our analysis beyond linear approximations to parametric estimations with higher order polynomials in Table A5 as a robustness check (Lee and Lemieux 2010; Card et al. 2014).[28] We weigh observations according to a triangular kernel function in all estimations, giving most importance to the closest elections (observations closest to the threshold).[29]

Given the balanced panel data structure of our baseline sample, we observe forest cover growth rates annually throughout the entire five-year election term while the running variable $margin_{i,t}$ naturally does not vary within a term.[30] Note that, following Prakash et al. (2019), the dependent variable is excluded for the election year since it could be influenced by the previous incumbent. Hence, all specifications include 4 year-of-term-variant annual forest growth observations and as many year-of-term-invariant observations of $margin_{i,t}$ and $criminalwinner_{i,t}$ for each election year t. The corresponding error terms $\mu_{i,t+j}$ are clustered on the constituency level, since forest cover growth is expected to be correlated over the course of an election term within a constituency.

Point estimation and inference for β_2 is conducted only based on those observations that lie within a range, or bandwidth, h around the threshold, chosen optimally according to the procedure developed by Calonico et al. (2014; 2020). In our main specification, this yields $h = 10.491$ percentage points, although the optimal bandwidth varies with both the choice of controls and the dependent variable.[31]

Since both point estimates and inference in local polynomial approximations tend to be sensitive to the bandwidth used, this choice is critical for the analysis (Cattaneo et al. 2019). Hence, we investigate how our estimates vary over a range of different bandwidths in Section 6.2.5 and dedicate the remainder of this section to describing valid inference procedures given the nature of the particular identifying assumption governing our close-election RD design.

[27] Note that any such covariate need to be continuous at the threshold of our running variable for valid estimation of ϕ (Cattaneo et al. 2019). See Section 6.2.1 for more details.

[28] Conceptually, this translates to the following specification:
$Y_{i,t+\mathbf{j}} = \beta_1 + \beta_2\, criminalwinner_{i,t} + \beta_3\, f(margin_{i,t}) + \Omega_{i,t+\mathbf{j}} + \mu_{i,t+\mathbf{j}} \;\; \forall\, \mathbf{j} \in [1,4]$
where $f(margin_{i,t})$ is a continuous control function of polynomial order n in the running variable on each side of the threshold, and all other variables correspond to equation (6).

[29] As per Calonico et al. (2020), a triangular kernel is the most feasible choice for bandwidth calculations in terms of point estimator optimality. In addition, since it assigns lower weights to observations further away from the cut-off, a triangular kernel may be a more appropriate choice in the presence of potential outliers (Cunningham 2021).

[30] In other words, RD estimations in this paper include repeated observations of the running variable. Our Stata command of choice, the rdrobust package developed by Calonico et al. (2017), adjusts for this by assuring that the optimally calculated bandwidths contain a minimal number of unique observations (Calonico et al. 2020).

[31] In accordance with Cattaneo et al. (2019) and for the sake of consistency, we impose the optimal bandwidth procedure for each covariate in the continuity tests of Section 6.2.1.

5.2 Local Polynomial Inference and the Continuity-Based RD Design

A recent debate over the validity of the assumptions invoked to identify the treatment effect in close-election RD designs highlights an important methodological distinction, with fundamental bearing on our analysis and especially our inference procedures.

In short, a critique by Caughey and Sekhon (2011) argued that, at least in the US context, predetermined characteristics for candidates in close elections could not be considered as-if-random. In an analysis of over 40,000 close races, Eggers et al. (2015) countered with the conclusion that this only seemed to be the case for a small subset of US elections – and thus that the close-election RD assumptions are likely to be valid in most applications (Cunningham 2021). The essence of this debate, however, is arguably not about the validity of the identifying assumption but rather about *which* assumption one invokes. In the words of de la Cuesta and Imai (2016), there are two different assumptions on which an RD design may rest – the above mentioned *as-if-random* assumption and the much less stringent *continuity* assumption. While the former essentially implies that victory of a certain candidate is randomly assigned near the threshold, this is not necessarily a consequence of the latter. In other words, an RD design resting merely on the continuity assumption need not be invalid if local randomization of treatment around the cut-off is violated (Cattaneo et al. 2015; de la Cuesta and Imai 2016). The first key point to be made here is thus that, since local randomization is not required for an RD design and given the ongoing debate concerning the validity of this stricter assumption in close-election settings, our identification strategy rests only on the continuity assumption.[32]

Crucially, our second important point here is an implication of choosing the continuity-based RD design – namely that conventional inference procedures are invalid for point estimates retrieved using our optimal bandwidths.[33] Since h is chosen as to minimize the mean squared error (or MSE) of the point estimator β_2, effectively optimizing the bias-variance trade-off intrinsic to the bandwidth decision, it nonetheless assumes that there is a non-zero bias in the polynomial fit (Lee and Lemieux 2010; Cattaneo and Vazquez-Bare 2017; Cattaneo et al. 2019).[34] This emphasizes the nonparametric nature of our local linear approximation. In stark contrast to this, conventional inference with standard confidence intervals conducted on the MSE-optimal estimator β_2 would implicitly assume null bias in the fit, hence viewing equation (6) as a parametric model correctly capturing the true relationship between potential outcomes and $margin_{i,t}$ (Cattaneo et al. 2019). Since this is obviously both unreasonable and inconsistent, we adopt the robust bias-corrected inference procedure proposed by Calonico et al. (2014) when analyzing our main specification based on the MSE-optimal bandwidth h.[35] While this procedure is both valid and the most convenient

[32]This does not mean that our RD design is less sensitive to threats to identification, including discontinuities in potential outcomes, manipulation or heaping in the running variable (Cunningham 2021). In Section 6.2, we test for these and other potential violations.

[33]In line with Cattaneo et al. (2019), conventional inference in this context refers to the same confidence intervals retrieved from a standard parametric least-squares estimation.

[34]While reducing the bandwidth improves the fit of the local polynomial approximation and hence lowers the bias of the point estimate, it also increases its variance since fewer observations are used. Conversely, a larger bandwidth increases the bias but reduces the variance of the estimate. Since the MSE of β_2 equals its squared bias plus its variance, minimizing the MSE corresponds to optimizing this trade-off (Cattaneo et al. 2019).

[35]Although robust bias-corrected confidence intervals enable both point estimation with optimal properties and valid inference using the same MSE-optimal bandwidth (and thus, conveniently, the same number of observations), the statistical properties of these confidence intervals cannot be considered optimal in terms of their coverage error. For more information on this and how the robust bias-corrected confidence intervals are constructed, see Cattaneo et al. (2019) and Calonico et al. (2020).

for our purposes, it is suboptimal in terms of the coverage error (CER) of its confidence intervals (Cattaneo et al. 2019). Hence, to strengthen our statistical conclusions, we conduct optimal inference based on an alternative bandwidth designed by Calonico et al. (2018) to minimize the CER in Table A6.[36] Nevertheless, there is suggestive evidence that robust-bias corrected inference is credible in our close-election context: Hyytinen et al. (2018) compare experimental results from tied elections resolved through randomization with corresponding MSE-optimal RD estimates using robust bias-corrected inference and conclude that these are aligned.

6 Results

In this section we present evidence showing that constituencies electing a criminal candidate to the Legislative Assembly experience significantly lower rates of forest cover growth in the years following the election, compared to where a non-criminal representative is elected. We commence with providing the main results of our analysis, followed by an extensive exercise to verify the validity of the RD design in Subsection 6.2. In Subsection 6.3, we perform further robustness checks and differentiate our estimates based on various subsamples.

6.1 Primary Findings

Table 2: Main Results

	(1)	(2)	(3)	(4)
Criminal Winner	-4.135***	-3.545*	-3.137**	-3.806***
	(1.532)	(2.002)	(1.537)	(1.376)
Robust P-value	0.002	0.052	0.037	0.003
Robust CI	[-7.663, -1.657]	[-7.820, 0.026]	[-6.221, -0.195]	[-6.823, -1.428]
Eff. no of obs.	1308	1288	1164	996
Opt. Bandwidth	12.265	12.083	10.491	8.542
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

[36] It should be noted that this alternative approach comes at the cost of a suboptimal point estimator (Calonico et al. 2018). Hence, following Cattaneo et al. (2019), inference statistics retrieved from using the CER-optimal bandwidth should be disconnected from their respective point estimators and rather be applied to the MSE-optimal estimates. As a consequence of using two different bandwidths for inference and point estimation, optimal statistical inference leads to the disadvantage of having to use a different number of observations for each purpose. See Cattaneo et al. (2019) for more details.

The primary finding of this paper, the RD effect of electing a criminal candidate on annual forest cover growth, is presented in Table 2 and Figure 1. Each column of Table 2 provides point estimates and robust bias-corrected inference retrieved from local linear approximations using a triangular kernel and the MSE-optimal bandwidth choice of Calonico et al. (2014; 2020). Our focus generally lies on the latter two estimations – where column (3) includes state-year fixed effects and corresponds to our main specification in equation (6), while the fourth column introduces additional controls. Note that valid estimation of our target parameter in the non-parametric continuity-based RD design requires these controls to be continuous at the threshold (Cattaneo et al. 2019).[37] The first two columns are included primarily for the sake of comparability, and should thus be interpreted with caution.[38] The optimal bandwidth and thus the number of observations included within this window varies across estimations since it is a function of variation partialled out upon inclusion of covariates.

Across all estimations, we consistently retrieve negative point estimates of similar magnitude around 3 – 4 percentage points which reach statistical significance. The estimation summarized in column (3) generates a coefficient on $criminalwinner_{i,t}$ which says that among constituencies in the neighbourhood of the threshold $margin_{i,t} = 0$, those electing a criminal candidate experience a yearly growth in average forest cover over the election term which is approximately 3.1 percentage points lower than in constituencies where a non-criminal candidate won. This estimate is statistically significant on the 5 % level and the robust bias-corrected confidence interval does not contain zero. The inclusion of additional controls in column (4) yields a slightly larger significant growth decline of about 3.8 percentage points. As expected, the precision of the covariate-adjusted estimate is marginally higher, but this comes at the cost of a narrower bandwidth. Recognizing that conclusions from Table 2 draw on suboptimal statistical inference, we proceed with estimating the same specifications using bandwidths designed to minimize the coverage error of the confidence intervals. These results are presented in Table A6 and confirm our conclusions above – all relevant point estimates remain statistically significant under CER-optimal inference.[39] Our main finding is also robust to parametric estimation of the effect based on the entire baseline sample, controlling for up to fourth degree polynomial functions of the running variable. As evident from Table A5, this consistently yields negative and significant effects, albeit smaller in magnitude for the point estimates corresponding to column (3) and (4) of Table 2.[40] With these assurances in mind, we proceed with a visualization of the impact of electing a criminal candidate on forest growth.

[37]The following additional covariates were included, based on their relevance and data availability: Log average forest cover in the election year, log number of electors, log number of valid votes, log constituency area and constituency seat reservation type (scheduled castes, SC, or scheduled tribes, ST, if applicable), as well as gender, log assets and type (ST or SC) of each of the top two candidates. On top of these, we also control for the age of the runner-up but *not the winner's age* since this variable is discontinuous at the cut-off (see Section 6.2.1).

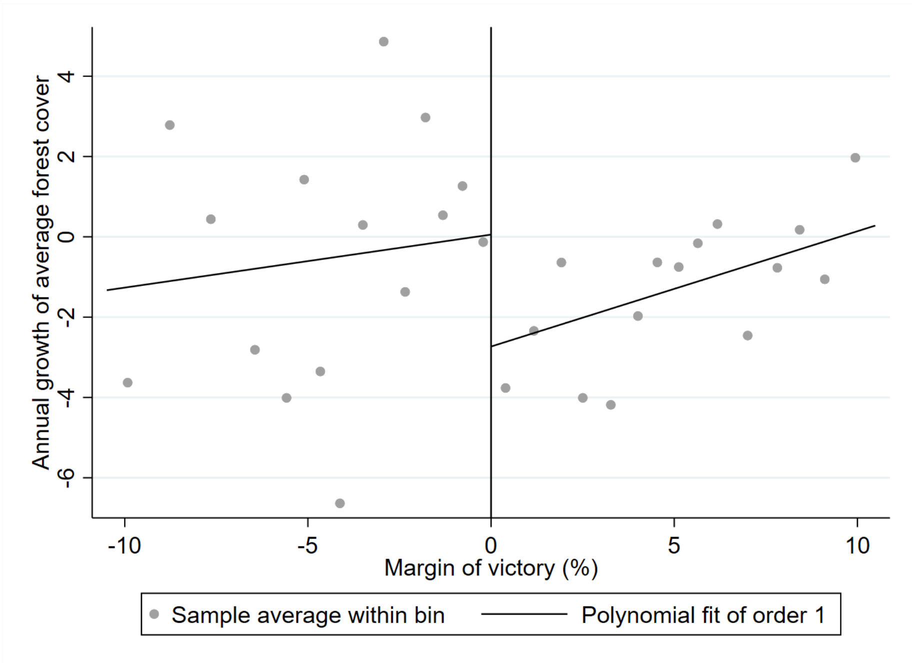
[38]By excluding year fixed effects, the simple estimation of column (1) ignores the high annual variability of the VCF discussed in Section 4.3. Consequentially, we cannot in this case reject that results might be driven by, for example, time-specific noise in the satellite measure. Likewise, a similar argument applies to column (2) where the exclusion of state fixed effects fails to account for the possibility that such unobserved noise may vary across state borders.

[39]Note that point estimates retrieved using CER-optimal bandwidths are not appropriate since they have too much variability relative to its bias (Calonico et al. 2018). Hence, as per Cattaneo et al. 2019, focus lies on the inferential results of Table A6 and these should be applied to the point estimates of Table 2. The only case where we cannot reject the null hypothesis with CER-optimal inference is for the less relevant local quadratic estimate when only year fixed effects are controlled for.

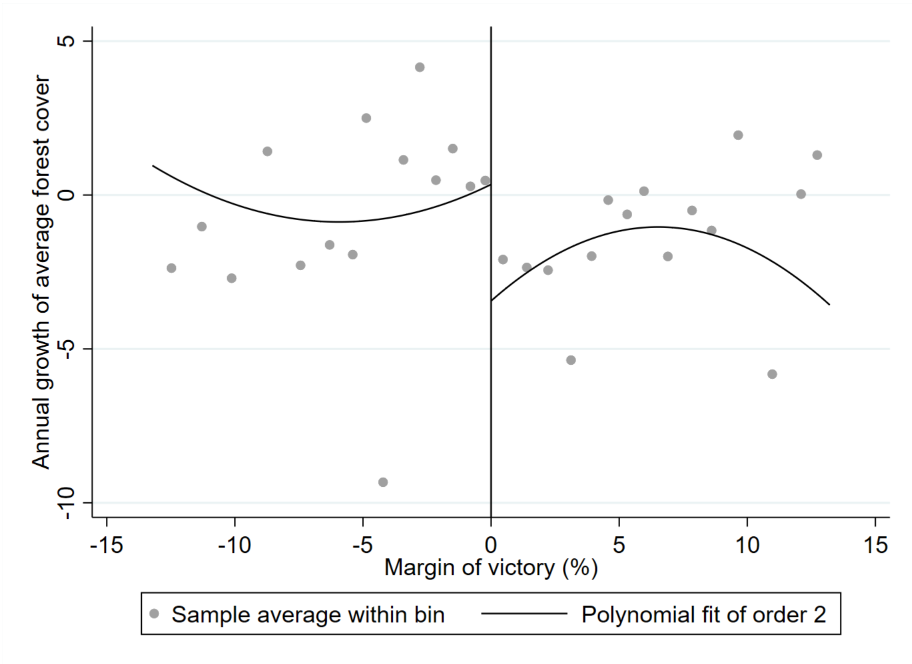
[40]For the same reasons discussed in Section 5.2, conventional inference is naturally more appropriate for parametric estimations and is thus what we present in Table A5.

Figure 1: Graphical Illustration of Main Results

(A) Linear Fit



(B) Quadratic Fit



Notes: The figure illustrates the local RD effect of electing a criminal politician on annual average forest cover growth graphically by plotting the conditional expectation of annual average forest cover growth, given the margin of victory of a criminal candidate. The curves represent linear (A) and quadratic (B) fits of the data within the corresponding MSE-optimal bandwidth, fitted separately on each side of the threshold. The optimal bandwidth, given by the selection procedure of Calonico et al. (2014; 2020), is 10.491 for (A) and 13.222 for (B). The dots depict sample averages within 15 quantile-spaced bins on each side, assuring that each bin contains roughly the same number of unique observations and adjusting for mass points in the running variable (see Calonico et al. (2015) for details).

Panel A of Figure 1 illustrates the local linear RD effect retrieved from the estimation in column (3) of Table 2 graphically, depicted as a sharp vertical decline at the threshold where the margin of victory equals zero. Panel B is analogous to the point estimate in column (3) of Table A7 in the appendix, which presents results from a quadratic rather than linear local estimation of equation (6) and yields a similar significant growth reduction of roughly 3.4 percentage points. Notably, the covariate-adjusted estimate is also robust to a quadratic fit – the coefficient in column (4) of Table A7 remains significant and is slightly larger in absolute value. In Figure 1, we plot observed annual forest growth outcomes conditional on the victory margin of the criminal candidate for constituencies within the corresponding MSE-optimal bandwidth.[41] As evident from the figure, the MSE-optimal bandwidth is slightly larger in the local quadratic estimation. In the proximity of the threshold, constituencies with a negative victory margin barely elected a non-criminal candidate while points to the right of the cut-off represent constituencies where a criminal candidate barely won. Each point in the figure represents the sample average within a certain interval of the running variable.[42]

Three main features of Figure 1 should be highlighted. The first is that in close elections, the effect of a criminal winner is sizeable and, as evident from Table 2, significant. Contrasting this against the average yearly growth among all constituencies within the bandwidth suggests that our estimate represents approximately a threefold growth reduction.[43] Bear in mind, however, that any such comparison should be made for the sole purpose of getting a grasp of the relative magnitude of the effect due to the inherently extrapolative nature of our estimate (Lee and Lemieux 2010). Secondly, both the linear and quadratic plots show that among the constituencies with the closest elections, those electing a non-criminal candidate seem to experience either slightly positive or near-null growth while those barely choosing a criminal candidate experience clearly negative forest cover growth rates. This could suggest that our estimated effects involve deforestation in these constituencies.[44] Relatedly, our final point regarding Figure 1 is that the above-mentioned tendencies do not necessarily apply to observations further away from the cut-off. Emphasizing once more the local nature of our estimates, any causally identified effect need not be present among constituencies where the margin between the top two candidates was larger. We leave this discussion to Section 7.2 and now turn to examining whether the results presented above may be interpreted as causal.

[41]Confidence intervals are deliberately excluded from these plots, due to two reasons. First, the purpose of the plots is to visualize the magnitude of the relevant point estimate and not its inferential properties – the latter are presented in the corresponding table. Second, the plotting tools of the rdrobust package only allows for graphing confidence intervals for each bin, which do not correspond to those of tables Table 2 and Table A7 and hence give an inaccurate picture.

[42]Data-driven bin length selection procedures assure that each of the 15 bins on each side of the cut-off consists of approximately the same number of unique observations, hence reflecting the distribution of the data within the window (Calonico et al. 2015; Cattaneo et al. 2019). Average bin length in terms of the running variable is about 0.7 % for Panel A and about 0.9 % for Panel B of Figure 1.

[43]Since the mean annual growth in average forest cover for our sample period within the MSE-optimal window of 10.491 % was about - 0.92, the point estimate of column (3) in Table 2 corresponds to a relative decrease of about 340 % (3.137 / 0.92). Note that the relative magnitude of the effect naturally depends on the benchmark used. Albeit increasing slightly when the bandwidth is widened, when instead the median growth is used, or when observations are residualized with respect to year-state fixed effects (since these generate benchmarks that tend to be smaller in absolute value), the relative effect size is roughly consistent across various benchmarks.

[44]We discuss this further in Section 7. See Sections 2.1 and 4.3 for comparisons between forest cover loss (degrowth) and deforestation.

6.2 Validity

The identifying assumption of our empirical strategy states that there should be no discontinuities in expected potential outcomes at the threshold value of the running variable. This is, as mentioned in Section 5, fundamentally untestable due to the unobservability of counterfactual outcomes (Cunningham 2021). Nevertheless, there are several ways to indirectly strengthen the credibility of the continuity assumption (Lee and Lemieux 2010; Cattaneo et al. 2019). This subsection conducts five tests in an attempt to confirm the validity of our continuity-based close-election RD design: we test the continuity of observed covariates in Subsection 6.2.1, Subsection 6.2.2 examines the density of our running variable around its threshold and Subsection 6.2.3 investigates the sensitivity of our results to excluding the closest elections. Moreover, Subsection 6.2.4 looks at our results at artificial thresholds and Subsection 6.2.5 concludes these validity tests by examining the sensitivity of our estimates to the choice of bandwidths. Based on these, our general conclusion is that the identifying assumption holds, allowing for a causal interpretation of the observed reduction in annual average forest cover growth following the election of a criminal candidate.

6.2.1 Continuity Tests

A major threat to causal identification of our estimated effects arises when there are systematic differences between observations just below and just above the threshold. In other words, we cannot credibly isolate and attribute the decline in forest growth to the election of criminal candidates if observed covariates that either could or should not have been affected by treatment display discontinuities at the cut-off. Note that for a continuity-based RD design, this is not equivalent to the experimental analogue of balance in pre-treatment covariates – it is merely required that the conditional expectation function of each covariate is sufficiently smooth (de la Cuesta and Imai 2016; Cattaneo et al. 2019).[45] Thus, we conduct continuity tests for observed constituency- and candidate-level characteristics, adopting the same local linear approach which we use for our main dependent variable.[46]

Results from these continuity tests are provided in Table 3. Based on robust-bias corrected inference, we cannot reject the null hypothesis of no discontinuity for any of the covariates at the 5 % significance level while only 1 of 19 variables displays a significant discontinuity at the 10 % level.[47] This also holds when we instead employ CER-optimal inference, as evident from Table A10.[48] Overall, our continuity tests show no apparent discontinuities in predetermined covariates

[45] In this context, simple t-tests of differences in means are not appropriate since they implicitly test for local randomization and rely on conventional inference (see Section 5.2).

[46] Covariates are chosen based on their relevance and availability, and may be either predetermined or placebo outcomes (Cattaneo et al. 2019). Except for the number of electors and valid votes, which are presumed to be unaffected by treatment, we consider constituency-level characteristics predetermined. The 10 observed candidate-level characteristics could be of either type – an implicit purpose of the test is to determine which one. Furthermore, seven variables are excluded from Table 3 since they contain many missing values and hence too few observations within their respective optimal bandwidth to conduct reliable inference. Nonetheless, the descriptive statistics and results for these are provided in Table A8 and Table A9, respectively, and the results do not display any significant jumps at the cut-off.

[47] The main results of Table 2 are robust to controlling for the age of the winner, which may indicate that this coefficient turned out significant on the 10 % level by chance. These results are available upon request.

[48] In Table A10, only the log number of valid votes is significant (on the 10 % level). Again, our results remain robust to the inclusion of this control (under both inference procedures). Output from these estimations are available upon request.

or placebo outcomes, neither among constituency- nor candidate-level characteristics, which speaks for the validity of the design. Furthermore, the results suggest an alternative, yet less conclusive, interpretation – namely that these variables do not appear to confound the causal relationship between criminal politicians and forest growth.

Table 3: Continuity Tests

Panel A	Constituency-level Characteristics				
	RD Estimate	Robust P-value	Robust CI	Eff. no of obs.	Opt. Bandwidth
Forest Growth (t-1)	5.737 (5.538)	0.176	[-3.363, 18.346]	259	8.971
Forest Growth (t-2)	-2.396 (6.752)	0.721	[-15.647, 10.821]	263	9.081
Forest Growth (t-3)	-1.516 (7.543)	0.672	[-17.973, 11.594]	322	12.11
Log Const. Area	0.093 (0.280)	0.628	[-0.414, 0.685]	266	9.272
Const. SC Reserved	0.039 (0.084)	0.688	[-0.131, 0.199]	319	11.96
Const. ST Reserved	0.030 (0.056)	0.771	[-0.093, 0.126]	274	9.602
Log Forest Cover (t)	0.001 (0.071)	0.857	[-0.126, 0.151]	292	10.63
Log Electors	0.005 (0.048)	0.820	[-0.082, 0.104]	308	11.61
Log Valid Votes	-0.051 (0.052)	0.181	[-0.171, 0.032]	216	7.053
Panel B	Candidate-level Characteristics				
Winner Gender	0.037 (0.067)	0.396	[-0.075, 0.189]	302	11.11
Runner-up Gender	-0.031 (0.055)	0.665	[-0.133, 0.085]	273	9.473
Winner Age	-4.389* (2.283)	0.073	[-8.564, 0.386]	328	12.43
Runner-up Age	0.647 (2.664)	0.892	[-4.860, 5.585]	309	11.69
Winner Education	0.024 (0.495)	0.858	[-0.881, 1.057]	336	13.43
Runner-up Education	-0.497 (0.530)	0.230	[-1.674, 0.402]	304	12.10
Winner Log Assets	-0.634 (0.445)	0.135	[-1.538, 0.207]	273	9.464
Runner-up Log Assets	-0.479 (0.381)	0.146	[-1.302, 0.193]	242	8.203
Winner Type (ST/SC)	0.090 (0.094)	0.419	[-0.108, 0.260]	332	12.54
Runner-up Type (ST/SC)	0.106 (0.093)	0.287	[-0.083, 0.281]	337	12.68

Notes: Continuity tests for 19 different constituency- and candidate-level characteristics, based on local linear estimations within the MSE-optimal bandwidth chosen separately for each variable. All tests use a triangular kernel function and control for election year-state fixed effects. Each row presents point estimate for the covariate, together with the corresponding robust bias-corrected standard errors, confidence intervals and p-values as well as the effective number of observations. The latter is comparatively low since all variables are invariant within election terms and thus only included for the election year. Robust bias-corrected standard errors are presented in parentheses. Asterisks denote significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

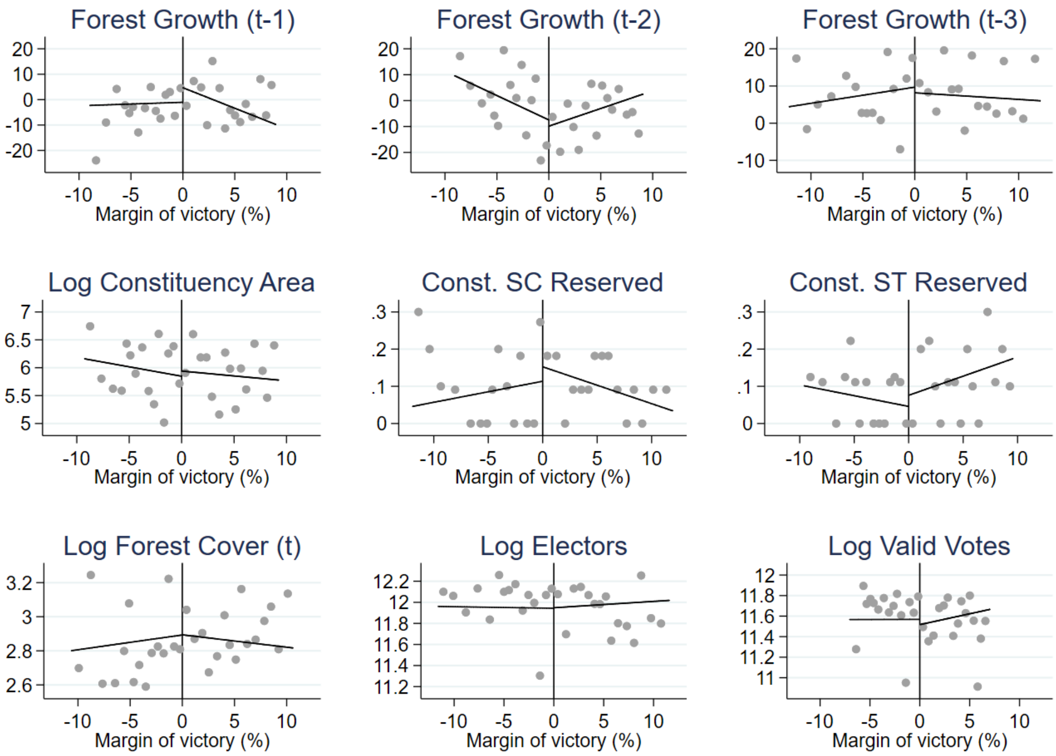
We depict constituency-level results from the continuity tests in Table 3 visually in Figure 2, which plots the conditional expectation of each covariate against the running variable. RD plots for the rest of the variables are found in Figure A2. Encouragingly, the general pattern in these graphs reflects the conclusions of our formal continuity tests – there are no significant discontinuities at the threshold for any of the examined covariates.[49] With this said, we move on to examining the density of our running variable around the cut-off.

6.2.2 Density of the Running Variable

Another cause of concern for the validity of our close-election RD design lies in whether manipulation of the running variable is possible. In the context of this paper, this is mainly exemplified by electoral fraud and vote manipulation by candidates or interest groups since, especially in close elections, criminal candidates may have an incentive intervene in election procedures (Prakash et

[49] Although some of the tested variables may seem to, at a glance, display a minor discontinuity at the cut-off, our formal statistical tests do not reject the null hypothesis of continuity in any case except for the age of the winner. Again, since results are robust to the latter, we conclude that any visually noticeable jumps are negligible. On this note, all plots in Figure 2 and Figure A2 deliberately excludes confidence intervals for the same reason as in Figure 1.

Figure 2: Continuity Tests - Constituency-level Characteristics



Notes: The figures show a local linear approximation of each observed constituency-level characteristic conditional on the margin of victory of a criminal candidate, fitted separately on each side of the threshold within the corresponding MSE-optimal bandwidth for each variable. The vertical difference in intercepts at the threshold is identical to the point estimates in the formal continuity test, which uses a triangular kernel function and control for election year-state fixed effects. The dots depict sample averages within 15 quantile-spaced bins on each side, assuring that each bin contains roughly the same number of unique observations.

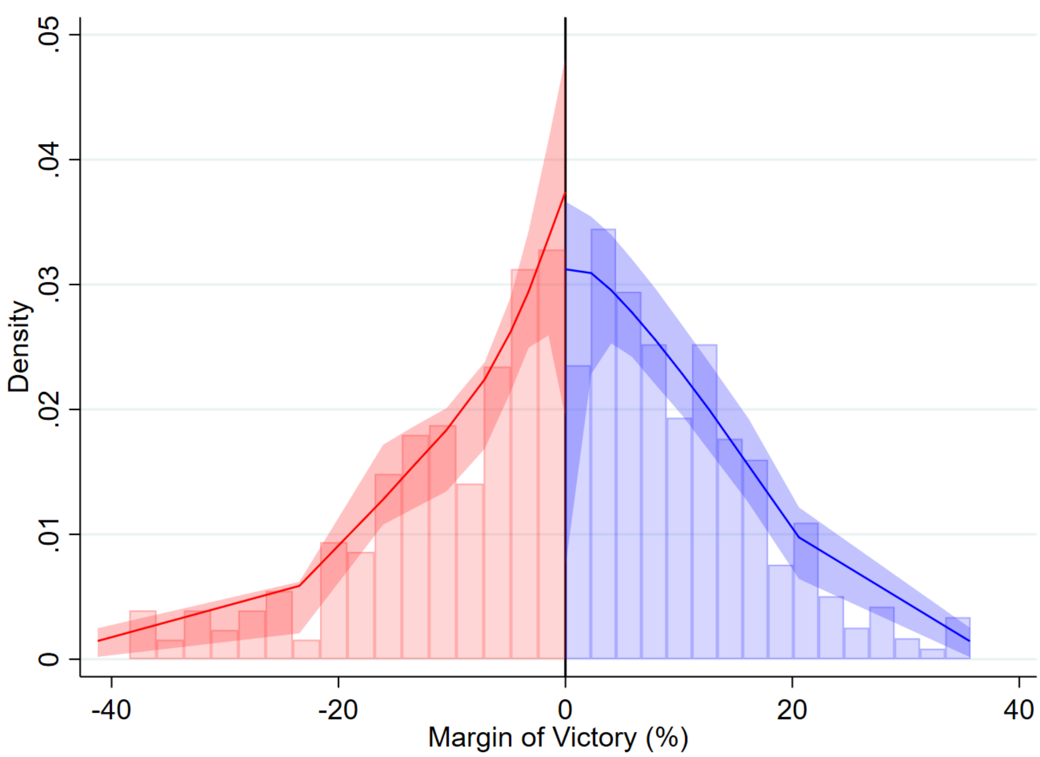
al. 2019). If institutional mechanisms to do so exist, this could imply systematic overrepresentation of criminal politicians in the Legislative Assembly and threaten to invalidate our analysis. While some studies have found indications of strategic manipulation in the Indian electoral setting, others have not (Eggers et al. 2015; Martin and Picherit 2020). The conventional strategy to examine this concern empirically is to test for discontinuities in the probability density function of the running variable at the threshold (Lee 2008; Lee and Lemieux 2010). We implement two versions of this test on the margin of victory of a criminal candidate in the neighbourhood of zero – the first is the standard procedure proposed by McCrary (2008) and the second is the local polynomial approach developed by Cattaneo et al. (2020).

Neither the local polynomial density test nor the McCrary test reject the null hypothesis of continuous density of the victory margin at the threshold, where the former is illustrated in Figure 3. Even though the figure suggests that there are slightly more elections where a criminal candidate barely lost than elections where a criminal candidate barely won, there is no significant difference in the density estimates (the vertical intercepts) at the cut-off – the 95 % robust bias-corrected confidence intervals overlap and the p-value of the formal test is 0.275.[50] Figure A3 depicts the fit

[50] Beyond its statistical insignificance, this tendency is interesting in its own right. While the converse would be

of a corresponding McCrary test, which yields the same conclusion. In sum, we do not find any evidence of candidates strategically manipulating vote shares.

Figure 3: Local Polynomial Density Test



Notes: Graphical illustration of the density of the margin of victory for a criminal candidate, estimated separately on each side of the threshold using a local polynomial approach, a triangular kernel function and an optimally selected bandwidth of about 12.7 with 335 unique observations within this window. The shaded area represents 95 % robust bias-corrected confidence intervals of the estimated density. In the associated histogram, observations within the range of three bandwidths on each side of the cut-off are divided into quantile-sized bins. For consistency, the estimated density is displayed over the entire range of the histogram. The test is implemented using the Stata command rddensity, described in Cattaneo et al. (2019).

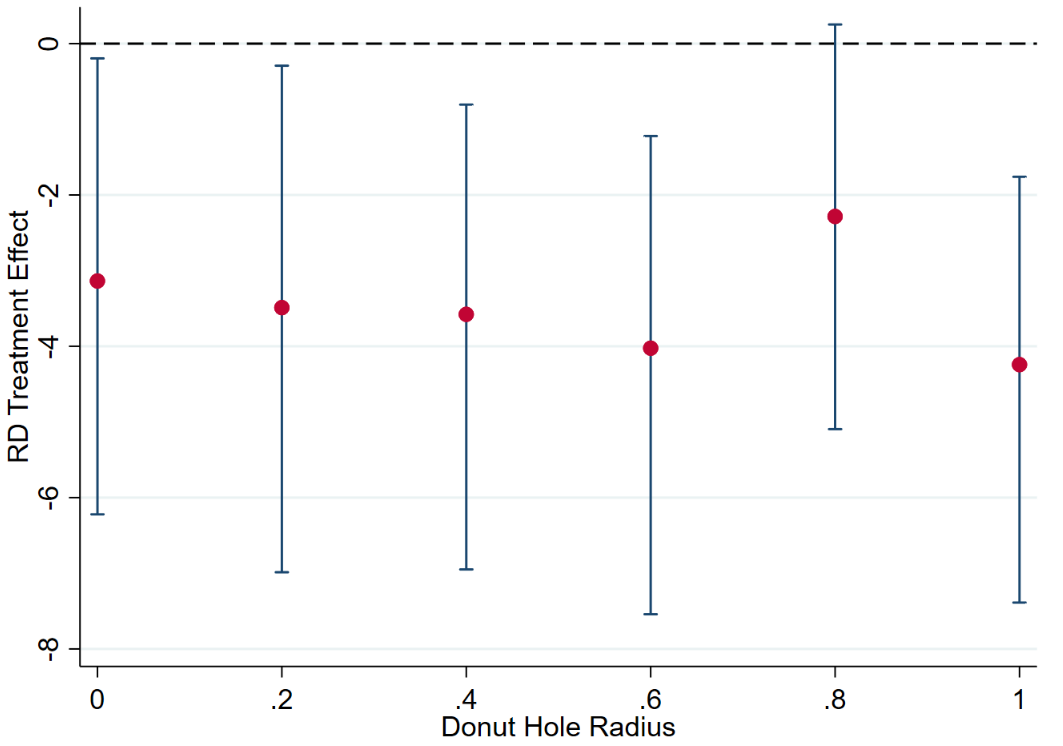
6.2.3 Sensitivity to the Closest Elections

As a complement to the two standard validation exercises above, we conduct three additional tests to provide further credibility to our results. The first of these, commonly referred to as the "donut hole" RD approach, examines the sensitivity of our conclusions to the constituencies with the very closest elections. Observations closest to the cut-off are not only disproportionately affecting our local polynomial fits due to the extrapolative nature of the estimates and our choice of a triangular kernel, but these are also the elections where incentives for manipulation are likely to be the strongest (Cattaneo et al. 2019; Cunningham 2021). Even in the absence of manipulation, however, our estimates may be biased due to non-random heaping on certain values of the running variable near the threshold (Barreca et al. 2016). This may occur, for instance, because of rounding

more worrying, it suggests non-criminals win more often than criminal candidates in the closest elections.

procedures at any stage of the data collection process. The donut hole RD examines these various concerns by excluding observations closest to the cut-off.

Figure 4: Donut Hole Plot



Notes: The figure plots point estimates and 95 % robust bias-corrected confidence intervals from five different donut hole RD estimations. Each estimation excludes constituencies for which the absolute value of the victory margin of the criminal candidate is less than the corresponding value on the horizontal axis. Estimates are retrieved from triangular kernel-weighted local linear estimations within MSE-optimal bandwidths selected separately for each donut hole radius, controlling for year-state fixed effects.

Figure 4 plots point estimates and 95 % robust bias-corrected confidence intervals for five donut hole RD estimations analogous to column (3) of Table 2, with the exception that constituencies where the absolute value of our running variable is less than the corresponding value on the horizontal axis are excluded. The formal results from these estimations are provided in Table A11.[51] In all estimations except for the case of a donut hole radius of 0.8, which nonetheless displays 10 % significance and a similar magnitude, the results remain significant at the 5 % level and lie close to our primary finding.[52] As such, our conclusions are robust to the exclusion of observations close to the threshold.

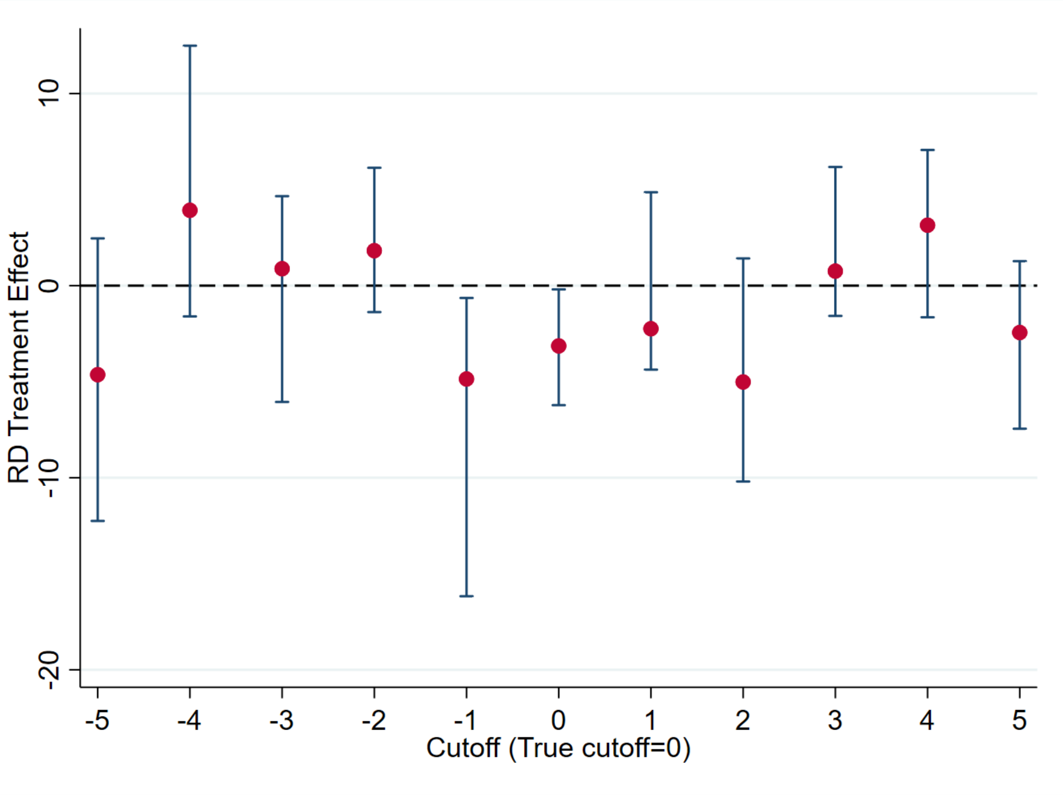
[51] As evident from Table A11, the MSE-optimal bandwidth varies and does not necessarily diminish with the radius of the donut hole. This is primarily a consequence of the dependence of the bandwidth selection procedure on the variation in the remaining observations (Cattaneo et al. 2019).

[52] These results are generally robust to employing CER-optimal inference and to the inclusion of controls. Output from these estimations is available upon request.

6.2.4 Placebo Thresholds

The second additional test we conduct examines whether there are any discontinuities where there should not be any. In other words, we employ the same techniques as in our main specification to test for treatment effects at artificial threshold values of the running variable for the subsample of constituencies which lies on each side of the true cut-off (Cattaneo et al. 2019). As per Imbens and Lemieux (2008), placebo tests in the positive domain of our running variable only compare constituencies where a criminal candidate won the election. Conversely, we only compare constituencies where the non-criminal candidate won when the artificial threshold is negative, to avoid placebo test results being driven by the true discontinuity.[53]

Figure 5: Placebo Plot



Notes: The figure plots point estimates and 95 % robust bias-corrected confidence intervals from ten different placebo tests where the true cut-off is replaced by an artificial value given by the horizontal axis. Estimates are retrieved from triangular kernel-weighted local linear estimations within MSE-optimal bandwidths selected separately for each artificial threshold, controlling for year-state fixed effects. The estimated RD effect for the true cut-off at zero is presented for reference.

Point estimates and 95 % robust bias-corrected confidence intervals from these placebo tests are illustrated graphically in Figure 5, with the corresponding formal results provided in Table A12. For all artificial thresholds but one, point estimates are statistically insignificant at any conventional level and the absolute value of all coefficients is similar in magnitude to the true estimate.[54] Hence,

[53]For example, for a placebo test where the artificial cut-off is set to -2, we only compare constituencies where a criminal candidate lost by more than -2 percentage points to constituencies where the vote share of the criminal candidate was larger than -2 but strictly smaller than 0.

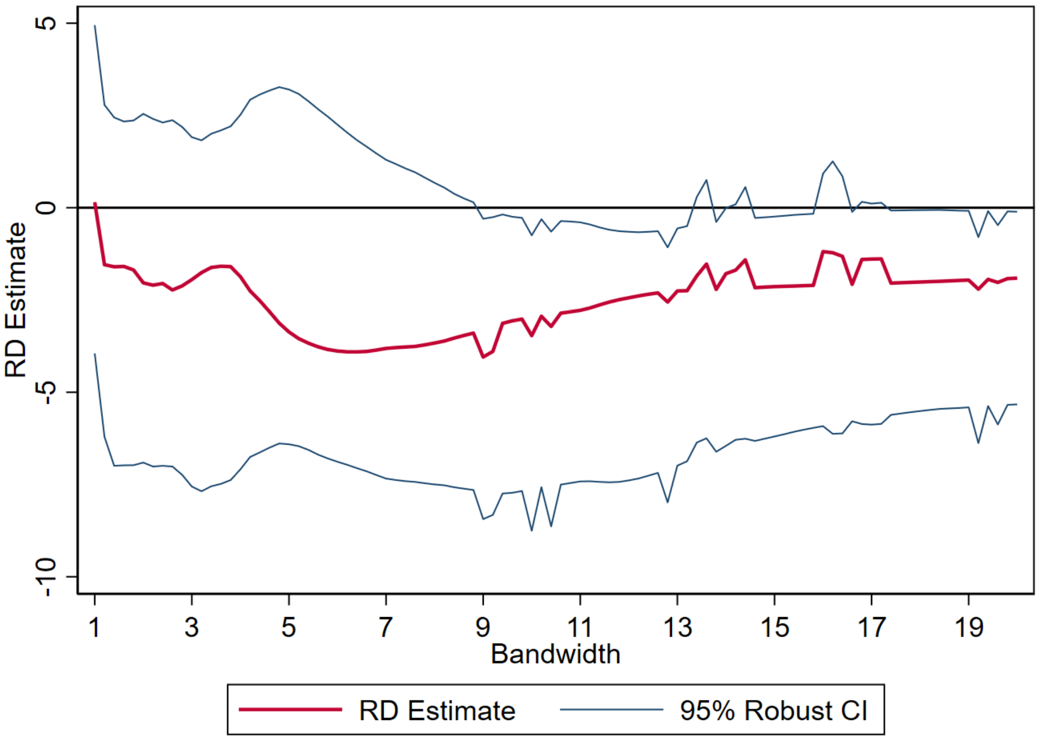
[54]Even though the significant discontinuity at the -1 artificial cut-off may appear worrying, the number of obser-

we generally do not find any noticeable discontinuities in forest growth at these placebo cut-offs.

6.2.5 Sensitivity to the Choice of Bandwidth

Finally, we investigate the sensitivity of our main results to the bandwidth used, since this is one of the most crucial choices in the empirical analysis (Lee and Lemieux 2010). We implement this by estimating our main specification in equation (6) over a broad spectrum of bandwidths.

Figure 6: Sensitivity to Bandwidth



Notes: The figure plots estimated RD effects against different bandwidths, retrieved using triangular kernel-weighted local linear estimations within each bandwidth in two decile increments over a range of 1-20 %. Each estimation controls for year-state fixed effects. Results for bandwidths between 0 and 1 are omitted due to insufficient observations to construct robust bias-corrected confidence intervals.

Results from this procedure are visualized in Figure 6, which plots coefficients and 95 % robust bias-corrected confidence intervals retrieved from estimating the equivalent of column (3) in Table 2 for bandwidths ranging from 1 to 20 in two decile increments.[55] Note that our ambition here is to

vations and thus the statistical power of this placebo test is relatively low – which is not least reflected by the width of the estimated confidence interval (see Table A12). To further examine the robustness of our results and increase the statistical power of the test, we also follow Imbens and Lemieux (2008) and implement the same procedure using instead the median victory margin on each side of the threshold as an artificial cut-off (the median victory margin in the negative domain is about -9.35 and 9.27 for positive values of the running variable). The estimated coefficients from these placebo tests are smaller in absolute value than the true estimate, and both are insignificant with p-values of 0.28 and 0.67, respectively.

[55] While the figure encompasses both the corresponding MSE-optimal bandwidth of 10.491 and the CER-optimal bandwidth of 7.665, note that the confidence intervals in Figure 6 are suboptimal in both senses and not equivalent to those constructed in the main estimations (see Calonico et al. 2020 for details). As such, using the depictions of the plot for inference would be inappropriate.

assess sensitivity of our point estimates to different bandwidths, and not to conduct any inference based on the figure. Given this, one should be cautious when interpreting results for bandwidths that lie too far away from the MSE- or CER-optimal choices – especially close to zero, since these tend to be predominantly driven by the bias-variance trade-off discussed in Section 5.1 (Cattaneo et al. 2019). With that said, it is evident from Figure 6 that our point estimates remain relatively stable across the range. Even though the more imprecise estimates for very narrow bandwidths seem to approach a null effect, Figure A4 displays consistently negative and similarly stable point estimates upon the inclusion of additional controls (the analogue of column (4) of Table 2). At large, this indicates that our results are robust to a wide range of bandwidths.

6.3 Further Robustness Checks and Effect Heterogeneity by Subsamples

This section presents results based on different sample limitations and variable definitions. In Subsection 6.3.1, we estimate our main specifications on alternative definitions of a criminal candidate. Next, we examine whether the effect of electing a criminal politician on forest growth varies over the years of the election term and whether it accumulates over time. Finally, Subsection 6.3.3 conducts our analysis on different subsamples of states. We elaborate on these results when discussing potential mechanisms in Section 7.1.

6.3.1 Redefining Criminal Politicians

We begin by examining the robustness of our main results to alternative treatment indicators, based on more conservative definitions of criminal candidates. This is primarily motivated by the concerns outlined in Section 4.2 – namely that our chosen definition of criminality (reporting at least one criminal charge ahead of the election) might also capture candidates charged with minor crimes or those who have a charge framed against them by their opponent. Since these candidates might differ from "real" criminals and corrupt politicians in characteristics or political motivations with bearing on forest outcomes, we estimate our main specification with and without additional controls using two sets of alternative definitions of criminality. The first test raises the minimum number of crimes a candidate must be accused of committing to be classified as criminal, from the original case of one or more crimes to at least two, three and four crimes, respectively. In the second set of estimations, treatment instead reflects the severity of the crime through three separate definitions. The first two estimations define candidates as criminal if they have been accused of committing at least one crime for which the punishment is at least one or three years of imprisonment, respectively, while the third estimation defines criminal candidates as those accused of at least one serious crime (based on classifications by the ADR, which include cases related to kidnapping, rape or murder).

Table 4: Effect Heterogeneity by Number of Crimes

	Panel A		Panel B		Panel C	
	(1)	(2)	(3)	(4)	(5)	(6)
Criminal Winner	-2.571*	-3.116**	-3.093**	-3.447**	-2.495	-3.719*
	(1.674)	([1.582)	(1.710)	(1.866)	(1.890)	(2.279)
Robust P-value	0.098	0.034	0.050	0.050	0.158	0.092
Robust CI	[-6.054, 0.509]	[-6.458, -0.258]	[-6.700, 0.003]	[-7.308, 0.007]	[-6.371, 1.036]	[-8.307, 0.628]
Eff. no of obs.	940	876	816	664	720	508
Opt. Bandwidth	10.382	9.282	11.382	8.068	11.360	6.390
Year FE	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function and estimated separately for different treatment indicators defined by the number of crimes as follows. Panel A: Candidates are defined as criminal if they are accused of 2 or more crimes and non-criminal if they are not accused of any crime. Panel B: Candidates are defined as criminal if they are accused of 3 or more criminal cases and non-criminal if they are not accused of any crime. Panel C: Candidates are defined as criminal if they are accused of 4 or more criminal cases and non-criminal if they are not accused of any crime. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1

Results for the first test are presented in Table 4. Panel A, B and C correspond to estimations where candidates are defined as criminal if they are accused of at least two, three or four crimes, respectively. In general, these estimates are similar to those of our main analysis, albeit slightly smaller in magnitude. The number of observations within the MSE-optimal bandwidth of each estimation is smaller, hence reducing the power of these tests compared to our main results in Table 2, and some statistical significance is lost. Notably, the covariate-adjusted estimate becomes gradually more negative as we move from the least strict definition in column (2) to the most conservative definition of criminality in column (6). This tendency is not evident among the estimations where only state-year fixed effects are controlled for.

Table 5 provides the estimates retrieved when using the three alternative definitions based on severity of crimes. Here, the first two panels rest on definitions of candidates being criminal if they risk at least one year (Panel A) or three years (Panel B) of imprisonment from at least one crime they are accused of. Panel C, in turn, corresponds to the subsample where criminality is defined as being accused of at least one serious crime. Again, these estimations yield conclusions similar to our main analysis. The forest growth reduction is comparatively larger in magnitude for the covariate-adjusted estimations, especially for more severe crimes.

In sum, the results in Table 4 and Table 5 confirm our previous findings – the estimated effect of a criminal politicians coming to power in close elections remains consistently negative and statistically significant in all but one of these estimations. As such, our conclusions are robust to alternative definitions of a criminal candidate.

Table 5: Effect Heterogeneity by Years of Punishment and Severity of Crime

	Panel A		Panel B		Panel C	
	(1)	(2)	(3)	(4)	(5)	(6)
Criminal Winner	-2.856**	-3.840***	-3.482**	-4.732***	-3.009*	-4.350***
	(1.521)	(1.596)	(1.722)	(1.620)	(1.804)	(1.570)
Robust P-value	0.046	0.008	0.035	0.002	0.085	0.008
Robust CI	[-6.011, -0.048]	[-7.336, -1.081]	[-6.998, -0.249]	[-8.168, -1.818]	[-6.638, 0.431]	[-7.216, -1.062]
Eff. no of obs.	1032	872	796	724	640	484
Opt. Bandwidth	10.322	8.199	10.199	8.955	11.134	7.566
Year FE	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function and estimated separately for different treatment indicators defined by the number of crimes as follows. Panel A: Candidates are defined as criminal if they are accused of at least one crime for which the punishment is at least one year of imprisonment and non-criminal if they are not accused of any crime. Panel B: Candidates are defined as criminal if they are accused of at least one crime for which the punishment is at least three years of imprisonment and non-criminal as before. Panel C: Candidates are defined as criminal if they are accused of at least one major crime as defined by the ADR and non-criminal as before. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1

6.3.2 Effects by Year of Term

Table 6: Cumulative Effect by Year of Term

	Panel A		Panel B		Panel C		Panel D	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Criminal Winner	1.896	1.954	-5.667	-4.794	0.066	-0.921	-11.596**	-15.648***
	(4.439)	(4.416)	(4.213)	(4.009)	(4.159)	(3.967)	(4.835)	(4.854)
Robust SE	[5.246]	[5.127]	[4.946]	[4.625]	[4.859]	[4.569]	[5.615]	[5.624]
Robust P-value	0.833	0.831	0.192	0.217	0.933	0.705	0.024	0.002
Robust CI	[-9.178, 11.385]	[-8.953, 11.145]	[-16.145, 3.244]	[-14.781, 3.350]	[-9.930, 9.117]	[-10.683, 7.226]	[-23.691, -1.679]	[-28.101, -6.056]
Eff. no of obs.	283	261	279	267	340	304	287	247
Opt. Bandwidth	10.071	9.096	9.843	9.371	12.917	11.397	10.382	8.314
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES	NO	YES

NOTES: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. Odd columns include year-state fixed effects and even columns year-state fixed effects and controls. The results demonstrate the effect of electing a criminally accused politician on the cumulative forest growth from the election year to the first year of term (Panel A), second year of term (Panel B), third year of term (Panel C) and last year of term (Panel D). Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses *** p<0.01, ** p<0.05, * p<0.1

Next, we investigate how the estimated impact of electing a criminal candidate on annual forest growth changes over the course of the election term by looking at the cumulative effect during this four-year period. Given that previous findings indicate a latent effect of political corruption on economic outcomes, the latter years of term may be the most important in retrieving the aggregate negative estimate (see for instance Khemani (2004), Prakash et al. (2019), Aidt et al. (2020) and Cisneros et al. (2021).

In Table 6, we present the estimated effect of electing a criminal candidate on cumulative forest growth from the election year to the first year of the election term (Panel A), second year of term (Panel B), third year of term (Panel C) and the fourth and last year of term (Panel D). Bearing in mind that these interpretations rest on relatively few observations, we note that none of the point estimates for the cumulative growth over the first three years of term are statistically significant. By the final year of term, however, the accumulated effect on forest growth is both very large in magnitude and statistically significant, especially upon the inclusion of additional controls. Notably, the coefficients are even positive for the year following the election. Overall, these results suggest that on average, the decline in forest cover following the election of a criminal politician is not immediate - instead, it seems to accumulate over time.

6.3.3 Effects by Level of Corruption and Development

Finally, we examine whether the effect of electing a criminal politician on forest growth varies depending on a set of state characteristics presumed to reflect the quality of institutions and level of development. The negative impact of political corruption, and deforestation in itself for that matter, may be accentuated in states with weaker institutions, lower levels of socio-economic development and widespread corruption (Culas 2007; Cheng and Urpelainen 2019). Following Prakash et al. (2019), we test this by estimating our main specifications on three different sets of states. These are: states classified as least developed by the Indian Ministry of Finance, states classified as highly corrupt by Transparency International India (TII) and states historically viewed as having weak institutions and low levels of economic development (often referred to as the "BIMAROU" states).

Results from these estimations are presented in Table 7. While all point estimates are consistently negative, interpretation of these coefficients warrants caution since the effective number of observations used in all estimations is alarmingly small. With this in mind, we note that the covariate-adjusted estimates for the least developed states (Panel B) and for states classified by TII as highly corrupt (Panel C) are particularly large in absolute value. The latter is almost three times as large in magnitude as the corresponding estimate of our main results in Table 2, and reaches statistical significance on the 1 % level. Estimations on the BIMAROU subsample (Panel A) yield negative and insignificant coefficients. For comparison, Table A13 provides analogous estimates on the complements of these three subsets (that is, for states not classified as BIMAROU, least developed or highly corrupt). Based on a larger number of effective observations, these estimations generally yield negative effects of smaller magnitude than in Table 7 and none of these coefficients is statistically significant.

While the above analysis suggests that the level of development and the degree of corruption in a state may matter for the effect of electing a criminal politician, small sample sizes and the fact that several states fall under more than one of these classifications make it difficult to draw any definite conclusions regarding these results.

Table 7: Effect Heterogeneity by State Level of Corruption and Development

	Panel A		Panel B		Panel C	
	(1)	(2)	(3)	(4)	(5)	(6)
Criminal Winner	-1.994	-3.144	-3.267	-4.751	-3.836	-11.055***
	(2.840)	(1.949)	(3.069)	(2.839)	(3.267)	(0.881)
Robust SE	[3.424]	[2.534]	[3.588]	[3.648]	[4.001]	[1.481]
Robust P-value	0.442	0.403	0.191	0.142	0.439	0.000
Robust CI	[-9.341, 4.080]	[-7.087, 2.846]	[-11.728, 2.338]	[-12.501, 1.797]	[-10.938, 4.744]	[-13.891, -8.085]
Eff. no of obs.	264	140	236	208	152	116
Opt. Bandwidth	9.196	4.539	7.310	5.862	5.953	4.417
Year FE	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function and estimated separately different subsamples defined as follows. Panel A: Only BIMAROU States (Bihar, Chhattisgarh, Jharkhand, Orissa/Odisha, Uttar Pradesh, and Uttarakhand). Panel B: Only Least Developed States as ranked by Ministry of Finance (Arunachal Pradesh, Assam, Bihar, Jharkhand, Orissa/Odisha and Uttar Pradesh). Panel C: Only Highly Corrupt States as ranked by Transparency International India (TII) on index of corruption (Tamil Nadu, Haryana, Jharkhand, Assam, and Bihar). Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

7 Discussion

In the previous section, we provided evidence of a negative effect of a criminal politician coming to power on annual forest cover growth in Indian constituencies where elections were close. Our validity tests generally confirm the causal identification of this effect within the context of competitive elections to Legislative State Assemblies in India. This section elaborates on the how and the why of these findings, as well as their external validity and what we recognize as the major remaining limitations of our analysis. But before delving into these discussions, we want to emphasize two important questions also related to how our findings align with the overarching ambition of this study – which is to shed light on the relationship between corruption and environment in developing countries.

Firstly, we ask ourselves to what extent our measure of criminality among politicians actually translates to political corruption and thus whether we can attribute, at least partially, our causal estimate of forest growth decline to corrupt behaviour. Due to the secretive nature of corruption, it is notoriously difficult to measure and, consequently, a plethora of innovative attempts to capture and estimate its implications for development outcomes have been made.[56] Although it has no clear-cut definition, we believe corruption as "the abuse of entrusted power for private gain", as described by Transparency International (2021), is one of the most comprehensive definitions. Given this, being criminal is not necessarily equivalent to being corrupt, even though the two are likely to be connected on an aggregate regional level. This is supported by the fact that many crime charges in our ADR affidavit data are related to corruption, such as embezzlement of public

[56] See for instance Reinikka and Svensson (2004), Bertrand et al. (2007), Ferraz and Finan (2008), Asher and Novosad (2017) and Banerjee et al. (2018).

funds, fraud and other offenses that fall under the 1988 Prevention of Corruption Act (Prakash et al. 2019). Arguably, this is strengthened by the accentuated negative effects retrieved in Table 5 since these charges tend to be classified as more severe.

On this note, politicians making and breaking laws is a plausible indicator of their general disregard of rules and regulations in a democratic system. Criminal politicians may abuse their power to avert holding offenders, like themselves, accountable for their crimes (Chemin 2008), or provide targeted benefits to certain targeted voter groups in a clientelist fashion (Robbins 2000; Bardhan and Mookherjee 2012). Following Besley's (2005) and Cheng and Urpelainen's (2019) notion of criminality as an indicator of self-interest, we could in this context view our treatment as a proxy for corruption – to the extent that it reliably captures the criminality of a candidate and insofar as criminal politicians are guided by incentives similar to those of corrupt politicians. On the other hand, corruption is an institutional phenomenon which could just as well enable the election of criminal politicians as much as it could arise as a consequence from our treatment. As such, criminals representing constituencies with weaker institutions could be given more leeway to further exert illicit influence – not least over environmental outcomes – in a vicious cycle. This line of reasoning is supported by the comparatively larger growth reductions found in highly corrupt states in Table 7. Nevertheless, without further empirical evidence, we cannot draw any conclusive parallels between our estimated causal effects and corrupt behaviour in India.

The second important question we want to address here relates to whether our causal estimates of annual forest cover growth reductions imply that the election of a criminal politician actually results in deforestation and not merely declined rates of reforestation. Despite the latter being interesting and environmentally significant in its own right, the consequences of deforestation for sustainable development are arguably more dire. In the context of our analysis, constituencies near the threshold can be said to experience deforestation as a result of electing a criminal politician if the estimated local RD effect encapsulates a shift from average annual forest growth to degrowth. Based on the plots of this effect in Figure 1, this seems to be the case visually. Applying our main estimates of about 3-4 percentage point growth reductions to the average forest cover growth rates in the full sample (about 1.89 % annually, as given by Table 1), as well as to most of the corresponding averages for constituencies within our range of optimal bandwidths, also suggests a deforestation effect. Furthermore, the estimated total growth reductions over the entire election term amount to roughly 11.5-15.5 percentage points as per Table 6. Beyond being reassuringly close the corresponding approximations of about 12 - 14 percentage points, given by compounding our main estimates over the election term, these effects might very well translate to forest degrowth and thus a legacy of deforestation at the end of the criminal politician's incumbency.[57] However, the arguments here rest not only on the imperfect ability of the VCF measure to distinguish between natural forests and plantations (see Section 4.3), but crucially also on extrapolations from a LATE. This is thus ultimately a question of measurement accuracy and, above all, external validity. In spite of this, our results may still be interpreted as moderate indications of a deforestation effect. After all, natural forests grow slowly, but can deteriorate much faster due to human activity.

[57] According to most long-term accounts, forest cover has marginally increased in India. For instance, the Forest Survey of India (1997; 2019) reports a total forest cover growth rate of 2.4 % over the last two decades.

7.1 Potential Mechanisms

Building on the results presented in Section 6.3, we proceed with discussing potential mechanisms behind our primary findings in Table 2. That is, why and how does the election of a criminal politician cause annual forest cover growth rates to decline by an estimated average of about 3.1 – 3.8 percentage points?

Differentiating the estimates by state characteristics and year of term provides some clues for attempting to answer this question. There is a sharp contrast in both magnitude and statistical significance between point estimates for cumulative forest cover growth in Table 6. Seemingly, the effect of electing a criminal politician accumulates over time and is by far the largest by the end of the incumbency. This is in line with the hypothesis of a latent effect, mediated through institutional channels. In other words, institutional change is slow, and it takes time for a politician to change previous standards and norms – which would imply that the actual legacy of a criminal politician in terms of forest outcomes is not evident until, at the earliest, the last year of the election term. While it is tempting to further examine this hypothesis by looking at forest cover changes even later in time, it would naturally be difficult to distinguish these from the impact of subsequent incumbents. What we can do, however, is to consider this notion in light of the results by state characteristics. With the low number of effective observations in mind, Panel C of Table 7 suggests that states with weaker institutions and higher levels of corruption (as proxied by the TII index) face disproportionately more forest cover loss following the election of a criminal candidate [58] A potential explanation for this is that law enforcement and environmental regulations in these states may be weaker in the first place, thus enabling more extractive and even illicit forestry practices. Not to say that the complementary subset of states considered in Table A13 (that is, those *not* classified as highly corrupt or institutionally malfunctioning) have strong institutions, the fact that estimates for these are both smaller and inconclusive strengthens this argument. Moreover, the notion that our estimated effect operates through institutional channels is consistent with previous research on how illicit or corrupt behaviour affects deforestation. This includes, for instance, rent-seeking behaviour undermining land use allocation processes and enforcement of land use plans (Damania et al. 2003; Asher and Novosad 2021) or bribes to local forestry officials (Barnett 1990; Southgate et al. 2000; Smith et al. 2003b). In sum, Tables 6, 7 and A13 together suggest that institutional quality matters for either accentuating or inhibiting the effect of electing a criminal politician on forest growth.

To shed further light on the potential mechanisms behind our results, we relate these and previous findings to the framework of Welsch (2004) mentioned in Section 3.1. That is, electing a criminal (or corrupt) politician may either directly affect forest outcomes, or the observed effect could operate through its impact on income. In the words of Welsch (2004), the latter can be interpreted as an indirect effect, which may or may not be mediated through the EKC.

Although we cannot establish whether our effect is direct or indirect in nature based on our data, several studies investigate the causal impact of electing a criminal politician on constituency-level economic outcomes in India using various proxies. Prakash et al. (2019) find a negative effect on

[58] Even if we cannot draw any definite conclusions based on the results for the samples of BIMAROU, least developed and highly corrupt states, respectively, in Table 7, our estimates generally align with previous findings in the literature (cf. Asher and Novosad 2017; Prakash et al. 2019).

GDP growth, Cheng and Urpelainen (2019) present reductions in household welfare and Nanda and Pareek (2016) show that this decreases total firm investment. These findings, combined with the tendency of larger estimates for least developed states in Panels A and B of Table 7, suggest that our treatment may affect deforestation indirectly through its impact on income. One explanation for this could be increased overexploitation of natural resources by impoverished households, depending on forests to earn their livelihoods. There is also some anecdotal evidence pointing towards a direct effect of electing a criminal or corrupt politician on deforestation. Although India has strict laws against illicit deforestation, there are still reports of corruption scandals and illegal activities in the forestry sector (as described in Section 2.3). However the secretive nature of illegal deforestation makes it difficult to determine to what extent our results are driven by such activities. On the other hand, intentional deforestation (e.g. for land clearing purposes) is easier to track. Previous studies find a link between rent-seeking behaviour among Indian MLAs and, for instance, environmental clearances (Kopas et al. 2021), mining clearances (Asher and Novosad 2021) and local construction projects (Lehne et al. 2018). Moreover, Cheng and Urpelainen (2019) argue that criminal incumbents can undermine environmental safeguards in the context of local infrastructure programmes, prioritizing quantity over quality. Even though all these factors may potentially affect intentional deforestation, it is hard to pin down any direct effects in the context of our analysis.

Another question that arises here is whether our results are driven by the criminality of politicians (or their inclination to engage in corrupt activities) *per se*, or if these characteristics mask other incentives or traits with bearing on our outcome. This could be the case if criminal incumbents cause deforestation not because they engage in illegal, extractive activities but rather because they are lower quality candidates (Prakash et al. 2019). That is, the observed effect may stem from poor forest management instead of intentional deterioration of natural resources. However, given that we cannot reject the null hypothesis of no difference between criminal and non-criminal candidates in observable indicators of candidate quality such as education and wealth (see Table 3), this seems unlikely for our sample. A different and potentially more worrying situation would be if either the personal political priorities, or the agenda of the party with which a criminal politician is aligned, neglect environmental issues. Arguably, this is less likely to be the case in light of the pattern suggested by Tables 4 and 5. That is, the magnitude of the effect seems to increase with the severity and number of crimes – which lends some credence to the notion that incentives associated with the criminality of politicians, and not their political agenda, drive the results. We can, however, not exclude this possibility, and attempting to do so would be beyond the scope of this paper.[59] This is but one among many other mechanisms potentially explaining our results – the signalling value of electing a criminal candidate and differential effects depending on whether the crime is financial or environmental in nature being yet another two – which we leave for future research to investigate.

7.2 External validity

While we present strong evidence for the internal validity of our estimates in Section 6.2, the external validity of the results in this study is limited. As previously mentioned, this is due to

[59] Asher and Novosad (2017) and Kopas et al. (2021) investigate the effects of partisan alignment and political affiliations in the context of Indian elections, albeit not for criminal politicians.

the fundamentally local nature of any estimates retrieved from an RD design. The causal effect observed for constituencies barely electing a criminal politician on annual forest cover growth, compared to constituencies where a non-criminal candidate won by a small margin, is limited to those in the neighbourhood of the threshold value of our running variable. More concisely, our analysis only applies to the subset of Indian constituencies where elections were close.

The coefficients presented in our main results thus capture a LATE, which is not necessarily relevant for constituencies where elections were less competitive. This point is crucial in any interpretation which extrapolates our local findings to the entirety of the sample, not least for the reasoning in the beginning of this section regarding whether our estimated effect captures deforestation or not. Although this suggests that such is the case, we are not inclined to accept this extrapolation as a casual finding. A major argument for this is that the political equilibrium arising from close elections is likely to differ from other constituencies, especially when it comes to how increased competition affects the incentives and behaviour of the winning candidate (Asher and Novosad 2017; Lehne et al. 2018). For instance, criminal politicians that won in competitive elections may be more prone to favour certain interest groups, such as the formal or informal forestry industry, or engage in rent-seeking behaviour in order to retain incumbency. On the other hand, the robustness of our results to the artificial thresholds examined in Section 6.2.4 speaks for some validity beyond the immediate proximity of the cut-off.

Moreover, our baseline sample consists only of elections where a criminal politician competed against a non-criminal, by construction of the RD design. These constituencies are likely to differ in various ways from constituencies where no criminal candidates contested in the election, or where two criminal candidates competed against each other. Naturally, this may also limit the relevance of this analysis in a broader perspective.

7.3 Limitations

Although the generalisability of our results to other constituencies and regions depends on the extent to which they differ from those within our optimal bandwidths, our causal estimates contribute to the literature on deforestation and the broader corruption-environment nexus. To the best of our ability, we have verified the validity and robustness of our findings based on available data. Nevertheless, we conclude this section by summarizing what we identify as the most immediate shortcomings of this study.

The first relates to the fact that we are unable to control for candidate-, election- and certain constituency-level characteristics associated with elections held before those of the sample. Since the affidavit data is limited by the Supreme Court order of 2003 and thus begins in 2004, while constituency boundaries changed in 2008 (see Section 2.2.1), this constrains our analysis. This is only a minor issue in terms of the continuity tests, since our data contains many predetermined characteristics already. Considering the plausibly latent nature of the effect of electing a criminal politician, however, it is a cause of concern. As discussed in Section 7.1, it may take longer than an election term for the true effect to materialize. Without data on previous elections, we cannot exclude the possibility that our results are biased by previous incumbents.

A second concern is the variation in the VCF data. We have taken several steps to control for this potential noise induced by the satellite-based measurement, including the exclusion of outliers, removing constituencies with less than 10 % average forest cover in the election year and controlling for state-year fixed effects (as described in Sections 4.3 and 5.1). Despite this, there is still a possibility that remaining unrepresentative variation affects our results.

Although the sample could have been extended significantly by including constituencies with low forest cover or by appending the panel with additional incomplete election terms, thus granting more statistical power to our tests, we deem the cost of doing so too high in terms of the bias it may introduce to the analysis. On a more general note, these shortcomings boil down to data availability. With that said, we hope that future studies on this topic will complement our results based on more extensive and even more accurate data.

8 Conclusion

This paper set out to investigate the impact of electing a criminal politician on a constituency-level measure of forest growth, with an overarching ambition to shed light on the complicated relationship between corruption and environmental outcomes in developing countries. We combine satellite data on forest cover with election outcomes and criminal records of candidates based on sworn affidavits into a balanced panel of Indian constituencies spanning from 2004 to 2014. By implementing a state-of-the-art regression discontinuity design on constituencies where elections were close, we estimate a causal and statistically significant negative effect of a criminal politician coming to power. This is our most important finding – barely electing a criminal politician results in about 3 – 4 percentage points lower annual forest cover growth over the course of the election term, compared to constituencies where a criminal candidate lost by a small margin. Albeit limited to competitive elections, our analysis suggests that this effect translates to deforestation over time. After all, natural forests grow slowly but may deteriorate much faster due to human activities.

In a series of extensive validity tests, we justify the causal interpretation of our estimates. In general, we also find that results are robust to various alternatives to choices made throughout the analysis. By considering more stringent definitions of criminal politicians, we show that the effect on forest growth is robust to and seems to increase with the severity and number of crimes. In an attempt to pin down potential mechanisms behind our findings, we further examine how estimates vary over the election term, by approximated institutional quality and by level of development. We find a distinct latent effect only in the final year of term and that the impact is accentuated in states considered to have weaker institutions and lower levels of development. As such, we interpret this as indications of the effect on forest growth operating through institutional channels.

Regardless of whether our estimates may be attributed to differences in quality, partisan alignment or inclinations to engage in corrupt activities, the analysis of this paper demonstrates that priorities between criminal and non-criminal politicians diverge. This broad tendency aligns with previous findings, such as Chemin (2012), Nanda and Pareek (2016), Asher and Novosad (2017), Cheng and Urpelainen (2019) and Prakash et al. (2019). When it comes to forest protection, this study clearly shows that electing a criminal politician is costly for the environment. Even though we cannot draw any definite parallels between our estimated causal effects and corrupt behaviour in India and elsewhere, we believe that our results may provide important insights for other settings and constitute an important contribution to the emerging corruption-environment nexus.

9 References

Abman, R. 2018. Rule of Law and Avoided Deforestation from Protected Areas. *Ecological Economics*, 146:282-289.

Association for Democratic Reforms. 2018. MPs/MLAs with Declared Criminal Cases. *Association for Democratic Reforms & National Election Watch.* Accessed 2021-11-01.

Aggarwal, A., Das, S. & Paul, V. 2009. *Is India ready to implement REDD Plus - A preliminary assessment.* The Energy Research Institute, New Delhi for COP 15.

Aidt, T., Asatryan, Z., Badalyan, L. & Heinemann, F. 2020. Vote buying or (political) business (cycles) as usual?. *Review of Economics and Statistics*, 102(3):409-425.

Amacher, G. S., Ollikainen, M. & Koskela, E. 2012. Corruption and forest concessions. *Journal of Environmental Economics and Management*, 63(1):92-104.

Angrist, J. D. & Lavy, V. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2):533-575.

Asher, S. & Novosad P. 2017. Politics and local economic growth: evidence from India. *American Economic Journal: Applied Economics*, 9(1):229–73.

Asher, S., Lunt, T., Matsuura, R. & Novosad, P. 2020a. *The Socioeconomic High-resolution Rural-Urban Geographic Dataset on India (SHRUG) Data Description and Codebook. v1.5 (samosa), December 2020.* Development Data Lab.

Asher, S., Garg, T. & Novosad, P. 2020b. The ecological impact of transportation infrastructure. *The Economic Journal*, 130(629):1173-1199.

Asher, S. & Novosad, P. 2021. Rent-Seeking and Criminal Politicians: Evidence from Mining Booms. *The Review of Economics and Statistics*, 1–44.

Asher, S., Lunt, T., Matsuura, R. & Novosad, P. 2021. Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform. *World Bank, Policy Research Working Paper*, 9540.

Ashutosh, S. & Roy, P.S. 2021. Three Decades of Nationwide Forest Cover Mapping Using Indian Remote Sensing Satellite Data: A Success Story of Monitoring Forests for Conservation in India. *Journal of the Indian Society of Remote Sensing*, 49:61–70.

Banerjee A., Green DP., McManus J. & Pande R. 2014. Are poor voters indifferent to whether elected leaders are criminal or corrupt? A vignette experiment in rural India. *Political Communication*, 31(3):391–407.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A. & Sumarto, S. 2018. Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia. *Journal of Political Economy*, 126(2):451-491.

Barbier, E.B. 2004. Explaining Agricultural Land Expansion and Deforestation in Developing Countries. *American Journal of Agricultural Economics*, 86(5):1347–53.

Barbier, E.B., Damania, R. & Léonard, D. 2005. Corruption, trade and resource conversion, *Journal of Environmental Economics and Management.* 50(2):276-299.

Bardhan, P. & Mookherjee, D. 2012. Political clientelism and capture: Theory and evidence from West Bengal, India. *WIDER Working paper*, 2012/97.

Barnett, T. E. 1990. *The Barnett report: a summary of the report of the commission of inquiry into aspects of the timber industry in Papua New Guinea.* Asia-Pacific Action Group.

Barreca, A. I., Lindo, J. M. & Waddell, G. R. 2016. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1):268-293.

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.

Bertrand, M., Djankov, S., Hanna, R. & Mullainathan, S. 2007. Obtaining a driver's license in India: an experimental approach to studying corruption. *The Quarterly Journal of Economics*, 122(4):1639-1676.

Besley, T. 2005. Political selection. *Journal of Economic Perspectives*, 19(3):43-60.

Biswas, A. K., Farzanegan, M. R. & Thum, M. 2012. Pollution, shadow economy and corruption: Theory and evidence. *Ecological Economics*, 75:114-125.

Borah, B., Bhattacharjee, A. & Ishwar, N. M. 2018. *Bonn challenge and India: Progress on restoration efforts across states and landscapes.* New Delhi, India: IUCN and MoEFCC, Government of India.

Brollo, F., Nannicini, T., Perotti, R. & Tabellini, G. 2013. The political resource curse. *American Economic Review*, 103(5):1759-96.

Brown, J. 2021. Data Smoothing - Reducing the "Noise" in NDVI. *United States Geological Survey.* Accessed 2021-11-14.

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, 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. & Titiunik, R. 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295-2326.

Calonico, S., Cattaneo, M. D. & Titiunik, R. 2015. Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753-1769.

Calonico, S., Cattaneo, M. D., Farrell, M. H. & Titiunik, R. 2017. rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372-404.

Calonico, S., Cattaneo, M. D. & Farrell, M. H. 2018. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, 113(522):767-779.

Calonico, S., Cattaneo, M. D. & Farrell, M. H. 2020. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192-210.

Card, D., Lee, D. S., Pei, Z. & Weber, A. 2014. Local polynomial order in regression discontinuity designs. *Brandeis University, Department of Economics, Working Paper.*

Cattaneo, M. D., Frandsen, B. R. & Titiunik, R. 2015. Randomization inference in the regression discontinuity design: An application to party advantages in the US Senate. *Journal of Causal Inference*, 3(1):1-24.

Cattaneo, M. D. & Vazquez-Bare, G. 2017. The choice of neighborhood in regression discontinuity designs. *Observational Studies*, 3(2):134-146.

Cattaneo, M. D., Idrobo, N. & Titiunik, R. 2019. *A practical introduction to regression discontinuity designs: Foundations.* Cambridge University Press.

Caughey, D. & Sekhon, J. S. 2011. Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008. *Political Analysis*, 19(4):385-408.

Cheng, C. Y. & Urpelainen, J. 2019. Criminal Politicians and Socioeconomic Development: Evidence from Rural India. *Studies in Comparative International Development*, 54(4):501-527.

Chemin, M. 2008. Do criminal politicians reduce corruption? Evidence from India. *Centre Interuniversitaire sur le Risque, les Politiques Èconomiques et l'Emploi (CIRPÉE).* Working Paper 08-25.

Chemin, M. 2012. Welfare effects of criminal politicians: a discontinuity-based approach. *The Journal of Law and Economics*, 55(3):667-690.

Choumert, J., Motel, P. C. & Dakpo, H. K. 2013. Is the Environmental Kuznets Curve for deforestation a threatened theory? A meta-analysis of the literature. *Ecological Economics*, 90:19-28.

Cisneros, E., Kis-Katos, K. & Nuryartono, N. 2021. Palm oil and the politics of deforestation in Indonesia. *Journal of Environmental Economics and Management*, 108, 102453.

Cole, M. A., Rayner, A. J. & Bates, J. M. 1997. The environmental Kuznets curve: an empirical analysis. *Environment and development economics*, 2(4):401-416.

Cole, M. A. 2007. Corruption, income and the environment: an empirical analysis. *Ecological economics*, 62(3-4):637-647.

Cozma, A. C., Vaidean, V. L. & Achim, M. V. 2021. Corruption, Shadow Economy and Deforestation: Friends or Strangers?. *Risks*, 9(9):153.

Culas, R. J. 2007. Deforestation and the environmental Kuznets curve: An institutional perspective. *Ecological Economics*, 61(2-3):429-437.

Cunningham, S. 2021. *Causal Inference. The Mixtape, 1 - Chapter 6 Regression Discontinuity.* Yale University Press.

Damania, R., Fredriksson, P. G. & List, J. A. 2003. Trade liberalization, corruption, and environmental policy formation: theory and evidence. *Journal of Environmental Economics and Management*, 46(3):490-512.

Davidar, P., Sahoo, S., Mammen, P. C., Acharya, P., Puyravaud, J. P., Arjunan, M., ... & Roessingh, K. 2010. Assessing the extent and causes of forest degradation in India: where do we stand?. *Biological Conservation*, 143(12):2937-2944.

de la Cuesta, B. & Imai, K. 2016. Misunderstandings about the regression discontinuity design in the study of close elections. *Annual Review of Political Science*, 19:375-396.

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D. H., Kelly, M. & Townshend, J. R. G. 2015. MOD44B MODIS/terra vegetation continuous fields yearly L3 global 250m SIN grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, 10.

Dinda, S. 2004. Environmental Kuznets curve hypothesis: a survey. *Ecological Economics*, 49(4): 431-455

Duflo, E., Greenstone, M., Pande, R. & Ryan, N. 2013. Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *The Quarterly Journal of Economics*, 128(4): 1499-1545.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B. & Snyder Jr, J. M. 2015. On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science*, 59(1):259-274.

Ehrhardt-Martinez, K., Crenshaw, E. M. & Jenkins, J. C. 2002. Deforestation and the environmental Kuznets curve: A cross-national investigation of intervening mechanisms. S*ocial Science Quarterly*, 83(1):226-243.

Eldeeb, O., Prochazka, P. & Maitah, M. 2015. Causes for deforestation in Indonesia: Corruption and palm tree plantation. *Asian Social Science*, 11(27):120.

Election Commission of India. 2018. The Functions (Electoral System of India). Accessed 2021-11-22.

Election Commission of India. 2020. Delimitation. Accessed 2021-11-22.

FAO & UNEP. 2020. *The state of the world's forests 2020: forests, biodiversity and people.* FAO & UNEP, Rome.

Ferraz, C. & Finan, F. 2011. Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274-1311.

Fisman, R., Schulz, F. & Vig, V. 2014. The private returns to public office. *Journal of Political Economy*, 122(4):806-862.

Fredriksson, P. G., Vollebergh, H. R. & Dijkgraaf, E. 2004. Corruption and energy efficiency in OECD countries: theory and evidence. *Journal of Environmental Economics and Management*, 47(2):207-231.

Forest (Conservation) Act. 1980. Parliament of India.

Forest Rights Act (FRA). 2006. Ministry of Tribal Affairs, Government of India and United Nations Development Programme.

Forest Survey of India. 1997. *State of Forest Report 1997*. Ministry of Environment Forest and Climate Change, Government of India.

Forest Survey of India. 2019. *State of Forest Report 2019*. Ministry of Environment Forest and Climate Change, Government of India.

Galinato, G. I. & Galinato, S. P. 2013. The short-run and long-run effects of corruption control and political stability on forest cover. *Ecological Economics*, 89:153-161.

Global Forest Watch. 2021. Summary of Key Statistics About Forests in India. Accessed 2021-11-22.

Greenstone, M. & Jack, B. K. 2015. Envirodevonomics: A research agenda for an emerging field. *Journal of Economic Literature*, 53(1):5-42.

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.

Halkos, G. E. & Tzeremes, N. G. 2013. Carbon dioxide emissions and governance: a nonparametric analysis for the G-20. *Energy Economics*, 40:110-118.

Hansen, M. C., Stehman, S. V. & Potapov, P. V. 2010. Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences*, 107(19):8650-8655.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. 2013. High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850-853.

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*, 54(4):471-494.

Haseeb, M. & Azam, M. 2021. Dynamic nexus among tourism, corruption, democracy and environmental degradation: a panel data investigation. *Environment, Development and Sustainability*, 23(4):5557-5575.

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O. & Tukiainen, J. 2018. When does regression discontinuity design work? Evidence from random election outcomes. *Quantitative Economics*, 9(2):1019-1051.

Imbens, G. W., and Angrist, J. D. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–75.

Imbens, G. W. & Lemieux, T. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615-635.

Jaffrelot, C. & Verniers, G. 2014. Indian Elections: Reaching the End of the Democratic Tether. Esprit, (7):75-87.

Jensenius, F. R. 2015. Development from representation? A study of quotas for the scheduled castes in India. *American Economic Journal: Applied Economics*, 7(3):196-220.

Jensenius, F. R. & Verniers, G. 2017. Studying Indian politics with large-scale data: Indian election data 1961–today. *Studies in Indian Politics*, 5(2):269-275.

Khemani, S. 2004. Political cycles in a developing economy: effect of elections in the Indian states. *Journal of Development Economics*, 73(1):125-154.

Kopas, J., Urpelainen, J. & York, E. A. 2021. Greasing the Wheels: the Politics of Environmental Clearances in India. *Studies in Comparative International Development*, 1-32.

Koyuncu, C. & Yilmaz, R. 2009. The impact of corruption on deforestation: a cross-country evidence. *The Journal of Developing Areas*, 213-222.

Koyuncu, C. & Yilmaz, R. 2013. Deforestation, corruption, and private ownership in the forest sector. *Quality & Quantity*, 47(1):227-236.

Kumari, R., Banerjee, A., Kumar, R., Kumar, A., Saikia, P. & Khan, M. L. 2019. *Forest degradation around the world - Chapter 4, Deforestation in India: consequences and sustainable solutions.* IntechOpen.

Laurance, W. F., Kakul, T., Keenan, R. J., Sayer, J., Passingan, S., Clements, G. R. & Sodhi, N. S. 2011. Predatory corporations, failing governance, and the fate of forests in Papua New Guinea. *Conservation Letters*, 4(2):95-100.

Lee, D. S. 2008. Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675-697.

Lee, D. S. & Lemieux, T. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281-355.

Lehne, J., Shapiro, J. N. & Eynde, O. V. 2018. Building connections: Political corruption and road construction in India. *Journal of Development Economics*, 131:62-78.

Leitão, A. 2010. Corruption and the environmental Kuznets curve: empirical evidence for sulfur. *Ecological Economics*, 69(11):2191-2201.

Lippe, M. 1999. Corruption and environment at the local level. *Transparency International*. Working Paper: Corruption and Environment.

Lopez, R. & Mitra, S. 2000. Corruption, pollution, and the Kuznets environment curve. *Journal of Environmental Economics and Management*, 40(2):137-150.

Mahadevan, M. 2019. The price of power: Costs of political corruption in Indian electricity. *Working Paper of the Department of Economics, University of Michigan.*

Martin, N. & Picherit, D. 2020. Special issue: electoral fraud and manipulation in India and Pakistan. *Commonwealth & Comparative Politics*, 58(1):1-20.

McCarthy, J. F. 2002. Turning in circles: district governance, illegal logging, and environmental decline in Sumatra, Indonesia. *Society & Natural Resources*, 15(10):867-886.

McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698-714.

Meehan, F. & Tacconi, L. 2017. A framework to assess the impacts of corruption on forests and prioritize responses. *Land Use Policy*, 60:113-122.

Milledge, S., Ahrends, A. & Gelvas, I. 2007. *Forestry, Governance and National Development: Lessons Learned from a Logging Boom in Southern Tanzania.* Traffic East/Southern Africa.

MoEFCC. 2018. *National REDD+ Strategy India.* Ministry of Environment, Forest and Climate Change, Government of India.

Nanda, V. K. & Pareek, A. 2016. Do Criminal Politicians Affect Firm Investment and Value? Evidence from a Regression Discontinuity Approach. *SSRN Working Paper.*

Pailler, S. 2018. Re-election incentives and deforestation cycles in the Brazilian Amazon. *Journal of Environmental Economics and Management*, 88:345-365.

Palmer, C. 2001. The extent and causes of illegal logging: an analysis of a major cause of tropical deforestation in Indonesia. *Centre for Social and Economic Research on the Global Environment, Working Paper.*

Prakash, N., Rockmore, M. & Uppal, Y. 2019. Do criminally accused politicians affect economic outcomes? Evidence from India. *Journal of Development Economics*, 141:102370.

Puyravaud, J. P., Davidar, P. & Laurance, W. F. 2010. Cryptic destruction of India's native forests. *Conservation Letters*, 3(6):390-394.

Ravishankar, N. 2009. The cost of ruling: Anti-incumbency in elections. *Economic and Political Weekly*, 92-98.

Reinikka, R. & Svensson, J. 2004. Local capture: evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*, 119(2):679-705.

Robbins, P. 2000. The rotten institution: corruption in natural resource management. *Political Geography*, 19(4):423-443.

Rubin, D. B. 2005. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322-331.

Santos, L. A., Ferreira, K. R., Camara, G., Picoli, M. C. & Simoes, R. E. 2021. Quality control and class noise reduction of satellite image time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 177:75-88.

Sarmah, D. 2020. *Agroforestry in Karnataka - A Golden Opportunity for Green Growth.* Notion Press.

Sasaki, N. & Putz, F. E. 2009. Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements. *Conservation Letters*, 2(5):226-232.

Sinha, A., Gupta, M., Shahbaz, M. & Sengupta, T. 2019. Impact of corruption in public sector on environmental quality: Implications for sustainability in BRICS and next 11 countries. *Journal of Cleaner Production*, 232:1379-1393.

Smith, R. J., Muir, R. D., Walpole, M. J., Balmford, A. & Leader-Williams, N. 2003a. Governance and the loss of biodiversity. *Nature*, 426(6962):67-70.

Smith, J., Obidzinski, K., Subarudi, S. & Suramenggala, I. 2003b. Illegal logging, collusive corruption and fragmented governments in Kalimantan, Indonesia. *International Forestry Review*, 5(3):293-302.

Smith, R. J. & Walpole, M. J. 2005. Should conservationists pay more attention to corruption?. *Oryx*, 39(3):251-256.

Smith, J., Colan, V., Sabogal, C. & Snook, L. 2006. Why policy reforms fail to improve logging practices: The role of governance and norms in Peru. *Forest Policy and Economics*, 8(4):458-469.

Southgate, D., Salazar-Canelos, P., Camacho-Saa, C. & Stewart, R. 2000. Markets, institutions, and forestry: the consequences of timber trade liberalization in Ecuador. *World Development*, 28(11):2005-2012.

Sundström, A. 2016. Understanding illegality and corruption in forest governance. *Journal of Environmental Management*, 181:779-790.

Tacconi, L., Downs, F. & Larmour, P. 2009. *Realising REDD+: National strategy and policy options - Anti-corruption policies in the forest sector and REDD+.* Center for International Forestry Research, 163-174.

Teye, J. K. 2013. Corruption and illegal logging in Ghana. *International Development Planning Review*, 35(1).

Thistlethwaite, D. L. & Campbell, D. T. 1960. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309.

Times of India. 2020a. Share of MPs with criminal cases has nearly doubled. Takur, A., Times of India, 2020-02-18. Accessed 2020-11-04.

Times of India. 2020b. 320 corrupt government officials given premature retirement. Times of India, 2020-03-04. Accessed 2021-11-04.

Times of India. 2020c. Gujarat ACB nabs forest official in disproportionate assets case. Times of India, 2020-11-06. Accessed 2021-11-04.

Times of India. 2021a. Andhra Pradesh: Forest officials trapped by ACB. Times of India, 2021-05-09. Accessed 2021-11-04.

Times of India. 2021b. Forest guard, relatives charge-sheeted in disproportionate assets case. Times of India, 2021-07-21. Accessed 2021-11-04.

Times of India. 2021c. Tree felling posing threat to one of Goa's highest peaks. Times of India, Kerkar, R. P., 2021-11-16. Accessed 2021-11-04.

Townshend, J., Hansen, M. C., Carroll, M., DiMiceli, C., Sohlberg, R. & Huang, C. 2011. *User Guide for the MODIS Vegetation Continuous Fields product Collection 5 Version 1.* University of Maryland.

Transparency International. 2020. *Global Corruption Barometer - Asia 2020: Citizens' Views and Experiences of Corruption.* Transparency International.

Transparency International. 2021. What is corruption?. Accessed 2021-11-05.

Urrunaga, J. M., Johnson, A., Orbegozo, I. D. & Mulligan, F. 2012. *The laundering machine: How fraud and corruption in Peru's concession system are destroying the future of its forests.* Washington, DC: Environmental Investigation Agency.

Vaishnav, M. 2017. *When crime pays.* Yale University Press.

Welsch, H. 2004. Corruption, growth, and the environment: a cross-country analysis. *Environment and Development Economics*, 9(5):663-693.

World Bank. 2006. *Strengthening Forest Law Enforcement and Governance: Addressing a Systemic Constraint to Sustainable Development.* The International Bank for Reconstruction and Development, The World Bank.

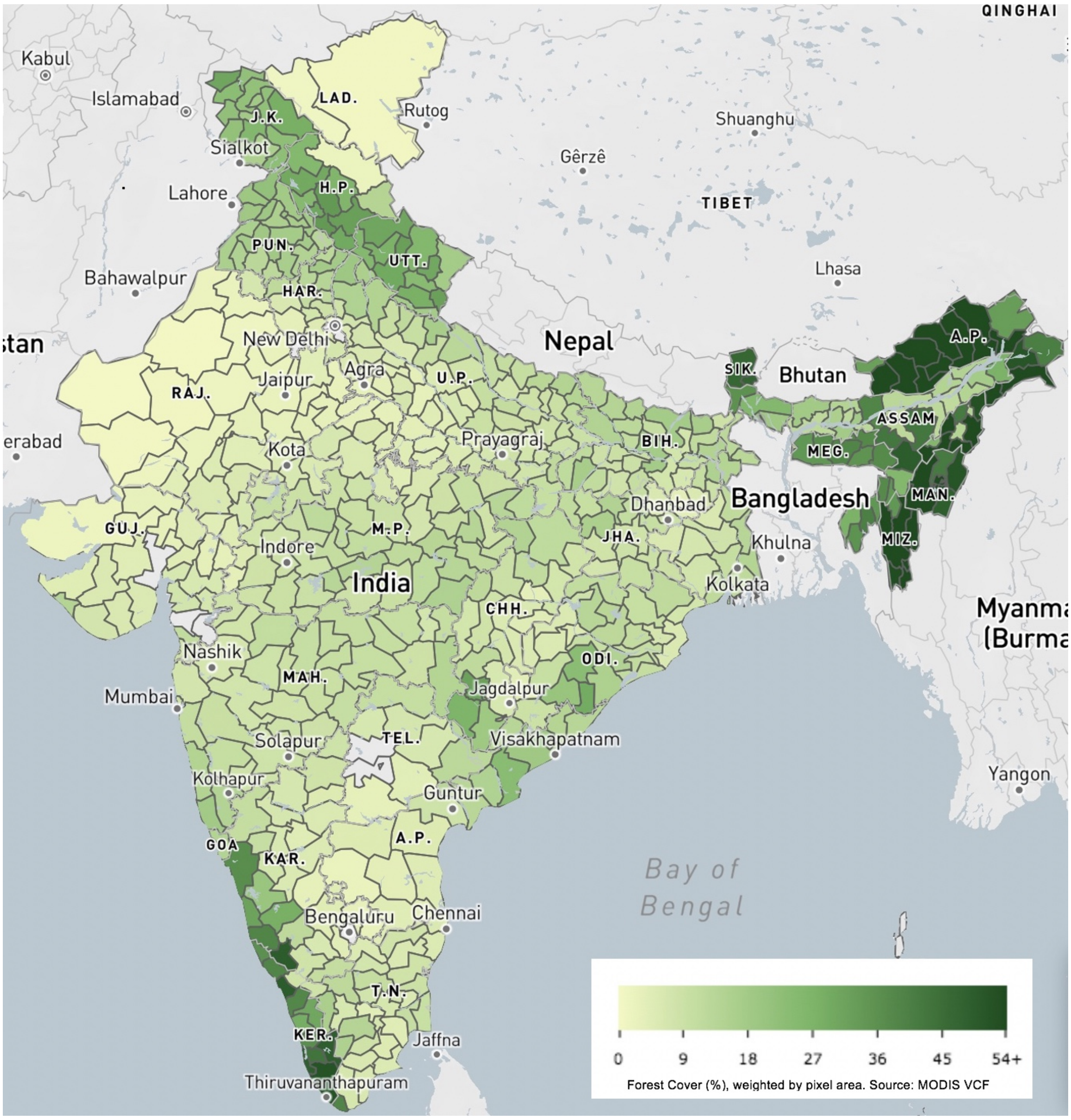
WRI. 2021. Forest Pulse: The Latest on the World's Forests. Accessed 2021-11-22.

Zhou, M., Wang, B. & Chen, Z. 2020. Has the anti-corruption campaign decreased air pollution in China?. *Energy Economics*, 91:104878.

10 Appendix

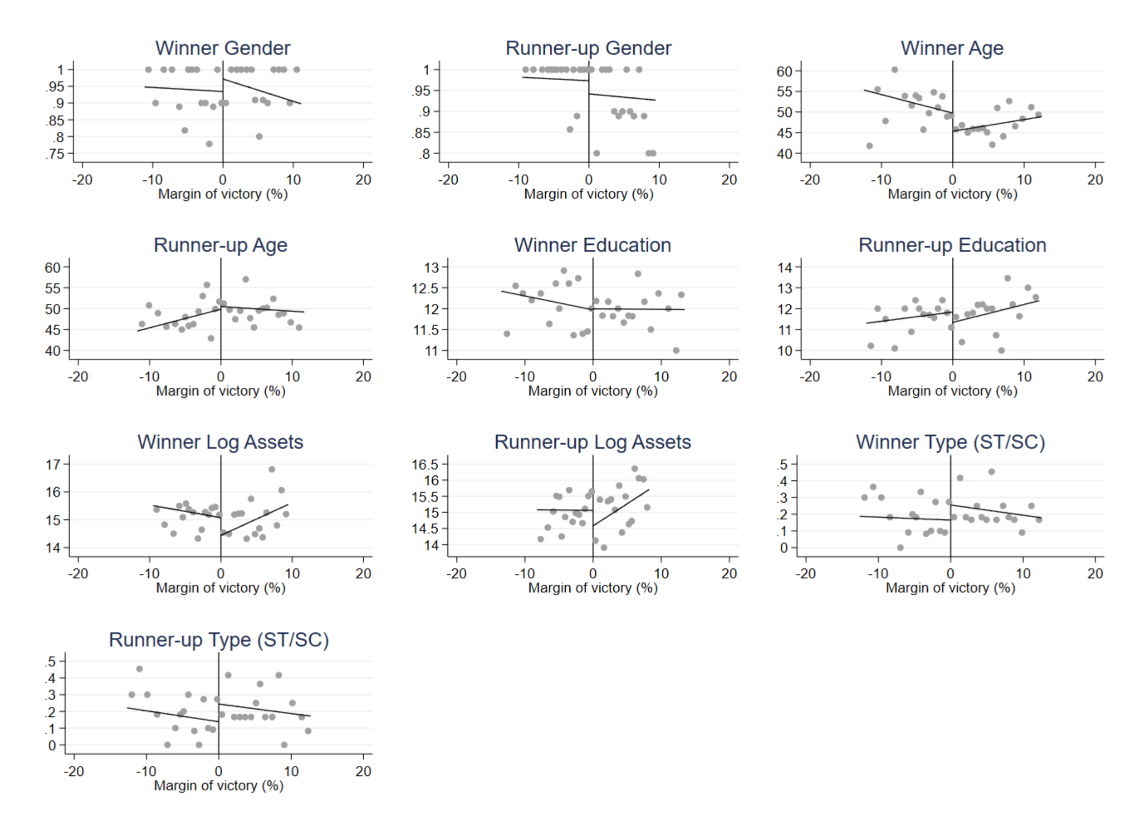
10.1 Figures

Figure A1: Forest Cover Map of India (2014)



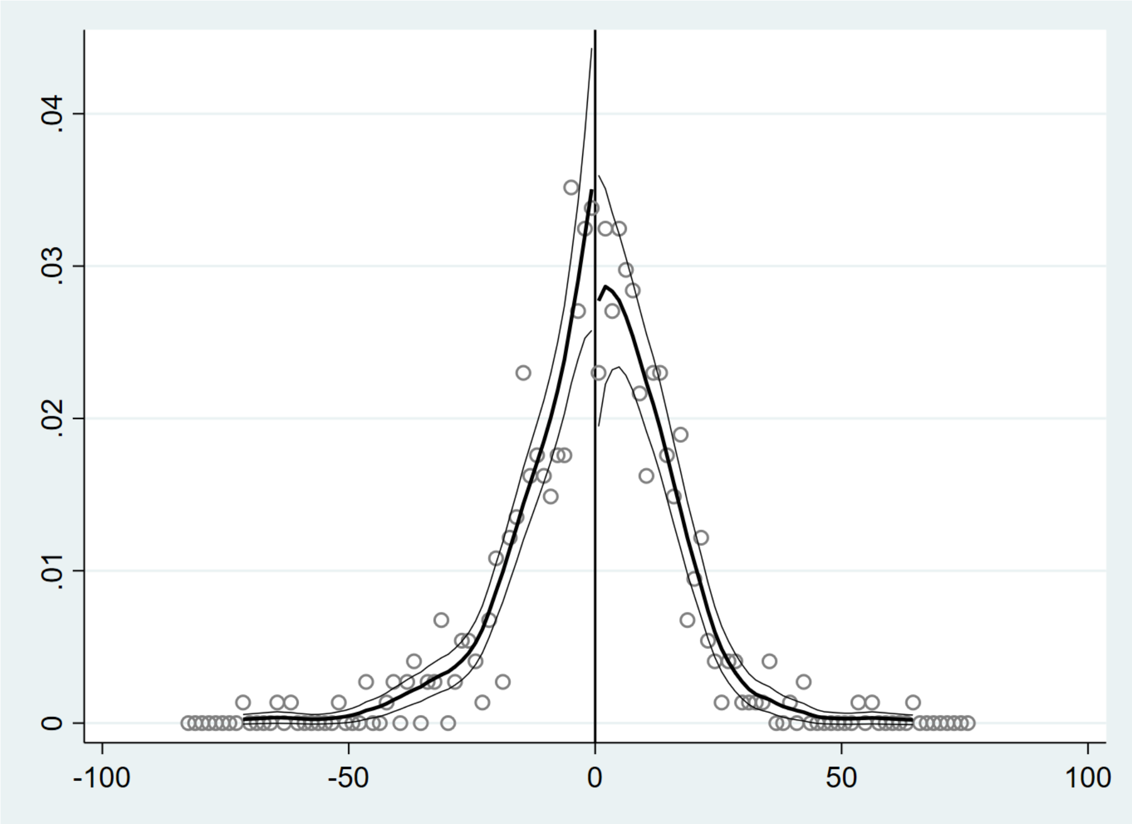
Notes: Visualization of average forest cover on the district level in India in 2014. The figure comes from the DDL (2021) SHRUG Atlas and is based on VCF MODIS data from DiMiceli et al. (2015).

Figure A2: Continuity tests - Candidate-level Characteristics



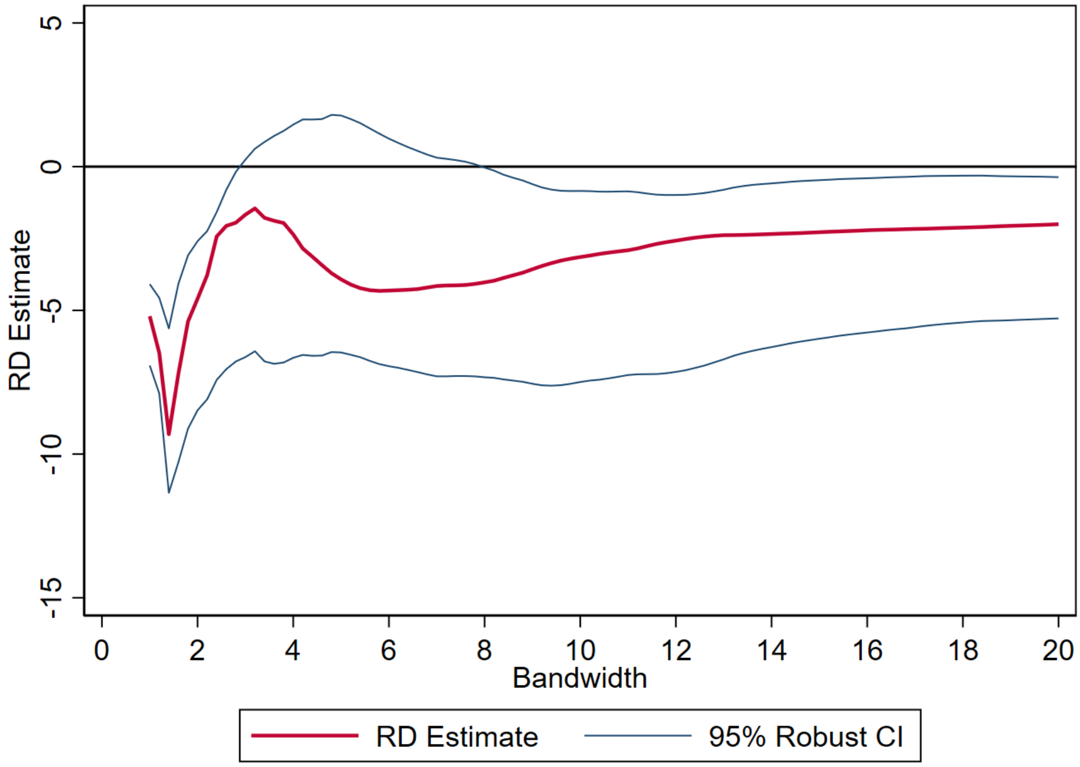
Notes: The figures show a local linear approximation of each observed candidate-level characteristic conditional on the margin of victory of a criminal candidate, fitted separately on each side of the threshold within the corresponding MSE-optimal bandwidth for each variable. The vertical difference in intercepts at the threshold is identical to the point estimates in the formal continuity test, which uses a triangular kernel function and control for election year-state fixed effects. The dots depict sample averages within 15 quantile-spaced bins on each side, assuring that each bin contains roughly the same number of unique observations.

Figure A3: McCrary Density Test



Notes: Graphical illustration of the density of the margin of victory for a criminal candidate, estimated separately on each side of the threshold using the procedure of McCrary (2008) within the MSE-optimal bandwidth of 10.491.

Figure A4: Sensitivity to Bandwidth - Including controls



Notes: The figure plots estimated RD effects against different bandwidths, retrieved using triangular kernel-weighted local linear estimations within each bandwidth in two decile increments over a range of 1-20 %. Each estimation controls for year-state fixed effects and additional covariates. Results for bandwidths between 0 and 1 are omitted due to insufficient observations to construct robust bias-corrected confidence intervals.

10.2 Tables

Table A1: Main Results - Excluding Thiessen Polygons

	(1)	(2)	(3)	(4)
Criminal Winner	-4.771***	-3.949**	-2.948**	-3.881***
	(1.677)	(1.975)	(1.462)	(1.438)
Robust P-value	0.002	0.026	0.027	0.003
Robust CI	[-8.595, -2.020]	[-8.252, -0.511]	[-6.100, -0.370]	[-7.052, -1.417]
Eff. no of obs.	1144	1260	1116	972
Opt. Bandwidth	10.618	12.046	10.195	8.385
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Estimates are retrieved from the baseline sample excluding states in which more than 50 % of the locations were Thiessen polygons (Arunachal Pradesh, Manipur, Meghalaya, Mizoram and Nagarand). Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A2: Main Results - Different Restrictions of Forest Cover

Panel A	Above 5 % Forest Cover			
	(1)	(2)	(3)	(4)
Criminal Winner	-0.983	-1.973	-1.985*	-1.907*
	(1.445)	(1.449)	(1.203)	(1.186)
Robust P-value	0.560	0.153	0.070	0.074
Robust CI	[-3.675, 1.990]	[-4.912, 0.770]	[-4.536, 0.179]	[-4.444, 0.207]
Eff. no of obs.	2968	2968	2472	2408
Opt. Bandwidth	12.131	12.150	9.424	9.287
Panel B	Above 0 % Forest Cover			
	(1)	(2)	(3)	(4)
Criminal Winner	0.065	-1.073	-1.196	-0.921
	(1.384)	(1.301)	(1.045)	(0.970)
Robust P-value	0.816	0.422	0.222	0.334
Robust CI	[-2.390, 3.036]	[-3.594, 1.506]	[-3.324, 0.771]	[-2.839, 0.964]
Eff. no of obs.	3664	3860	3500	3664
Opt. Bandwidth	10.151	11.081	9.568	10.415
Panel C	Below 10% Forest Cover			
	(1)	(2)	(3)	(4)
Criminal Winner	1.515	-0.352	-0.034	0.195
	(1.890)	(1.763)	(1.397)	(1.367)
Robust P-value	0.332	0.820	0.942	0.945
Robust CI	[-1.871, 5.538]	[-3.856, 3.054]	[-2.840, 2.635]	[-2.584, 2.773]
Eff. no of obs.	2332	2332	2152	2088
Opt. Bandwidth	9.057	9.109	8.179	8.036
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Panel A: Baseline sample excluding constituencies with below 5% average forest cover in the year of election. Panel B: Full baseline sample without any sample limitations based on average forest. Panel C: Baseline sample excluding constituencies with above 10% average forest cover in the year of election. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Main Results - Including Outliers

	(1)	(2)	(3)	(4)
Criminal Winner	-4.512***	-3.506**	-2.986**	-3.956***
	(1.628)	(1.942)	(1.391)	(1.386)
Robust P-value	0.002	0.043	0.019	0.002
Robust CI	[-8.226, -1.845]	[-7.728, -0.115]	[-6.000, -0.545]	[-7.020, -1.587]
Eff. no of obs.	1172	1336	1144	992
Opt. Bandwidth	10.616	12.535	10.104	8.335
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Estimates are retrieved from the baseline sample including the 6 constituencies defined as outliers. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Main Results - Unbalanced Panel

	(1)	(2)	(3)	(4)
Criminal Winner	-2.264	-2.410	-2.078*	-3.031**
	(1.919)	(1.858)	(1.334)	(1.348)
Robust P-value	0.141	0.132	0.056	0.009
Robust CI	[-6.587, 0.934]	[-6.441, 0.843]	[-5.162, 0.066]	[-6.179, -0.897]
Eff. no of obs.	1561	1613	1502	1248
Opt. Bandwidth	10.018	10.636	10.636	7.804
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function. Estimates are retrieved from an unbalanced panel, excluding outliers and constituencies with less than 10 % average forest cover in the election year. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Main Results - Parametric Estimations with Higher Order Polynomials

Polynomial of	Degree 2		Degree 3		Degree 4	
	(1)	(2)	(3)	(4)	(5)	(6)
Criminal winner	-1.864**	-1.708**	-2.091**	-2.227**	-2.692**	-2.841**
	(0.863)	(0.833)	(1.060)	(1.021)	(1.286)	(1.233)
Observations	2132	2116	2132	2116	2132	2116
Conventional P-value	0.031	0.040	0.049	0.029	0.036	0.021
Conventional CI	[-3.555, -0.173]	[-3.342, -0.074]	[-4.168, -0.013]	[-4.229, -0.226]	[-5.213, -0.171]	[-5.257, -0.425]
Year FE	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES

Notes: Results are from parametric estimations on the full baseline sample using a triangular kernel function and second (columns 1-2), third (columns 3-4) and fourth (columns 5-6) degree polynomials. Each column presents point estimates and conventional standard errors, confidence intervals and p-values, as well as the number of observations used in the estimation. Odd columns include year-state fixed effects and even columns introduce additional controls. Conventional errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Main Results - CER-Optimal Bandwidth (Optimal Inference)

	Linear Fit				Quadratic Fit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Criminal Winner	-4.863***	-3.934*	-3.746***	-4.297***	-5.234**	-3.121	-4.283**	-3.992**
	[1.646]	[2.175]	[1.503]	[1.471]	[2.069]	[2.810]	[2.159]	[1.864]
Robust P-value	0.002	0.058	0.009	0.002	0.010	0.287	0.040	0.028
Robust CI	[-8.369, -1.917]	[-8.384, 0.141]	[-6.864, -0.973]	[-7.366, -1.599]	[-9.355, -1.243]	[-8.500, 2.515]	[-8.659, -0.194]	[-7.759, -0.452]
Eff. no of obs.	1036	1024	920	808	1256	1160	1056	1008
Opt. Bandwidth	8.960	8.828	7.665	6.243	11.853	10.469	9.236	8.757
Year FE	NO	YES	YES	YES	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES	NO	NO	YES	YES
Controls	NO	NO	NO	YES	NO	NO	NO	YES

Notes: Results are from local linear estimations within CER-optimal bandwidths, using a triangular kernel function. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. Columns 1-4 are local linear while columns 5-6 are local quadratic approximations. No controls are included in columns (1,5), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2,6), (3,7) and (4,8), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A7: Main Results - Quadratic Fit

	(1)	(2)	(3)	(4)
Criminal Winner	-4.868***	-3.571	-3.425**	-4.120**
	(1.899)	(2.518)	(1.910)	(1.730)
Robust P-value	0.008	0.201	0.044	0.010
Robust CI	[-8.773, -1.329]	[-8.156, 1.715]	[-7.585, -0.097]	[-7.824, -1.042]
Eff. no of obs.	1668	1536	1388	1316
Opt. Bandwidth	16.969	14.987	13.222	12.531
Year FE	NO	YES	YES	YES
Year*State FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES

Notes: Results are from local quadratic estimations within MSE-optimal bandwidths, using a triangular kernel function. Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values, as well as the effective number of observations used in the estimation. No controls are included in column (1), whereas year fixed-effects, year-state fixed effects and additional controls are introduced in columns (2), (3) and (4), respectively. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Descriptive Statistics for Excluded Covariates

	Full Sample						Baseline Sample					
	N	Median	Mean	SD	min	max	N	Median	Mean	SD	min	max
Tot. Pop. Growth (2001-11)	2844	15.2	15.7	9.76	-19.5	234.0	312	12.5	12.8	8.18	-13.6	41.3
Rural Pop. Growth (2001-11)	2831	13.6	12.6	15.1	-311.1	234.0	308	10.7	9.94	11.2	-51.6	39.7
Paved Road Access (2001)	2787	0.74	0.70	0.24	0	1	301	0.72	0.69	0.25	0.11	1
Nightlights Growth in Prev. Term	2765	22.8	24.1	52.9	-337.7	392.5	304	20.7	20.1	43.6	-105.6	250.0
Log Nightlight Density (t)	2780	1.18	0.84	1.39	-6.49	4.12	307	1.06	0.77	1.35	-4.11	3.55
Winner Log Liabilities	2375	13.2	13.2	1.89	4.46	20.4	287	13.2	12.9	1.97	4.46	17.8
Runner-up Log Liabilities	2398	13.1	13.0	1.86	1.10	20.1	301	13.0	12.9	1.87	2.71	19.1

Variable Descriptions: *Tot. Pop. Growth (2001-11)* is the total population growth between 2001 and2011; *Rural. Pop. Growth (2001-11)* is the rural population growth between 2001 and 2011; *Paved Road Access (2001)* is the share of locations within a constituency with accessibility to paved (pucca) roads; *Nightlights Growth in Prev. Term* is the total nightlight density growth over the previous election term; *Log Nightlight Density (t)* is the log nightlight density at the time of election; *Winner/Runner-up Log Liabilities* is the log liabilities of the winner/runner-up.

Table A9: Continuity Tests for Excluded Covariates

	RD Estimate	Robust P-value	Robust CI	Eff. no of obs.	Opt. Bandwidth
Tot. Pop. Growth (2001-11)	-0.169 (2.347)	0.948	[-4.751, 4.448]	179	10.55
Rural. Pop. Growth (2001-11)	1.458 (2.889)	0.527	[-3.834, 7.491]	135	7.241
Paved Road Access (2001)	-0.039 (0.041)	0.236	[-0.128, 0.032]	183	11.49
Nightlights Growth in Prev. Term	-5.375 (12.227)	0.615	[-30.123, 17.807]	158	9.026
Log Nightlight Density (t)	-0.107 (0.367)	0.788	[-0.817, 0.620]	171	9.647
Winner Log Liabilities	0.245 (0.607)	0.655	[-0.918, 1.460]	168	9.968
Runner-up Log Liabilities	-0.557 (0.688)	0.389	[-1.940, 0.756]	170	10.48

Notes: Continuity tests for the 7 additional covariates excluded from the main analysis, based on local linear estimations within the MSE-optimal bandwidth chosen separately for each variable. All tests use a triangular kernel function and control for election year-state fixed effects. Each row presents point estimate for the covariate, together with the corresponding robust bias-corrected standard errors, confidence intervals and p-values as well as the effective number of observations. The latter is comparatively low since all variables are invariant within election terms and thus only included for the election year. Moreover, these variables contain many missing values. Robust bias-corrected standard errors are presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Continuity Tests - CER-Optimal Bandwidths (Optimal Inference)

Panel A	Constituency-level Characteristics				
	RD Estimate (Robust SE)	Robust P-value	Robust CI	Eff. no of obs.	Opt. Bandwidth
Forest Growth (t-1)	5.960 (5.963)	0.244	[-4.742, 18.633]	209	6.554
Forest Growth (t-2)	-2.349 (7.119)	0.756	[-16.163, 11.742]	211	6.634
Forest Growth (t-3)	-4.397 (8.184)	0.517	[-21.338, 10.743]	257	8.845
Log Const. Area	0.218 (0.287)	0.398	[-0.320, 0.805]	213	6.774
Const. SC Reserved	0.013 (0.093)	0.899	[-0.170, 0.193]	254	8.741
Const. ST Reserved	0.032 (0.058)	0.676	[-0.089, 0.137]	216	7.015
Log Forest Cover (t)	0.016 (0.075)	0.769	[-0.124, 0.168]	232	7.763
Log Electors	0.002 (0.049)	0.934	[-0.092, 0.100]	251	8.485
Log Valid Votes	-0.088* (0.056)	0.0770	[-0.208, 0.011]	167	5.153
Tot. Pop. Growth (2001-11)	-0.271 (2.489)	0.924	[-5.117, 4.639]	145	7.917
Rural. Pop. Growth (2001-11)	0.786 (2.985)	0.746	[-4.884, 6.818]	105	5.437
Paved Road Access (2001)	-0.045 (0.042)	0.231	[-0.132, 0.032]	147	8.639
Nightlights Growth in Prev. Term	-3.806 (12.200)	0.720	[-28.292, 19.530]	129	6.782
Log Nightlight Density (t)	-0.047 (0.402)	0.930	[-0.823, 0.753]	136	7.245
Panel B	Candidate-level Characteristics				
Winner Gender	0.062 (0.073)	0.320	[-0.071 0.216]	242	8.116
Runner-up Gender	-0.047 (0.055)	0.416	[-0.151 0.063]	215	6.921
Winner Age	-3.421 (2.365)	0.164	[-7.924 1.345]	263	9.084
Runner-up Age	0.024 (2.733)	0.967	[-5.469 5.244]	251	8.542
Winner Education	0.094 (0.526)	0.811	[-0.905 1.156]	268	9.830
Runner-up Education	-0.444 (0.589)	0.371	[-1.680 0.627]	244	8.866
Winner Log Assets	-0.458 (0.467)	0.314	[-1.385 0.445]	215	6.914
Runner-up Log Assets	-0.463 (0.398)	0.195	[-1.295 0.264]	196	5.995
Winner Type (ST/SC)	0.031 (0.101)	0.814	[-0.175 0.223]	264	9.165
Runner-up Type (ST/SC)	0.045 (0.100)	0.675	[-0.155 0.239]	266	9.261
Winner Log Liabilities	0.311 (0.630)	0.593	[-0.899 1.572]	138	7.512
Runner-up Log Liabilities	-0.467 (0.700)	0.477	[-1.870 0.874]	137	7.880

Notes: Continuity tests for 27 different constituency- and candidate-level characteristics, based on local linear estimations within the CER-optimal bandwidth chosen separately for each variable. All tests use a triangular kernel function and control for election year-state fixed effects. Each row presents point estimate for the covariate, together with the corresponding robust bias-corrected standard errors, confidence intervals and p-values as well as the effective number of observations. The latter is comparatively low since all variables are invariant within election terms and thus only included for the election year. Robust bias-corrected standard errors are presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Sensitivity to Closest Elections - Formal Results

Donut Hole Radius	RD Estimate	Robust P-value	Robust CI	Opt. Bandwidth	Tot. eff. obs	Excl. obs. left	Excl. obs. right
0	-3.137** (1.537)	0.037	[-6.221, -0.195]	10.49	1164	0	0
0.2	-3.489** (1.708)	0.033	[-6.986, -0.292]	10.08	1100	25	15
0.4	-3.578** (1.567)	0.013	[-6.949, -0.806]	9.039	992	40	20
0.6	-4.025*** (1.612)	0.007	[-7.541, -1.221]	8.912	952	65	35
0.8	-2.284* (1.364)	0.076	[-5.095, 0.253]	12.09	1188	70	55
1	-4.242*** (1.436)	0.001	[-7.387, -1.760]	9.747	988	80	70

Notes: Results are from local linear estimations within MSE-optimal bandwidths selected separately for each donut hole radius, using a triangular kernel function and controlling for state-year fixed effects. Each estimation excludes constituencies for which the absolute value of the victory margin of the criminal candidate is less than the corresponding value in column (1). The other columns present point estimates and robust-bias corrected inferential statistics, as well as the number of observations excluded on each side of the threshold. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A12: Placebo Thresholds - Formal Results

Placebo Cut-off	RD Estimate	Robust P-value	Robust CI	Obs left	Obs right	Opt. Bandwidth
-5	-4.636 (3.752)	0.192	[-12.251, 2.459]	140	200	3.069
-4	3.923 (3.595)	0.130	[-1.602, 12.492]	152	168	2.562
-3	0.882 (2.733)	0.798	[-6.056, 4.657]	172	172	2.575
-2	1.822 (1.917)	0.215	[-1.378, 6.135]	204	144	3.104
-1	-4.860** (3.962)	0.034	[-16.172, -0.642]	288	64	4.189
0	-3.137** (1.537)	0.037	[-6.221, -0.195]	552	612	10.49
1	-2.243 (2.356)	0.917	[-4.373, 4.862]	56	144	2.759
2	-5.015 (2.963)	0.138	[-10.198, 1.416]	104	176	2.816
3	0.757 (1.978)	0.245	[-1.577, 6.178]	168	224	3.117
4	3.148 (2.222)	0.223	[-1.647, 7.064]	176	200	3.036
5	-2.440 (2.226)	0.165	[-7.452, 1.275]	168	152	2.578

Notes: Results are from local linear estimations within MSE-optimal bandwidths selected separately for each placebo cut-off, using a triangular kernel function and controlling for state-year fixed effects. Instead of the true cut-off of 0 (presented for reference), each row presents results from estimations using the 10 different artificial thresholds of column (1). The other columns present point estimates and robust-bias corrected inferential statistics, as well as the number of observations used on each side of the threshold. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A13: Effect Heterogeneity by State Characteristics - Complementary Subsample of States

	Panel A		Panel B		Panel C	
	(1)	(2)	(3)	(4)	(5)	(6)
Criminal Winner	-2.103	-2.160	-2.030	-1.786	-0.974	-0.975
	(1.637)	(1.639)	(1.691)	(1.841)	(1.667)	(1.684)
Robust P-value	0.164	0.159	0.221	0.460	0.580	0.607
Robust CI	[-5.487, 0.929]	[-5.523, 0.902]	[-5.385, 1.244]	[-4.968, 2.248]	[-4.188, 2.345]	[-4.167, 2.435]
Eff. no of obs.	892	812	880	716	936	884
Opt. Bandwidth	13.535	11.944	12.805	9.766	13.509	12.517
Year FE	YES	YES	YES	YES	YES	YES
Year*State FE	YES	YES	YES	YES	YES	YES
Controls	NO	YES	NO	YES	NO	YES

Notes: Notes: Results are from local linear estimations within MSE-optimal bandwidths, using a triangular kernel function and estimated separately different subsamples defined as follows. Panel A: All states except Bihar, Chhattisgarh, Jharkhand, Orissa/Odisha, Uttar Pradesh, and Uttarakhand. Panel B: All states except those ranked as least developed by Ministry of Finance (Arunachal Pradesh, Assam, Bihar, Jharkhand, Orissa/Odisha and Uttar Pradesh). Panel C: All states except those ranked by Transparency International India (TII) as highly corrupt (Tamil Nadu, Haryana, Jharkhand, Assam, and Bihar). Each column presents point estimates and robust bias-corrected standard errors, confidence intervals and p-values. Robust bias-corrected standard errors are clustered at the constituency level and presented in parentheses. Asterisks denote significance levels: *** p<0.01, ** p<0.05, * p<0.1.