WIDER Working Paper 2022/68

Is economic development affected by the leaders' education levels?

Evidence from India

Chandan Jain,[1] Shagun Kashyap,[2] Rahul Lahoti,[3] and Soham Sahoo[4]

June 2022

Abstract: Although formal education is often considered an indicator of political leaders' quality, the evidence on the effectiveness of educated leaders is mixed. Besides, minimum education qualifications are increasingly being used as requirements for contesting elections, making it critical to understand the role of politicians' education in their performance. We investigate the impact of electing an educated politician on economic development in the politician's constituency in India. We use constituency-level panel data on the intensity of night-time lights to measure economic activity. Our identification strategy is based on a regression discontinuity design that exploits quasi-random outcomes of close elections between educated and less-educated politicians. We find that narrowly electing a graduate leader, as compared to a non-graduate leader, in the state assembly constituency increases the growth rate of night-time lights by about three percentage points in the constituency. As pathways, we find that graduate leaders improve the provision of roads, electricity, and power; however, they do not significantly impact the overall provision of public goods. In comparison with findings from other studies in the literature, our result suggests that the impact of formal education of the leader is weaker than the leader's other characteristics, such as gender or criminality.

Key words: educated political leaders, night-time lights, close elections, public goods, India

JEL classification: D72, H11, H41, O40

Acknowledgements: We thank Emmerich Davies and participants in the American Political Science Association Annual Conference 2020 for helpful discussions. We acknowledge research support received from IIM Bangalore through a Research Seed Grant. We also acknowledge IZA for publishing a version of this work as discussion paper (DP 15278).

[1] International Initiative for Impact Evaluation (3ie), New Delhi, India; [2] Indian Institute for Management Bangalore, Bengaluru, India; [3] UNU-WIDER, Helsinki, Finland; [4] Indian Institute of Management Bangalore, Bengaluru, India, and Institute of Labor Economics (IZA), Bonn, Germany, corresponding author: soham.sahoo@iimb.ac.in

This study is published within the UNU-WIDER project Academic excellence.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-199-0

https://doi.org/10.35188/UNU-WIDER/2022/199-0

Typescript prepared by Siméon Rapin.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

In recent literature, education has been considered a marker of leaders' quality, assuming that formal education equips leaders with the necessary skill and competence to execute complex functions associated with holding a political office. While much of empirical research has used leaders' formal education level as a proxy for leadership quality, the evidence supporting this claim is ambiguous. Some studies claim a direct link (Besley et al. 2011; Peveri 2021; Yu and Jong-A-Pin 2020), while others suggest that it is tenuous (Bastos and Sánchez 2021; Carnes and Lupu 2016; Curto-Grau and Gallego 2019). The issue of the impact of educated leaders becomes far more relevant in a setting like India where minimum education requirements are increasingly being mandated for contesting elections. These mandates have been justified by courts and legislatures on the basis of assumed higher efficiency of educated political leaders, without much supporting evidence.

The existing literature has highlighted several reasons why educated leaders, compared to less educated leaders, may have an ambiguous effect on economic development. Educated legislators may promote growth if education influences their policy preferences, technical skills, or dedication to serving public interest (Besley et al. 2011; Congleton and Zhang 2013; Dreher et al. 2009). They may drive more investment from businesses leading to higher growth (François et al. 2020) or improve chances of reforms as educated leaders understand them better (Dreher et al. 2009). On the other hand, education may impart human capital, but that may not transfer into political ability to carry out policy tasks efficiently (Carnes and Lupu 2016; Lahoti and Sahoo 2020), or there may be a mismatch between preferences of educated leaders and their constituents (Curto-Grau and Gallego 2019).

In this paper, we investigate whether politicians with higher levels of formal education lend to better economic development outcomes in areas under their jurisdiction. To explore the different pathways, we also analyse if educated leaders who have spent time in office create developmental infrastructure or reduce crime that ultimately affects the level of economic activity in a constituency.

To shed light on this issue, we use data on elections of leaders to state legislatures in India. We assemble a dataset on the educational qualifications of candidates contesting elections along with the election results. For the main outcome variable, we use the intensity of night-time lights captured in satellite imagery. In absence of reliable data that can measure economic development at the assembly constituency level, this measure serves as a proxy variable. Previous studies have established the use of nigh-time lights as a proxy for measuring sub-national GDP (Chen and Nordhaus 2011; Doll et al. 2006; Donaldson and Storeygard 2016; Henderson et al. 2012). Night-time lights data are available annually at a spatially granular level and can be aggregated to provide estimates at the constituency level. Combining data on leaders' education and night-time lights outcome, we construct a comprehensive panel data at the constituency level for the period 2009–13.[1]

A challenge in estimating the causal effect of legislators' education is that election of educated leaders to state legislature might be non-random and driven by unobserved voter preferences and other constituency-level characteristics, thus making it an endogenous process. To overcome this challenge and estimate the casual effect of educated leaders, we use a regression discontinuity design (RDD) based on close elections between educated and less-educated candidates. The identification strategy relies on the quasi-random outcomes of close elections; this set-up has been widely used in the related literature (Eggers et al. 2015; Lee 2008). Comparing constituencies where an educated leader wins by a narrow margin with constituencies where a less-educated leader wins by a narrow margin, we can isolate the casual effect of educated legislator.

[1] The choice of this period for the main analysis is due to restrictions in the availability of data for our key variables. We discuss this aspect further in the later sections.

Using the RDD we find that, on average, the election of a graduate leader results in approximately three percentage points higher growth rate of annual average intensity of night-time lights in the constituency. The point estimates are stable and robust across alternate bandwidths and to the inclusion of fixed effects that control for unobserved heterogeneity at the regional level. Our results are also robust to alternate specifications and exclusion of extreme values of the outcome variable. Using existing estimates from the literature on the elasticity of GDP growth to night-time lights growth (Baskaran et al. 2021; Bickenbach et al. 2016), we find that the election of an educated leader results in roughly 0.3 to 0.45 percentage points higher GDP growth rate per year in the constituency. Additionally, we find that the impact of educated leaders is driven by leaders who have a college-graduate degree and not by those who have only completed schooling. The estimate also varies by the initial level of development in the states, with the effect being larger in the least developed states. We also explore further heterogeneity in the effect of graduate leaders based on their gender, age, criminality, and affiliation with the state ruling party.

We examine the internal validity of our empirical strategy by performing a range of checks on the RD design. We show that constituency-level pre-determined characteristics are balanced around the threshold of discontinuity. Besides, we find no evidence to suggest that educated leaders are more likely to win or lose in close elections, implying that the outcome of close elections are not manipulated. We also conduct balance check of candidate characteristics to isolate the effect of education from any other characteristics that may be correlated. In addition, we conduct various placebo tests and find that the placebo estimates are insignificant and smaller than the true estimate. These analyses help us establish that the outcome of close elections are indeed quasi-random and our estimates reflect the causal effect of leaders' education. Further, we inspect the external validity and broader applicability of our results beyond close elections. We compare a range of candidate and constituency-level characteristics across close and non-close elections and mixed and non-mixed elections, finding no substantive differences in them.

We also explore the mechanisms through which educated leaders may impact economic growth; specifically, we examine the impact of educated leaders on the provision of infrastructure facilities, crimes, and corruption. We find that the election of graduate legislators result in around 0.3 standard deviation increase in access to both roads and power supply. Similarly, the percentage of households that have access to electricity as the main source of lighting is 5.3 percentage points higher if a graduate candidate is elected. With regard to crime outcome, we find that a ten percentage point increase in the proportion of graduate legislators in a district on average causes a 1.2 to 1.9 per cent decline in the reported crime in the constituency. This result indicates that an educated legislator may lead to better environment for economic activity and growth. We do not find any statistically significant difference between graduate and non-graduate legislators' growth in private assets which we use as a proxy for corruption following Fisman et al. (2014).

Comparing our results to the existing literature on leaders' identity in the Indian context, we find that the effect of leader's education on economic growth is relatively smaller. Prakash et al. (2019) show that electing a criminally accused politician lowers the growth rate in night-time lights by about 24 percentage points, and Baskaran et al. (2021) estimate that electing a woman legislator increases growth in night-time lights by about 15 percentage points. As compared to these studies that follow a similar identification strategy and outcome variable, we find that electing an educated legislator has a relatively small impact of only a three percentage point increase in night-time lights.

Our paper contributes to various strands of literature. First and most directly, our paper contributes to the literature exploring whether education is an important marker in determining the skill and competence of leaders holding public office. Specifically, our findings contribute to discover whether educated leaders enhance economic development in their jurisdiction by facilitating provision of public infrastructure (Bastos and Sánchez 2021; Besley et al. 2011; Carnes and Lupu 2016; Curto-Grau and Gallego 2019; Lahoti and Sahoo 2020; Mitra 2020). Our estimates based on RD specification and data from a developing country adds to the existing body of evidence. To the best of our knowledge, Lahoti and Sahoo (2020) is the only other paper that analyses the effect of educated leaders in the Indian context. They investigate the impact of graduate legislators on quality of education in schools in the legislators' constituencies. In contrast, we analyse the impact on economic growth and its channels. Moreover, our study suggests that with regard to the level of formal education of the leader, attainment of a college-graduate degree is the main differentiating factor.

Secondly, our findings add to the literature on the role of identity and characteristics of leaders in development and policy outcomes. We find evidence in support of citizen candidate models put forth in Besley and Coate (1997) and Osborne and Slivinski (1996). These models suggest that demographic characteristics contribute to the identity of a leader and it has potential to significantly impact policy outcomes (Besley 2005). Candidate characteristics that have been found to matter include gender (Bhalotra and Clots-Figueras 2014; Bhalotra et al. 2014, 2018; Chattopadhyay and Duflo 2004), caste (Pande 2003), criminality (Prakash et al. 2019), political affiliation (Gulzar and Pasquale 2017; Hill and Jones 2017), and political alignment (Asher and Novosad 2017). We add to this strand by considering education to be an important dimension of a leader's identity and add evidence to the related literature (Besley et al. 2011; Carnes and Lupu 2016; Lahoti and Sahoo 2020).

Our analysis is motivated by policy measures being considered across Indian states to institute a mandate for minimum educational qualifications for candidates contesting elections. Haryana[2] and Rajasthan[3] put in place policies mandating minimum educational requirements for candidates contesting local body elections, and Assam and Maharashtra have considered incorporating a similar policy.[4] Policy mandates for minimum education are crucial as they have implications for equality and representation, especially for marginalized sections of the society.[5] Any restriction placed on limiting the ability of individuals to contest has been considered undemocratic, elitist, and discriminatory.[6] It also affects political selection by changing the pool and subsequent selection of candidates contesting elections. Afzal (2014) and Curto-Grau and Gallego (2019) find that educational elitism has consequences for representation. Altering the identity of candidates who will hold office may change the nature of policies and credibility with which it is implemented (Arora 2022; Besley 2005). In this regard, although our paper bypasses the aspect of political selection by exploiting quasi-random variations emanating from close elections, we show that the education level of leaders may have a limited impact compared to other characteristics of leaders. Thus, ballot access restriction by educational mandates may not have the intended consequence of vastly improving efficiency of the elected legislators.

The rest of the paper is organized as follows. In Section 2, we provide a note on background and context, in Section 3 we provide a description of the dataset, and in Section 4 we lay out the identification strategy. Section 5 evidences the validity of the RD design, and Sections 6, 7, and 8 discuss the results. In Section 9, we conduct robustness and sensitivity checks on results from our main specification. In Section

[2] For Haryana, the Panchayati Raj Amendment Act (2015) mandates general candidates to have completed at least ten years of schooling, women and Dalit candidates to have completed eight years, and Dalit women candidates to have completed at least five years of schooling to contest in local body elections. See https://www.thehindu.com/news/national/other-states/minimum-qualification-set-as-haryana-passes-panchayati-raj-bill/article7626719.ece.

[3] Rajasthan has since scrapped this policy, which was in place for one election term. See https://www.thehindu.com/news/national/other-states/rajasthan-to-scrap-education-criterion/article26241205.ece.

[4] https://www.business-standard.com/article/pti-stories/assam-to-move-resolution-for-educational-qualification-to-118031401158_1.html and https://timesofindia.indiatimes.com/city/mumbai/directly-elected-sarpanch-now-schooling-till-class-vii-must/articleshow/59432314.cms

[5] For instance, in Haryana, this mandate has led to disqualifying 68% of Dalit women and over 50% of all women from contesting panchayat elections (Bhaskar 2016).

[6] https://indianexpress.com/article/opinion/columns/narendra-modi-pm-degree-arvind-kejriwal-delhi-university-ba-degrees-of-exclusion-2792374/

10 we discuss external validity of our results. Finally, in Section 11, we discuss the relevance of our findings.

2 Background

2.1 Conceptual framework

From a theoretical point of view, our paper contributes to the body of literature suggesting that identity and characteristics of the leaders matter for policy outcomes (Besley 2005). The early literature on elected leaders followed a Downsian approach conceptualizing that candidates once elected to office fully commit to the policy preferences of the electorate (Downs 1957). In such a model, leaders are willing to adopt the policy preferences of the median voter and subsequently implement it. In contrast to the Downsian approach, the citizen candidate models in the new political economy literature focus on the identity of politicians, assuming that in a representative democracy politicians are elected from a pool of candidates with varied preferences. Once elected, they implement policies according to their preference which may not align with the preference of the median voter. This idea, theoretically formalized in Besley and Coate (1997) and Osborne and Slivinski (1996), received empirical support by studies linking leaders' identity such as caste (Besley et al. 2004; Pande 2003), gender (Chattopadhyay and Duflo 2004), or political affiliation (Hill and Jones 2017) to differences in the choice of policy outcomes.

An important dimension that captures a leader's identity and quality is education. Leaders with higher levels of education may be more skilled, competent and exhibit high levels of civic awareness. Educated leaders may be equipped to drive change through two channels. On the one hand, they may have a preference for certain policies and intrinsically value development. If so, they may implement policies in the broader interests of the electorate. On the other hand, education may have imparted a level of competency allowing them to ensure the effective implementation of policies while they are in power (Besley et al. 2011; Persson and Tabellini 2002). If these skills and values are important for leaders holding office, then it is likely that more education will result in more efficient leaders. Alternatively, it is also possible that while formal education may equip individuals with more human capital, it may not be relevant for discharging responsibilities associated with a political office (Carnes and Lupu 2016).

The evidence on the effect of leaders' education has been inconclusive. Studies have found that a leader's education level matters for economic performance (Besley et al. 2011; Congleton and Zhang 2013; Peveri 2021; Yu and Jong-A-Pin 2020), educational attainment of citizens (Diaz-Serrano and Pérez 2013; Martinez-Bravo 2017), and investment in public goods (Mitra 2020). On the other hand, in an extensive analysis, Carnes and Lupu (2016) find no evidence of electing graduate leaders on measures of economic performance, re-election, and corruption. Bastos and Sánchez (2021) find that in the context of Brazil, educated leaders do not produce better socio-economic outcomes and they also do not have higher likelihood of being re-elected, implying that educated leaders may not be perceived as better leaders by the citizens. Lahoti and Sahoo (2020) find no effect of leaders' education on educational outcomes at the all-India level. Curto-Grau and Gallego (2019) suggest that there is a possibility of mismatch in preferences of more educated leaders and their constituents, resulting in policies that may be far from being pro-social. In light of theoretical ambiguity and conflicting evidence on the subject of leaders' education, this remains an empirically interesting question to explore.

2.2 Political structure of India

In this paper, we analyse data on state assembly constituencies and representatives elected to state legislatures in India. Below, we provide a brief context on the political structure in India.

India is a federal republic consisting of 28 States and eight Union Territories, and with a parliamentary system for government. At the national level, elected representatives to the Parliament of India are referred to as the Members of Parliament (MPs). Similarly, at the state level, elections are held to the State Legislative Assembly (or Vidhan Sabha). The elected representatives to state legislatures are referred to as Members of Legislative Assembly (MLAs). The states are sub-divided into assembly constituencies (each of which is represented by an MLA) based on population to ensure equal representation. Owing to the federal system of governance, every state has autonomy in deliberating upon and taking legislative action on certain subjects.[7] Elections, at both the state and national levels, are conducted once every five years, with the electoral cycles being asynchronous across the states. The elections follow a 'first-past-the-post' electoral system and once elected, MLAs hold office for a term of five years. In this paper, our focus is on MLAs.

2.3 Importance of the Members of Legislative Assembly

Elected MLAs have legislative, executive, and financial powers conferred upon them by the Indian Constitution (Article 246).[8] All items of the state and concurrent lists fall under the purview of their responsibilities as elected officials, to debate and legislate upon. MLAs influence policy and facilitate change directly or indirectly. Directly, MLAs can ensure the provision of public goods by lobbying for government funds for their constituencies or by allocating funds according to their discretion. Under the Local Area Development Scheme (LADS) introduced in 1994–95, MLAs have access to an annual fund that they can use as per their discretion to sanction construction of public goods and assets such as schools, community buildings, roads and bridges, water and sanitation facilities, etc.[9] Indirectly, as representatives for their respective constituents, MLAs debate and vote on legislation that may positively affect the level of economic activity and development in their respective assembly constituencies. Additionally, they can provide oversight to the schemes and programmes implemented by the executive, ensuring that developmental works are completed in time. They can also leverage their position to influence bureaucrats to improve implementation of local development schemes that ensure electoral benefits (Gulzar and Pasquale 2017).

The existing literature has highlighted the various characteristics of MLAs that matter for development outcomes. For instance, while Baskaran et al. (2021) find that election of female MLAs results in higher economic growth, Prakash et al. (2019) show that electing criminally accused MLAs hampers economic growth measured by night-time lights. Multiple studies establish the role of MLAs for a host of development outcomes (Asher and Novosad 2017; Bhalotra and Clots-Figueras 2014; Bhalotra et al. 2014, 2018; Baskaran et al. 2021; Burchi 2013; Clots-Figueras 2011, 2012; Ejaz Ghani et al. 2013; Halim et al. 2016; Iyer et al. 2012; O'Connell 2018; Pande 2003). While gender, caste, religion, and criminality of leaders have been extensively studied, the effect of educational qualification of leaders has not been adequately explored in the Indian context. Lahoti and Sahoo (2020) is one study that analyses the effect of leaders' education on educational outcomes, finding heterogeneity in the effect based on the initial level of development of the states.

[7] The subjects where the states can independently legislate upon are laid out in the 7th Schedule of the Indian Constitution.

[8] https://legislative.gov.in/constitution-of-india

[9] The amount accessible to MLAs varies across states. For example, MLAs in Karnataka have access to an annual sum of five crore rupees (INR), while those in Delhi have access to an annual sum if ten crore rupees.

3 Data

For our analysis we merge data from various sources. In the following sub-sections we describe the different datasets used to measure the outcomes and main explanatory variables used in our analysis.

3.1 Night-time lights data

There is no reliable dataset that measures economic activity or output at the assembly constituency level for India. National Income Accounts data are available at the national and state levels only, and the accuracy of these data is often debated (Subramaniam 2019). Alternatively, the existing sample surveys may provide indicators for development outcomes at the district level; however, even these datasets are not available on an annual basis and are not always representative of the district population. The night-time lights data on the other hand, have the twin benefits of being available on an annual basis and being spatially granular, allowing us to construct time series data at a sub-regional level. Hence, we proxy for the level of economic activity in an assembly constituency by the intensity of night-time lights. Doll et al. (2006), Chen and Nordhaus (2011), and Henderson et al. (2012) have assessed the reliability of night-time lights as a proxy, analysing the relationship between the intensity of night-time lights and regional GDP. Henderson et al. (2012), for instance, note that night-time lights serve as a useful proxy for economic growth as well as for measuring the short-run fluctuations in growth. Recent studies, in this regard, have used night-time lights to predict various developmental outcomes (Alesina et al. 2016; Bruederle and Hodler 2018; Hodler and Raschky 2014; Michalopoulos and Papaioannou 2014; Storeygard 2016; Weidmann and Schutte 2017).[10] Bruederle and Hodler (2018) and Weidmann and Schutte (2017) both note that a higher intensity of night-time lights is associated with better development outcomes at the local level. Within the Indian context as well, studies such as Bhandari and Roychowdhury (2011) and Bickenbach et al. (2016) have examined the association between night-time lights and district-level GDP, reporting a high correlation between the two. Prakash et al. (2019) and Baskaran et al. (2021) use night-time lights data for India at the regional level to study the impact of leaders' gender and criminality, respectively.

We use imagery of earth gathered at night by multiple satellites orbiting the earth under the U.S. Air Force Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS). The DMSP-OLS is equipped with the unique ability to capture low-light imaging data. Since 1992, the data have been processed to remove the effect of late evening sunlight, moonlight, and the presence of clouds from the images. Further, the images are filtered to remove the effect of forest fires and other ephemeral lights, and background noise is removed by setting thresholds based on visible band values found in areas known to be free of detectable lights. The images are projected onto geo-referenced 30 arc-second grid (roughly $1km^2$ at the equator) pixels. Each pixel is encoded with a digital number (DN) that signifies annual average brightness on a 6-bit scale from 0 to 63, where 0 signifies absence of detectable light and 63 signifies maximum measurable value for night-time light. The higher the number, the greater is the intensity of night light.[11]

To estimate light intensity at constituency level, we overlay a map of assembly constituency boundaries on the night-time lights file. We extract the sum of pixel values within the constituency boundary and divide by the total area of the constituency. This gives us annual estimates of mean night-time lights for constituencies. Since the DMSP-OLS programme was discontinued in 2014, data on night-time lights are available for the period 1992–2013. Another aspect of the data is that the sensors measuring

[10] Further, Donaldson and Storeygard (2016) discuss various studies that have used night-time lights in different contexts in economics.

[11] Henderson et al. (2012) detail a discussion on the processing of night-time images captured by the satellite into usable night-time lights and mention the limitations of interpreting economic activity from night-time lights.

night-time lights might generate measurement error because of saturation and low sensitivity. Saturation occurs because of the limitation in the sensors in recording high levels of brightness as the highest value recorded by the sensors is censored at 63. This might lead to underestimation in growth in large cities which have already reached the saturation value of night-time lights. To overcome this limitation, we conduct robustness analysis by dropping all observations having these extreme values.

3.2 Political data

We merge candidate level data from two different sources to construct a comprehensive political dataset for all candidates contesting state assembly elections. We use detailed data on assembly constituency results from reports provided by the Election Commission of India (ECI). This includes information on all contesting candidates, affiliated political parties, the number of votes polled for each candidate, and the size of electorate in each assembly constituency.

Following a Supreme Court judgment in 2003, all individuals contesting elections are required to file an affidavit with the ECI listing their education level, assets, and criminal cases, among other details.[12] This information has been processed and made available by the Association for Democratic Reforms (ADR), an election watchdog.[13] We merge the data from ECI and ADR at the candidate level to compile panel data on election outcome for each individual candidate and their characteristics, including the number of votes polled in the election and their educational qualifications.

3.3 Merged data

The main focus of our paper is on the educational qualification of the contesting candidates. Given that the Supreme Court judgement was implemented in 2004, we have information on candidates' educational qualification post 2004 only. In addition, elections to State Legislative Assembly take place once in five years, and the election cycles are different across the states. Therefore, not all states have information on education levels of candidates from 2004 onwards; rather, this information becomes available when a state has election for the first time in the post-2004 period. Thus, in our data, 2009 is the initial year when we have information on the education level of MLAs in power for all the states.

Additionally, the Delimitation Act, passed in 2002, amended the assembly constituency boundaries based on population figures from the 2001 census.[14] After much administrative delay, it was implemented from 2008 onwards, implying that elections post 2008 followed newly drawn assembly constituency boundaries. This means that constituencies are not comparable across elections that took place in the pre-delimitation and post-delimitation periods. To tackle this issue, we take into consideration the constituency boundaries prior to delimitation as well as after it and merge the political data with the night-time lights data for the respective year depending on whether delimitation was already implemented or not.[15] We construct an annual panel by merging the data on leader who is in power in a given year with the night-time lights outcome corresponding to that year. Thus, we have two sets of annual

---

[12] https://globalfreedomofexpression.columbia.edu/cases/union-india-uoi-v-respondent-association-democratic-reforms-another-peoples-union-civil-liberties-pucl-another-v-union-india-uoi-another/

[13] ADR has processed these affidavits from different state elections over time. The same are also available on the ECI website in PDF format.

[14] The constituency boundaries account for the population. Given changes in population growth, the boundaries have been historically adjusted to ensure equal representation in legislature through the Delimitation Act of 1952, 1963, 1973, and 1977; following each decadal census. The last delimitation was conducted in 1977, before new boundaries were proposed in the Delimitation Act of 2002. Further, Iyer and Reddy (2013) study the delimitation enforcement in Rajasthan and find the redrawing of boundaries to be politically neutral.

[15] We use shapefiles for both pre- and post-delimitation assembly constituency boundaries and extract night-time lights data using QGIS. These data are then merged with the political data for the respective assembly constituencies. The post-delimitation

panel data at the constituency level, specific to pre- and post-delimitation periods. For our analysis, we pool these two datasets.[16] As explained earlier, owing to varied electoral cycles of states, we have political data on candidates from 2009 onwards. Discontinuation of the DMSP-OLS programme from 2014 onwards results in night-time lights data being available till 2013. Thus, our analysis is restricted to the period 2009–13. We report the descriptive statistics of the key variables in Table 1.

3.4 Additional outcome variables

Data on public goods and infrastructure

In addition to investigating the relationship between educated leaders and night-time lights, we also explore the mechanisms through which educated leaders might have an effect on economic growth. Specifically, we explore the effect of electing educated leaders on the availability of public goods and infrastructure in the assembly constituency. For this purpose, we use village-level data from the 2011 round of census.[17] From the census data, we obtain information on access to public goods and infrastructure such as schools, colleges, nutritional centres, hospitals, drinking water, sanitation facilities, financial institutions, roads connectivity, power supply, and electrification at the village level.[18] Provision of public goods and infrastructure might be influenced by the MLA in a constituency through channels discussed in Section 2.3.

Since our analysis is at the constituency level, we match villages from the census data to assembly constituencies.[19] We aggregate the village-level data at the constituency level to construct indices measuring availability of public goods and infrastructure. Table A1 provides information on the proportion of villages in a constituency that have access to a specific infrastructure facility captured in the data. Electricity, however, measures the percentage of households in an assembly constituency that have access to electricity as the main source of lighting. Based on these variables, we construct three indices at the constituency level using principal component analysis (PCA). The first index captures overall access to public goods, considering all the variables mentioned in Table A1. The second index considers access to power supply, constructed using information on access to power for domestic, agricultural, and industrial use. The third index is for road infrastructure in the constituency, constructed using variables that measure access to national highways, state highways, paved roads, unpaved roads, all-weather roads, and other major roads that enable connectivity to agricultural markets and administrative offices. All the indices are standardized to have mean 0 and standard deviation 1. Additionally, we also consider an outcome variable measuring the percentage of households having access to electricity in a constituency. We merge these variables with the election data at the constituency level. Unlike our night-time lights outcome where we have annual panel data at the assembly constituency level, the data from census are a single cross section. Therefore, for the additional analysis, we estimate a cross-sectional regression where leader of the previous year is used to estimate the effect.[20]

---

shapefiles used have been sourced from Susewind (2014). The pre-delimitation shapefiles were made publicly available by Sandip Sukhtankar and Manasa Patnam; see https://uva.theopenscholar.com/sandip-sukhtankar/data.

[16] The pooled data essentially are constituency-year panel data that are unbalanced due to the non-comparability of pre- and post-delimitation constituencies. Although constituency boundaries have changed after delimitation, no constituency is spread across multiple districts in both pre- and post-delimitation period. Since constituencies are always contained within districts, we include district fixed effects in our regression presented in the next section.

[17] The census data can be accessed from: https://censusindia.gov.in/2011-Common/Archive.html.

[18] Power includes access to electricity for domestic, agricultural, and commercial uses.

[19] We overlay village shapefiles on shapefiles for assembly constituency boundaries to match villages to assembly constituencies using QGIS.

[20] Here, the leader holds office in 2010, a year before Census 2011 was conducted. However, these leaders could have been elected to office at any time in the past five years. This gives the leader more than a year in office to effect change. In Table A2,

Data on crime

To investigate if educated leaders uphold law and order in the constituency, we use the total number of documented crimes as an outcome variable. This variable is sourced from the 'Crime in India' publication of the National Crime Records Bureau (NCRB), Government of India. We consider any crime under the Indian Penal Code, recorded at the police station in the form of First Information Report (FIR), and aggregated at the district level. Thus, we have district-level annual panel data on total crimes for the period from 2005 to 2018. We merge this variable with the political data aggregated at the district level to conduct the analysis.

Data on politicians' assets

We also consider an outcome variable that measures the growth in net asset of leaders over an election cycle. Candidates contesting elections must declare their assets and liabilities, in addition to other demographic details, per law mentioned in Section 3.2. However, asset growth can only be estimated for candidates who decide to re-contest and file affidavits in two subsequent elections. This limits the analysis to a subset of leaders who re-contest.

We consider data on leaders' assets and liabilities declared in election affidavits for all state elections in the period 2004–17. We consider 30 states, covering the asset growth of leaders who held political office for two election cycles in 17 states and one election cycle in the remaining 13 states. We then match these leaders by their names with candidates who re-contested in subsequent elections in the same state. In this process, we use a probabilistic matching algorithm first and then manually verify the matches, as the commonalities between Indian names and different spellings in affidavits make it hard to find exact matches based on candidate names. We were able to match 4,476 leaders out of a total of 5,689 leaders re-contesting elections.

4 Empirical strategy

In this section, we present our empirical strategy for estimating the causal effect of electing an educated leader on the economic activity in a constituency. We use the annual average growth rate of night-time lights in the constituency as our main outcome variable, defined below:

$$Y_{idst} = log(l_{idst+1}) - log(l_{idst}) \tag{1}$$

l_{idst} measures the annual average intensity of night-time luminosity in assembly constituency i belonging to district d from state s in time period t.[21] The outcome variable Y_{idst} captures the growth rate of night-time lights in assembly constituency i in between the periods t and $t+1$. For each assembly constituency i, we use data on night-time lights for the period 2009–13.

Using constituency-year-level data, our basic specification exploring whether an educated MLA affects the growth rate of night-time lights in the constituency is given by the following equation:

$$Y_{idst} = \alpha + \delta_d + \gamma_t + \beta \, Graduate_{idst} + {}_{idst} \tag{2}$$

we present the distribution of years in which leaders holding office in 2010 were elected, and our analysis controls for these election-year fixed effects.

[21] We add 1 to the value of night lights before taking the logarithm, since there are some observations with the value 0 (1.78% of the sample). We later show that our results are not sensitive to dropping extreme values of 0 and 63.

where t identifies the time period while i, d, and s identify a constituency, district, and state, respectively. Our interest lies in the coefficient of the main explanatory variable $Graduate_{idst}$, taking the value 1 if the elected leader in constituency i in district d from state s at time t has a graduate degree and 0 otherwise. We define a graduate leader as one who has completed college education, using this as a cut-off to differentiate between a more educated leader from a less educated one.[22] In further analysis, we also consider an alternative threshold of education, i.e. school completion (at least 12 years of education), to define an educated leader. The coefficient β captures the marginal effect of having a graduate leader in constituency i during time period t on the growth of average night-time lights between time periods t and $t+1$. Thus, our model is similar to the one estimated by Prakash et al. (2019) to determine the effect of criminally accused leaders on economic growth. In the set of explanatory variables, we include district fixed effects (δ_d) to control for various time-invariant unobserved factors at the district level and year fixed effects (γ_t) to capture any secular change in the growth rate of night-time lights over time. Note that district fixed effects also subsume any unobserved heterogeneity at the state level, since districts are nested within states. To check the robustness of our results, we also present alternate estimates dropping the fixed effects in one specification, and including only state and year fixed effects in another specification.

The main identification problem in Equation (2) is that the election of an educated leader may not be random, and constituencies that elect an educated leader may systematically differ from those that do not. For example, constituencies with more educated electorate may have higher economic growth and also elect an educated politician. Thus, the election of an educated leader is likely to be correlated with unobserved voter- and constituency-specific characteristics, making the ordinary least squares (OLS) estimate biased and inconsistent.

4.1 Regression discontinuity design

To identify the causal effect of educated leaders, we utilize a regression discontinuity design (RDD). In a first-past-the-post electoral system, among the top two contestants, the probability that a candidate wins is a function of the vote margin between the candidate and the opponent; this probability changes discontinuously at the point where vote margin is zero (Lee 2008). Given the objective of this study, we consider elections where one of the top two candidates is a graduate while the other is not, and define the forcing variable as the vote margin between the graduate and the non-graduate candidate. The RDD framework assumes that in a close neighbourhood of the discontinuity at zero margin, the difference between constituencies that elect a graduate leader and those that elect a non-graduate leader diminishes as the margin becomes smaller. Therefore, constituencies that barely elect a non-graduate candidate in a close election serves as a valid counterfactual for constituencies that barely elect a graduate candidate in a close election. Thus, the RDD framework allows us to estimate the causal effect of graduate leaders on outcomes.

Accordingly, we define the forcing variable $Margin_{idst}$ as the difference in vote shares between the graduate and the non-graduate candidate for every election where either the winner or the runner-up is a graduate and the other is not. Owing to this definition:

$$Graduate_{idst} = \begin{cases} 1, & \text{if } Margin_{idst} > 0 \\ 0, & \text{if } Margin_{idst} \leq 0 \end{cases}$$

[22] The distribution of education level of candidates (considering the winner and runner-up for each election) is presented in Table A3. Nearly 61 percent of the top two candidates contesting in election have a graduate degree (i.e., at least 15 years of education) in our data; hence, compared to other education levels, the cutoff of graduate degree allows us to have a more balanced distribution between more educated and less educated leaders.

where $Margin_{idst} > 0$ implies that the graduate candidate has won the election, while $Margin_{idst} \leq 0$ would imply that the graduate candidate has lost. Thus, the probability of having a graduate leader changes discontinuously at the cut-off where $Margin_{idst} = 0$. We exploit this discontinuity in treatment assignment to identify the causal effect of electing a graduate leader on economic activity in the constituency. Given this set-up, we estimate the following specification using constituencies that lie in the close neighbourhood (h) around the cut-off ($Margin_{idst} = 0$).

Finally, the estimated model as per the RDD specification is:

$$\begin{aligned} Y_{idst} \ = \ \alpha \ + \ \delta_d \ + \ \gamma_t \ + \ \beta \ Graduate_{idst} \ + \ f(Margin_{idst}) \ + \ {}_{idst}, \\ \forall \ Margin_{idst} \in (-h, h) \end{aligned} \tag{3}$$

where Y_{idst}, δ_d, γ_t, and $Graduate_{idst}$ are all defined the same way as in Equation (2). By construction, $Margin_{idst}$ is positive when a graduate leader wins against a non-graduate leader and negative the other way around.

We estimate the model using local linear regression with triangular kernel as suggested by Gelman and Imbens (2019). Concurrently, we also report results for several bandwidth choices ($2h, h/2$) including the optimal bandwidth procedure suggested in Imbens and Kalyanaraman (2012) and Calonico et al. (2014).

5 Validity of the RD design

In this section, we provide supporting evidence that the outcomes of close elections are quasi-random and other validity tests for the RD design. We conduct the various tests suggested by Imbens and Lemieux (2008). The first is a density test that investigates any evidence of sorting around the threshold (McCrary 2008). For instance, if graduate politicians have the means to manipulate election outcomes, they may be more likely to win close elections against non-graduate politicians. If it were so, we would observe a larger frequency of graduate candidates compared to non-graduate candidates in the neighbourhood of the cut-off of forcing variable, which in our analysis is margin of victory. The upper and lower panels in Figure 1 plot the density of the forcing variable and the density test, respectively. Both panels of Figure 1 show that the difference in density of the victory margin above and below the cut-off is not statistically significant, indicating that there is no evidence of manipulation in the outcome of a close election.

The second test investigates whether pre-determined constituency-level characteristics change at the threshold of discontinuity. If the RDD is internally valid, these factors should not change discontinuously at the cut-off. We conduct this test on a set of constituency-level variables including growth rate of night-time lights in the previous election term, whether the constituency was reserved for Scheduled Caste or Scheduled Tribe candidates (SC/ST), total number of contestants, size of electorate, number of voters, and poll percentage. Figure 2 shows that indeed, there is no significant discontinuity in these characteristics around the cut-off of zero margin of victory between the graduate and non-graduate candidates. Analogously, considering the sample of close elections involving a graduate candidate and a non-graduate candidate, we conduct a t-test to compare the means of these characteristics between constituencies with a graduate winner and a non-graduate winner. Consistent with the graphical evidence, results from Table A4 show no significant difference in these covariates across the two types of constituencies. In addition, we also find that the probability of a graduate candidate winning in close elections does not systematically depend on the constituency characteristics (Table A5), indicating that the outcomes of close elections are indeed random.

Further, we consider the possibility that the education of a leader may be correlated with other characteristics of the leader. If other characteristics of the leader also change discontinuously at the cut-off, then we will not be able to separate out the causal effect of a leader's education from these other characteristics. Figure 3 shows continuity checks in terms of the leaders' gender, age, criminality, ruling party affiliation, assets, and liabilities. The graphical evidence does not indicate any discontinuous change in these characteristics at the cut-off. The analogous comparison of means presented in Table A5 indicates minor differences in terms of the leader's criminality and gender; however, we later show in robustness analysis that our estimate of the effect of the leader's education is unlikely to be driven by these factors.

6 Results

6.1 Main results

Our main results are based on the RDD specification defined by Equation (3).[23] We find that electing a graduate leader in close elections results in higher growth in night-time lights, and the finding is robust across specifications. We report the point estimates in Table 2. First, we estimate the model without any fixed effects and present the results in Panel A. Results in Panel B are from a specification that includes year and state fixed effects. Our preferred specification, presented in Panel C, controls for district fixed effects and thus takes into account any unobserved heterogeneity at the level of state, district, and year.

In Column (1), we estimate the model using the optimal bandwidth criterion proposed by (Imbens and Kalyanaraman 2012) (IK). According to our preferred specification, there is a 3.28 percentage point increase in annual average growth rate of night-time lights resulting from electing a graduate candidate over a non-graduate candidate in close elections. This estimate is statistically significant at 1 per cent level of significance. We show estimates using alternate bandwidths as well. Using the optimal bandwidth criterion of Calonico et al. (2014) (CCT), Column (2) shows an effect size of 3.1 percentage points. In the last two columns, we double $(2h)$ and halve $(h/2)$ the IK bandwidth, and find that the estimated effects are 2.1 and 3.6 percentage points, respectively. Thus, the estimates are quantitatively similar and relatively stable across different bandwidths and with the inclusion of fixed effects. Using estimates of elasticity of GDP to night-time lights, a three percentage point difference in luminosity growth translates into a 0.3 to 0.45 percentage point difference in GDP growth, based on the elasticity used.[24] The average growth in India during the period of study was about seven per cent per year. Our estimates hence indicate that the growth premium for constituencies stemming from them having a graduate legislator could be between 4 and 6 per cent.

We also illustrate the impact graphically in Figure 4. Each dot in the figure reflects the average annual growth rate in night-time lights in successive intervals of 0.5 per cent of the margin of victory. In the plot, a positive margin of victory suggests that a candidate with a graduate degree has been elected over a candidate without a graduate degree; the reverse would be true when the margin of victory is negative. The figure shows a discontinuous jump in growth rate of night-time lights at the cut-off (i.e. at $Margin_{idst} = 0$). The magnitude of the jump also corresponds to the effect size found in our main results presented in Table 2. This suggests that constituencies that elect a graduate leader in close

[23] We also present OLS results for Equation (2) in Table A6 of the Appendix.

[24] The potential impact on GDP growth should be interpreted as indicative at best. The elasticity has been estimated to be around 0.10 using district-year level panel data for a few years by Bickenbach et al. (2016) and around 0.15 using state-year level panel data by Baskaran et al. (2021); however, these estimates can vary when we consider smaller regions like constituencies in our case.

elections experience higher growth in night-time lights, implying a higher level of economic activity, than constituencies that elect non-graduate leaders in close elections.

6.2 Alternate thresholds for leaders' education

In this section, we consider completion of schooling as an alternate threshold to differentiate between educated and less educated leaders. Using the RDD, we establish that a graduate degree is the threshold of education that is most relevant when we consider the impact of leaders' education on night lights. In other words, we find that leaders who have a graduate degree drive most of the impact of educated leaders and not leaders who have completed schooling but do not have a graduate degree.

First, we consider a more educated candidate to be one who has completed schooling and a less educated candidate to have not completed schooling.[25] Using this cut-off, we estimate the effect on growth of night-time lights in Table 3. We find that when candidates who have completed schooling are elected, the growth rate of night-time lights is 3.1 percentage points higher than when a candidate who has not completed schooling is elected. These estimates are statistically significant for all bandwidths considered $(h, 2h, h/2, CCT)$. In terms of magnitude, the effect of school-educated leaders is similar to the effect of graduate leaders found in Table 2.

Next, we show that the observed effect from using school completion as a cut-off for education in Table 3 is driven by the sub-sample of candidates who not only have completed school education, but also have a graduate degree. We show this in three steps that involve pairwise comparison of three types of leaders: those who have not completed schooling (years of education < 12), those who have completed schooling but not attained a graduate degree (years of education $\in [12, 15)$), and graduate leaders (years of education ≥ 15). Firstly, we restrict our sample to elections contested by candidates who have a graduate degree and candidates who have not completed schooling (excluding those who have completed schooling but not attained a graduate degree). The point estimates presented in Panel A of Table 4 suggest that the election of a graduate candidate over a candidate who has not completed schooling results in a 3.9 percentage point increase in annual average growth rate of night-time lights. The estimates for all bandwidths are statistically significant and quantitatively similar to the estimates in analysis of graduate versus non-graduate leaders in Table 2. Secondly, in Panel B of Table 4, we restrict our sample to elections contested by candidates who have a graduate degree and candidates who have completed school education but not attained a graduate degree. Similarly, we find that the effect of election of a graduate candidate over a candidate with only school education is statistically significant and quantitatively similar to the estimates of having a graduate leader in Table 2. Finally, we restrict our sample to elections contested by candidates who have completed schooling but are not graduates and candidates who have not completed schooling. The estimates in Panel C of Table 4 are statistically insignificant and far smaller in magnitude. Thus, in comparison with candidates who have not completed schooling, candidates who have completed schooling but do not hold a graduate degree has no impact on growth of night-time lights. Findings from the three panels, put together, indicate that it is the candidates' attainment of a graduate degree rather than school education that has an impact on the growth of night-time lights in constituencies.

---

[25] In the Indian context, schooling consists of 12 years of education, which we consider to be the threshold in this case. Nearly 75 per cent of the top two candidates contesting in elections have at least 12 years of education. The RDD is based on the sample of elections where one of the top two candidates has completed schooling (years of education ≥ 12) while the other has not (years of education < 12). Apart from school completion and graduate degrees, we do not consider other thresholds for the leaders' education because other thresholds would make the distribution of sample between more educated and less educated leaders very unbalanced.

7 Heterogeneity analysis

In this section, we explore if the effect of a graduate leader on growth rate of night-time lights varies by the development level of states and other characteristics of the leader.

7.1 Heterogeneity based on the level of development of states

The pre-existing level of development of a state may moderate the effectiveness of educated leaders. On one hand, it may be easier for educated leaders to improve economic outcomes if the base level of development is low. On the other hand, more developed states may provide the required infrastructure to support educated leaders, resulting in greater impact. To explore this angle, we divide our sample into three groups of states based on Human Development Index (HDI) scores.[26] We estimate the RD model separately for each sub-sample and present the estimates in Table 5. We find that the effect is larger in the states with the lowest HDI, where electing a graduate leader results in a 5.7 percentage point increase in growth rate of night-time lights. The effect is 1.9 percentage points in the highest HDI states, and the difference in the effect sizes between the highest and the lowest HDI states is statistically significant. Thus, underdeveloped states seem to benefit more from having an educated leader.

7.2 Heterogeneity by candidate characteristics

We examine whether the effect of graduate leaders varies across candidate characteristics. Figure 5 plots heterogeneity by leader's gender, age, party alignment with state ruling party, and criminality. The point estimates are reported in Tables A7–A10.

We find that for the sub-sample consisting of male leaders, the effect of graduate leaders is a 3.4 percentage point higher growth rate of night-time lights, but the effect is not significant among female leaders. However, this result may be driven by the fact that the sample size for the sub-sample of female leaders is much lower than that of male leaders.

Next, we hypothesize that formal education may play a larger role in leadership skills of younger leaders, while it may be less relevant for older leaders who are likely to be more experienced. To examine this possibility, we categorize leaders into older and younger groups based on their age. Indeed, whether the leader has a graduate degree matters only for the sub-sample of younger leaders for whom the effect is precisely estimated. Although the effect of graduate leaders is slightly lower and imprecise among older leaders, it is statistically not distinguishable from the effect among younger leaders.

If the politicians' party is the same as the ruling party in the state, it might help them in gaining access to public resources and promote growth in their constituencies. So, we estimate the impact on growth in night-time lights for the aligned and not-aligned samples of legislators separately. We do not find any substantial difference between these two types of leaders.

Finally, investigating heterogeneity by criminal record of the legislator, we find no substantial difference in the effect of graduate leaders between the sub-samples of criminally accused and non-accused legislators.

[26] We classify Bihar, Chhattisgarh, Jharkhand, Madhya Pradesh, Arunachal Pradesh, Assam, Meghalaya, Odisha, Rajasthan, and Uttar Pradesh as least developed states; Andhra Pradesh, Delhi, Gujarat, Himachal Pradesh, Jammu and Kashmir, Karnataka, Mizoram, Manipur, Nagaland, Puducherry, Sikkim, Tripura, and West Bengal to have a higher development than least developed states. Lastly, we classify Haryana, Uttarakhand, Maharashtra, Punjab, Tamil Nadu, Kerala, and Goa as the most developed states.

8 Mechanisms

We delve into exploring various pathways through which a graduate leader may affect economic activities in the constituency, resulting in higher growth rate of night-time lights, as found in our main analysis.

8.1 Effect on provision of public goods

First, we consider the effect of graduate leaders on the public goods index described in the first subsection of Section 3.4. The point estimates are presented in Table 6. We find that there is no significant effect of electing a graduate leader on overall provision of public goods in the constituency.

8.2 Effect on provision of roads, electricity, and power

We postulate that graduate leaders might be able to influence the provision of certain public goods as opposed to all, as some of these can be built easily and in a shorter time span. Hence, we consider separate indices for roads, electricity, and power supply; these variables are likely to be captured as a part of night-time lights in satellite images as well (Henderson et al. 2012). The point estimates for these outcomes are presented in Table 7. With regard to roads, we find that the election of graduate leaders results in an increase in accessibility to roads by 0.35 standard deviation. Similarly, for power supply, we find that there is a statistically significant increase of 0.35 standard deviation in access. Estimates also suggest that the election of a graduate leader over a non-graduate leader results in a 5.37 percentage point increase in the proportion of households having electricity as the main source of lighting in the constituency.

8.3 Effect on crime

Another mechanism through which graduate leaders might impact economic growth is through a reduction in the level of crimes in their constituency. To test for this channel, we estimate the impact of electing a graduate leader on aggregate crimes recorded. Since we do not have data for crimes being reported at the constituency level, we use district-level annual panel data on the number of crimes recorded by the police. To carry out the estimation at the district level, we instrument the fraction of seats won by a graduate politician by the fraction of seats won by graduates in close elections. As additional control variables, the regression includes the district-level proportion of close elections between graduate and non-graduate candidates, the vote margin between such candidates, district fixed effects, and year fixed effects.[27] We find that a ten percentage point increase in the proportion of graduate legislators in a district leads to a 1.2 to 1.9 per cent reduction in average incidence of crimes in the district (Table 8).

8.4 Effect on corruption

Besides crime, another potential mechanism through which a graduate leader might promote economic growth is through reduction in corruption. Leaders holding office may benefit from rent-seeking activities when they hold the political office, thus indicating higher financial returns and corruption during their tenure. A general reduction in corruption might lead to better business environment, and more competition and efficiency in implementing public works in the constituency. Following Fisman et al.

[27] This empirical strategy aggregates the constituency-specific discontinuities at the district level to construct the instrumental variable. This method has been extensively used in the literature identifying the socio-economic effect of political leaders' identity using district-level data (Bhalotra and Clots-Figueras 2014; Clots-Figueras 2012; Lahoti and Sahoo 2020; Prakash et al. 2022).

(2014), we proxy for corruption by measuring change in net asset growth of candidates between two subsequent elections. Using the RDD, we find no statistically significant difference in net asset growth of graduate versus non-graduate leaders who won in close elections (Table 9).[28]

9 Sensitivity analysis

In this section, we conduct a range of robustness and sensitivity tests, such as controlling for higher order polynomials of the forcing variable, excluding extreme values of the dependent variable, and placebo tests. We find that our main results remain unperturbed in terms of both magnitude and significance.

9.1 Sensitivity to higher order polynomial

We explore the sensitivity of the RD estimates to controlling for quadratic and cubic functions of vote margin (i.e. the forcing variable). We report these results for different bandwidths and fixed effects in Table A11. Across these specifications, we find that the estimates are positive, mostly significant, and quantitatively similar to the effects estimated in Table 2, where linear function of vote margin was controlled for.

9.2 Extreme values

As discussed in Section 3.1, night-time lights data captured through satellite imagery are top coded at 63, i.e. no pixel value in an image exceeds this number. While this is not a problem in less developed areas, this could be an issue in constituencies that have areas with high luminosity, where we cannot measure any changes in intensity of night-time lights beyond 63. Also in some dimly lit areas, satellites might not be able to capture dim lights and the measure could be coded as zero. For about 4.95 per cent observations in our data, constituency-level mean night-time luminosity has a value of either 0 or 63.

As a part of robustness check, we address this issue by excluding extreme values (both 0 and 63) and present the estimates from the RD model in Table A13. In Columns (1)–(4), we drop all observations where the constituency-year pixel mean take the value of 0 or 63. In Columns (5)–(8), we drop any constituency in which mean pixel intensity takes the value of 0 or 63 for any of the years in the period 2009–13. We find that the results are unaffected by exclusion of extreme values, and the estimates remain quantitatively similar and statistically significant.

9.3 Placebo tests

To evaluate the RD design, we conduct two placebo tests and find that the placebo coefficient estimates in both cases are insignificant and smaller in size than the true effect. In the first test, we estimate Equation (3) with 22 placebo thresholds of the forcing variable. Specifically, we estimate 11 coefficients on a sub-sample of graduate leaders; here, instead of defining the threshold of discontinuity when margin of victory is zero, we redefine the cut-off in steps of 0.5 within the interval $[5, 10]$. Similarly, taking the sub-sample of non-graduate leaders, we redefine the cut-off in steps of 0.5 within the interval $[-10, -5]$ on the margin of victory variable. We plot these estimates based on the 22 placebo cut-offs and contrast them with the true coefficient of graduate leaders in Figure A1. We find that the placebo estimates are insignificant and smaller in size than the true estimate of the causal effect of graduate leaders.

---

[28] Further, considering alternate definitions of the dependent variable such as total change in net asset between two elections and growth rate of total assets of a leader, we find no significant effect in either.

In the second placebo test, we consider the dependent variable to be defined as the growth rate of night-time lights in the previous election term. Thus, we estimate Equation (3) for the effect of electing a graduate leader in the current election on the economic outcomes in the previous election term. We find that the estimates are statistically insignificant and much smaller in magnitude than the true estimate (Table A14).

9.4 Candidate characteristics

If a leader's education is systematically correlated with other characteristics, we may not be able to separate out the causal effect of education from those characteristics. While Figure 3 does not show any major difference in individual characteristics between graduate and non-graduate legislators winning in close elections, Table A4 indicates that they may slightly differ in terms of gender and criminality. In order to check if some of the impact on night-time lights of having a graduate candidate win in a narrow election is driven by these characteristics, we control for them in our model in Table A15. The effect of graduate leader remains almost unchanged in terms of both magnitude and statistical significance. This alleviates the concern that the impact we observe may be due to the legislators' education bundled with other characteristics.

10 Generalizability of the RD estimates

Our broad results from the RD analysis show that economic growth is significantly affected by electing educated legislators as compared to less educated legislators in narrow elections. But do our results generalize and hold across all elections? Figure 4 indicates a slightly higher growth rate of night-time lights associated with having a graduate leader across various margins of victory. Our OLS results for all (close and non-close) elections between graduate and non-graduate candidates also show positive and significant association. However, this correlation is lower in magnitude than RD estimates and loses significance when we include district fixed effects. These OLS results are also potentially contaminated by endogenous selection of graduate leaders. Since our identification strategy relies on close elections, we now investigate how representative our close election sample is to all elections in India.

We find that close elections between graduate and non-graduate candidates are largely representative of all mixed elections (i.e. all elections between graduate and non-graduate candidates) in India. Overall, we find that 47 per cent of all mixed elections are close elections. The median margin of mixed elections is 7.9 and our RD estimates use a bandwidth between 3.7 and 14.85. Close elections are also geographically spread across the country. The percentage of close elections in various states varies from 28 to 75 per cent. Further, we compare a range of constituency and candidate characteristics between mixed and non-mixed elections (Table A16) and between close and non-close elections (Table A17). In both cases, the different types of elections look comparable on the basis of most of the observable factors.

Another concern with our RD analysis is that a graduate (or a non-graduate) candidate contesting in a close election may be different in unobservable characteristics from an average graduate (or non-graduate) politician.[29] For instance, it is possible that voters prefer competent leaders and they infer a politician's competence by observing their education level. It is also likely that voters would perceive graduate candidates to be more competent than non-graduate candidates. Then, a graduate candidate must possess a lower level of competence in order to be involved in a close contest with a non-graduate candidate. Similarly, since graduates are on average preferred by voters, non-graduate candidates need

[29] Following a related argument of Marshall (2022), it is possible that our RDD identifies the effect of a leader's education and all compensating differentials—candidate-level characteristics that ensure elections remain close between graduate and non-graduate candidates.

to be especially competent to overcome this disadvantage and be in a close election against graduate candidates. While this is a plausible scenario, it is likely to result in underestimation of the effect of graduate leaders when the empirical analysis is based on close elections; thus, our estimates might be a lower bound of the true effect of graduate leaders on growth of night-time lights.

11 Conclusion

In the political economy literature, emphasis is placed on the importance of leader's education as it is assumed to instill more valuable skills in leaders, resulting in better performance. Recent academic as well as policy discourses have considered formal education of leader as a proxy for leader's quality, suggesting that education sets leaders apart. Using a detailed dataset from India and a regression discontinuity design, we find a positive but small effect on growth in night-time lights due to the election of graduate leaders to state assembly constituencies. The estimated effect size of leaders' education, however, is smaller in comparison with previous studies that have estimated the effect of a leaders' criminality (Prakash et al. 2019) and gender (Baskaran et al. 2021) on growth of night-time lights in the Indian context. Analysing the various thresholds of education level to define an educated leader, we show that the positive effects are due to those leaders who have completed college education, whereas we do not find any effect of leaders who have completed only school education. We also provide evidence for heterogeneity in our results, especially showing that the effect of educated leaders is stronger in states with lower levels of development.

We also explore the pathways through which graduate leaders affect development. While there is no change in the overall provision of public goods, we do find that graduate leaders have a positive and significant effect on the availability of road infrastructure and electricity supply in the constituency. Graduate leaders also reduce crime rates that may in turn promote an environment conducive for economic development. Further, we show that the estimated effect of graduate leaders is robust to various specification tests.

There are some limitations to the study due to data availability and the nature of the analysis that may be relevant for interpreting the results. Our data spans the period from 2009 to 2013 due to limits imposed by the availability of information on candidates' characteristics and comparable night-time lights. The data on public goods, roads, electricity, and power are cross-sectional in nature, limiting the extent to which we can draw conclusions based on these data. We show that close elections are largely similar to non-close elections on observable characteristics, supporting the external validity of our findings. However, our empirical strategy cannot fully alleviate the concern that unobservable characteristics may vary across close and non-close elections. Addressing these issues constitutes an agenda for future research.

References

Afzal, M. (2014). 'Do Barriers to Candidacy Reduce Political Competition? Evidence from a Bachelor's Degree Requirement for Legislators in Pakistan'. *Public Choice*, 161(1/2): 51–72. https://doi.org/10.1007/s11127-013-0126-2

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). 'Ethnic Inequality'. *Journal of Political Economy*, 124(2): 428–88. https://doi.org/10.1086/685300

Arora, A. (2022). 'Election by Community Consensus: Effects on Political Selection and Governance'. *The Review of Economics and Statistics*, 104(2): 321–35. https://doi.org/10.1162/rest_a_00970

Asher, S., and Novosad, P. (2017). 'Politics and Local Economic Growth: Evidence from India'. *American Economic Journal: Applied Economics*, 9(1): 229–73. https://doi.org/10.1257/app.20150512

Baskaran, T., Bhalotra, S., Min, B., and Uppal, Y. (2021). 'Women Legislators and Economic Performance'.

CEPR Discussion Paper DP16605. London: Centre for Economic Policy Research.

Bastos, P., and Sánchez, C. (2021). 'The Effects of Educated Leaders on Policy and Politics: Quasi-Experimental Evidence from Brazil'. Mimeo.

Besley, T. (2005). 'Political Selection'. *Journal of Economic Perspectives*, 19(3): 43–60. https://doi.org/10.1257/089533005774357761

Besley, T., and Coate, S. (1997). 'An Economic Model of Representative Democracy'. *The Quarterly Journal of Economics*, 112(1): 85–114. https://doi.org/10.1162/003355397555136

Besley, T., Montalvo, J. G., and Reynal-Querol, M. (2011). 'Do Educated Leaders Matter?' *The Economic Journal*, 121(554): F205–21. https://doi.org/10.1111/j.1468-0297.2011.02448.x

Besley, T., Pande, R., Rahman, L., and Vijayendra, R. (2004). 'The Politics of Public Good Provision: Evidence from Indian Local Governments'. *Journal of European Economic Association*, 2(2-3): 416–26. https://doi.org/10.1162/154247604323068104

Bhalotra, S., and Clots-Figueras, I. (2014). 'Health and the Political Agency of Women'. *American Economic Journal: Economic Policy*, 6(2): 164–97. https://doi.org/10.1257/pol.6.2.164

Bhalotra, S., Clots-Figueras, I., Cassan, G., and Iyer, L. (2014). 'Religion, Politician Identity and Development Outcomes: Evidence from India'. *Journal of Economic Behavior & Organization*, 104(-): 4–17. https://doi.org/10.1016/j.jebo.2013.09.00

Bhalotra, S., Clots-Figueras, I., and Iyer, L. (2018). 'Pathbreakers? Women's Electoral Success and Future Political Participation'. *The Economic Journal*, 128(613): 1844–78. https://doi.org/10.1111/ecoj.12492

Bhandari, L., and Roychowdhury, K. (2011). 'Night Lights and Economic Activity in India: A Study Using DMSP-OLS Night Time Images'. *Proceedings of the Asia-Pacific Advanced Network*, 32(-): 218–36. https://doi.org/10.7125/APAN.32.24

Bhaskar, A. (2016). 'Damage to Democracy: Elitist Judgment of Supreme Court in *Rajbala v State of Haryana*'. *Economic & Political Weekly*, 51(40): 47–54.

Bickenbach, F., Bode, E., Nunnenkamp, P., and Söder, M. (2016). 'Night Lights and Regional GDP'. *Review of World Economics*, 152(2): 425–47. https://doi.org/10.1007/s10290-016-0246-0

Bruederle, A., and Hodler, R. (2018). 'Nighttime Lights as a Proxy for Human Development at the Local Level'. *PLoS ONE*, 13(9): e0202231. https://doi.org/10.1371/journal.pone.0202231

Burchi, F. (2013). 'Women's Political Role and Poverty in Educational Dimension: A District-Level Analysis in India'. Discussion Paper 23/2013. Bonn: German Development Institute. https://doi.org/10.2139/ssrn.2369248

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). 'Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs'. *Econometrica*, 82(6): 2295–326. https://doi.org/10.3982/ECTA11757

Carnes, N., and Lupu, N. (2016). 'What Good Is a College Degree? Education and Leader Quality Reconsidered'. *The Journal of Politics*, 78(1): 35–49. https://doi.org/10.1086/683027

Chattopadhyay, R., and Duflo, E. (2004). 'Women as Policy Makers: Evidence from a Randomized Policy Experiment in India'. *Econometrica*, 72(5): 1409–43. https://doi.org/10.1111/j.1468-0262.2004.00539.x

Chen, X., and Nordhaus, W. D. (2011). 'Using Luminosity Data as a Proxy for Economic Statistics'. *Proceedings of the National Academy of Sciences*, 108(21): 8589–94. https://doi.org/10.1073/pnas.1017031108

Clots-Figueras, I. (2011). 'Women in Politics: Evidence from the Indian States'. *Journal of Public Economics*, 95(7-8): 664–90. https://doi.org/10.1016/j.jpubeco.2010.11.017

Clots-Figueras, I. (2012). 'Are Female Leaders Good for Education? Evidence from India'. *American Economic Journal: Applied Economics*, 4(1): 212–44. https://doi.org/10.1257/app.4.1.212

Congleton, R. D., and Zhang, Y. (2013). 'Is it all about Competence? The Human Capital of U.S. Presidents and Economic Performance'. *Constitutional Political Economy*, 24(2): 108–24. https://doi.org/10.1007/s10602-013-9138-7

Curto-Grau, M., and Gallego, A. (2019). 'The Education of Politicians: Effects on Performance and Fiscal Policy'. Working Paper.

Diaz-Serrano, L., and Pérez, J. (2013). 'Do More Educated Leaders Raise Citizen's Education?'. IZA DP 7661. Bonn: Institute of Labor Economics. https://doi.org/10.2139/ssrn.2342533

Doll, C. N., Muller, J.-P., and Morley, J. G. (2006). 'Mapping Regional Economic Activity from Night-Time Light Satellite Imagery'. *Ecological Economics*, 57(1): 75–92. https://doi.org/10.1016/j.ecolecon.2005.03.007

Donaldson, D., and Storeygard, A. (2016). 'The View from Above: Applications of Satellite Data in Economics'. *The Journal of Economic Perspectives*, 30(4): 171–98. https://doi.org/10.1257/jep.30.4.171

Downs, A. (1957). 'An Economic Theory of Political Action in a Democracy'. *Journal of Political Economy*, 65(2): 135–50. https://doi.org/10.1086/257897

Dreher, A., Lamla, M. J., Lein, S. M., and Somogyi, F. (2009). 'The Impact of Political Leaders' Profession and Education on Reforms'. *Journal of Comparative Economics*, 37(1): 169–93. https://doi.org/10.1016/j.jce.2008.08.005

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and 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): 259–74. https://doi.org/10.1111/ajps.12127

Ejaz Ghani, S., Mani, A., and D. O'Conell, S. (2013). 'Can Political Empowerment Help Economic Empowerment? Women Leaders and Female Labor Force Participation in India'. Policy Research Working Paper 6675. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-6675

Fisman, R., Schulz, F., and Vig, V. (2014). 'The Private Returns to Public Office'. *Journal of Political Economy*, 122(4): 806–62. https://doi.org/10.1086/676334

François, A., Panel, S., and Weill, L. (2020). 'Educated Dictators Attract more Foreign Direct Investment'. *Journal of Comparative Economics*, 48(1): 37–55. https://doi.org/10.1016/j.jce.2019.11.006

Gelman, A., and Imbens, G. (2019). 'Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs'. *Journal of Business & Economic Statistics*, 37(3): 447–56. https://doi.org/10.1080/07350015.2017.1366909

Gulzar, S., and Pasquale, B. J. (2017). 'Politicians, Bureaucrats, and Development: Evidence from India'. *American Political Science Review*, 111(1): 162–83. https://doi.org/10.1017/S0003055416000502

Halim, N., M. Yount, K., A. Cunningham, S., and P. Pande, R. (2016). 'Women's Political Empowerment and Investments in Primary Schooling in India'. *Social Indicators Research*, 125(-): 813–51. https://doi.org/10.1007/s11205-015-0870-4

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). 'Measuring Economic Growth from Outer Space'. *The American Economic Review*, 102(2): 994–1028. https://doi.org/10.1257/aer.102.2.994

Hill, A. J., and Jones, D. B. (2017). 'Does Partisan Affiliation Impact the Distribution of Spending? Evidence from State Government's Expenditures on Education'. *Journal of Economic Behavior and Organization*, 143(-): 58–77. https://doi.org/10.1016/j.jebo.2017.09.008

Hodler, R., and Raschky, P. A. (2014). 'Regional Favoritism'. *The Quarterly Journal of Economics*, 129(2): 995–1033. https://doi.org/10.1093/qje/qju004

Imbens, G., and Kalyanaraman, K. (2012). 'Optimal Bandwidth Choice for the Regression Discontinuity Estimator'. *The Review of Economic Studies*, 79(3): 933–59. https://doi.org/10.1093/restud/rdr043

Imbens, G., and Lemieux, T. (2008). 'Regression Discontinuity Designs: A Guide to Practice'. *Journal of Econometrics*, 142(2): 615–35. https://doi.org/10.1016/j.jeconom.2007.05.001

Iyer, L., Mani, A., Mishra, P., and Topalova, P. (2012). 'The power of political voice: Women's political representation and crime in india'. *American Economic Journal: Applied Economics*, 4(4): 165–193.

Iyer, L., and Reddy, M. (2013). 'Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World's Largest Democracy?'. HBS Working Paper 14-051. Boston: Harvard Business School.

Lahoti, R., and Sahoo, S. (2020). 'Are Educated Leaders Good for Education'. *Journal of Economic Behavior and Organization*, 176(-): 42–62. https://doi.org/10.1016/j.jebo.2020.03.026

Lee, D. S. (2008). 'Randomized Experiments from Non-Random Selection in US House Elections'. *Journal of Econometrics*, 142(2): 675–97. https://doi.org/10.1016/j.jeconom.2007.05.004

Marshall, J. (2022). 'Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?'. Working Paper.

Martinez-Bravo, M. (2017). 'The Local Political Economy Effects of School Construction in Indonesia'. *American Economic Journal: Applied Economics*, 9(2): 256–89. https://doi.org/10.1257/app.20150447

McCrary, J. (2008). 'Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test'. *Journal of Econometrics*, 142(2): 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

Michalopoulos, S., and Papaioannou, E. (2014). 'National Institutions and Subnational Development in Africa'. *The Quarterly Journal of Economics*, 129(1): 151–213. https://doi.org/10.1093/qje/qjt029

Mitra, A. (2020). 'Should You Want An Educated Mayor? Evidence from Close Elections in Italy'. Working Paper. https://doi.org/10.2139/ssrn.3616729

O'Connell, S. D. (2018). 'Political Inclusion and Educational Investment: Estimates from a National Policy Experiment in India'. *Journal of Development Economics*, 135(-): 478–87. https://doi.org/10.1016/j.jdeveco.2018.08.004

Osborne, M. J., and Slivinski, A. (1996). 'A Model of Political Competition with Citizen-Candidates'. *The Quarterly Journal of Economics*, 111(1): 65–96. https://doi.org/10.2307/2946658

Pande, R. (2003). 'Can Mandated Political Representation Increase Policy Influence for Disadvantages Minorities? Evidence for India'. *The American Economic Review*, 93(4): 1132–51. https://doi.org/10.1257/000282803769206232

Persson, T., and Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

Peveri, J. (2021). ‘The Wise, the Politician and the Strongman: National Leaders’ Type and Quality of Governance’. AMSE Working Paper 2021-20. Marseille: Aix-Marseille School of Economics, France.

Prakash, N., Rockmore, M., and Uppal, Y. (2019). ‘Do Criminally Accused Politicians Affect Economic Outcomes? Evidence from India’. *Journal of Development Economics*, 141(-): 102370. https://doi.org/10.1016/j.jdeveco.2019.102370

Prakash, N., Sahoo, S., Saraswat, D., and Sindhi, R. (2022). ‘When Criminality Begets Crime: The Role of Elected Politicians’. IZA Discussion Paper 15259. Bonn: Institute of Labor Economics. https://doi.org/10.2139/ssrn.4114835

Storeygard, A. (2016). ‘Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa’. *The Review of Economic Studies*, 83(3): 1263–95. https://doi.org/10.1093/restud/rdw020

Subramaniam, A. (2019). ‘India’s GDP Mis-estimation: Likelihood, Magnitudes, Mechanisms, and Implications’. CID Faculty Working Paper 354. Cambridge, MA: Center for International Development, Harvard University.

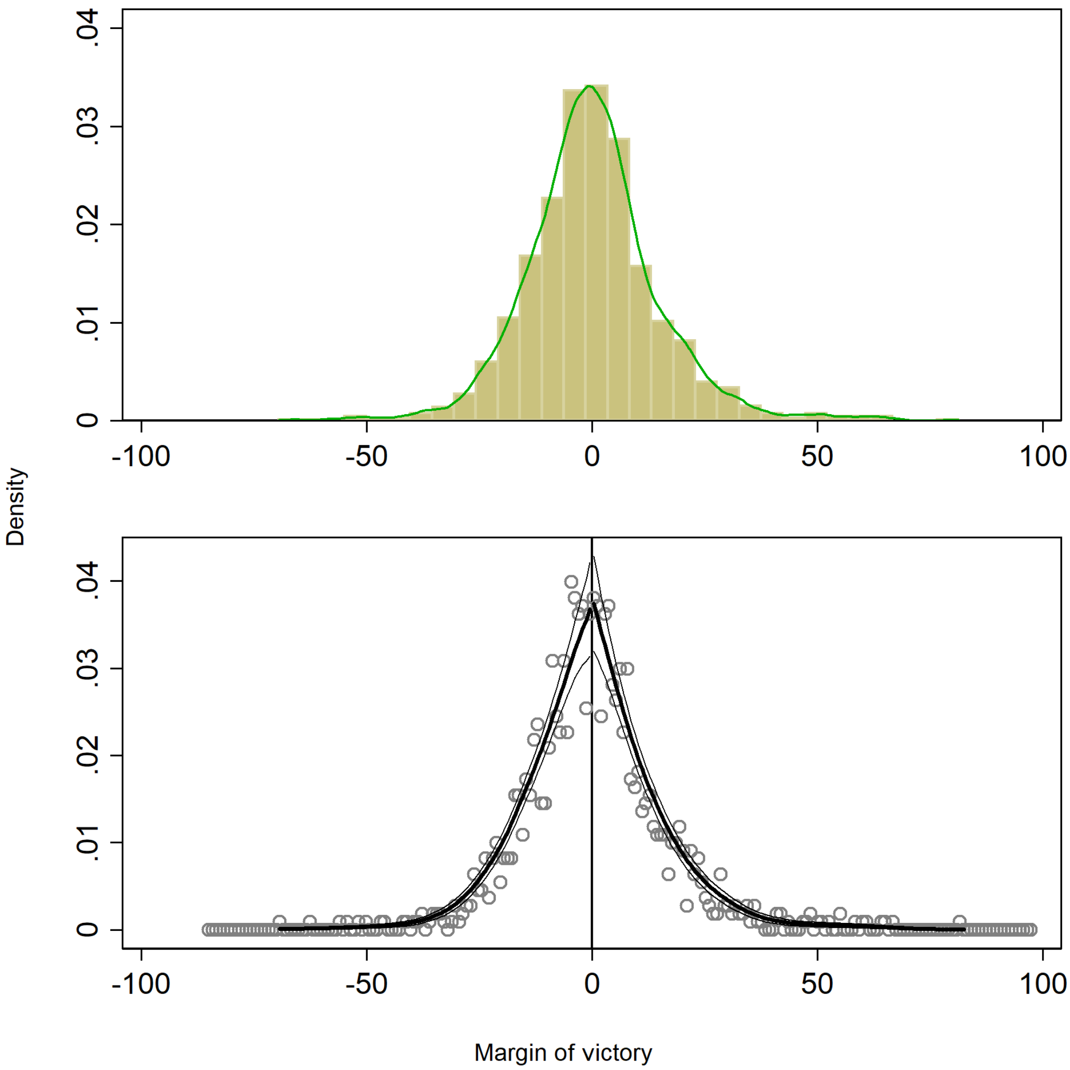
Susewind, R. (2014). *GIS Shapefiles for India’s Parliamentary and Assembly Constituencies Including Polling Booth Localities.* Bielefeld: Bielefeld University. https://doi.org/10.4119/unibi/2674065

Weidmann, N. B., and Schutte, S. (2017). ‘Using Night Light Emissions for the Prediction of Local Wealth’. *Journal of Peace Research*, 54(2): 125–40. https://doi.org/10.1177/0022343316630359

Yu, S., and Jong-A-Pin, R. (2020). ‘Rich or Alive? Political (In)Stability, Political Leader Selection and Economic Growth’. *Journal of Comparative Economics*, 48(3): 561–77. https://doi.org/10.1016/j.jce.2019.11.004

Figures and tables

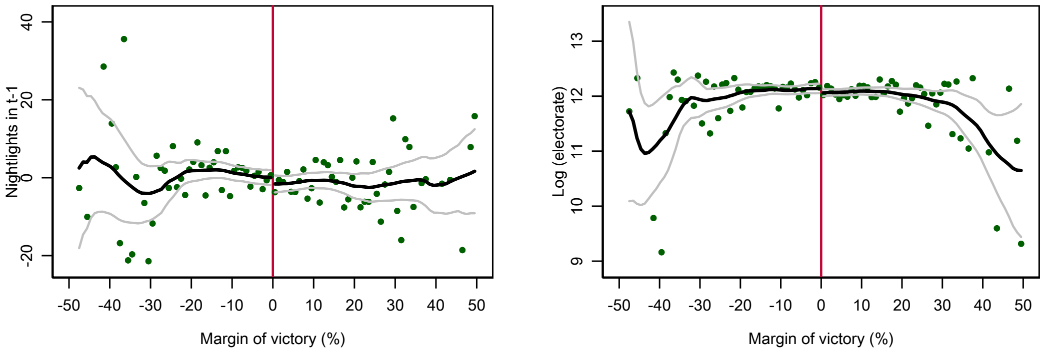
Figure 1: Density test of the forcing variable



Note: the figure shows the continuity of the forcing variable which is the margin of victory, defined as the difference in vote share between the graduate and the non-graduate candidates in mixed elections (i.e. elections where one of the top two candidates is graduate while the other is non-graduate). By construction, margin of victory is positive for graduate leaders and negative for non-graduate leaders. The upper panel plots the kernel density of victory of margin and the lower panel plots the density test for a discontinuity at the cut-off where margin of victory is 0. The figure shows that there is no significant discontinuity in the density of the victory margin above and below the cut-off.

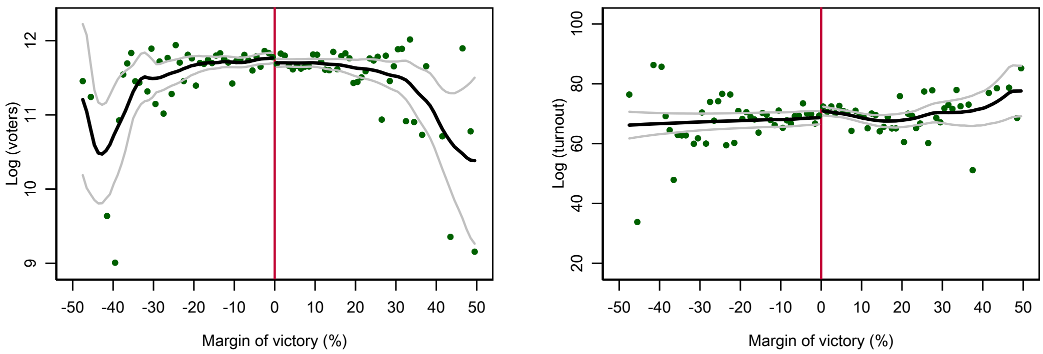
Source: authors' illustration.

Figure 2: Continuity checks of constituency-level covariates



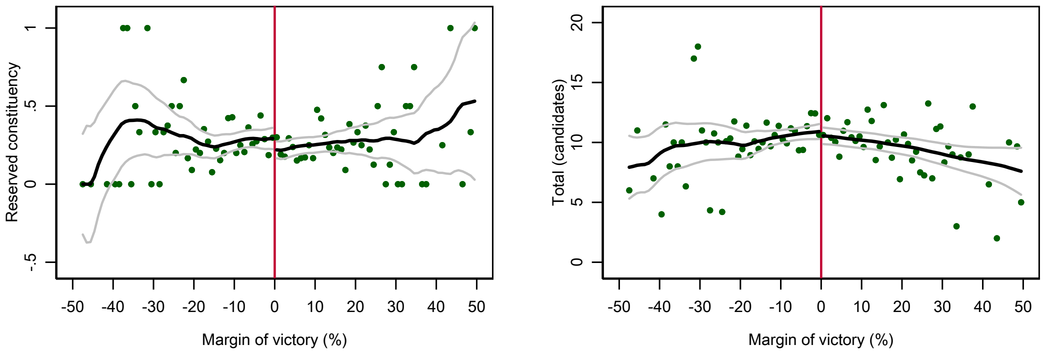
(a) Night-time light (previous term)

(d) Electorate size



(e) Voters

(f) Turnout



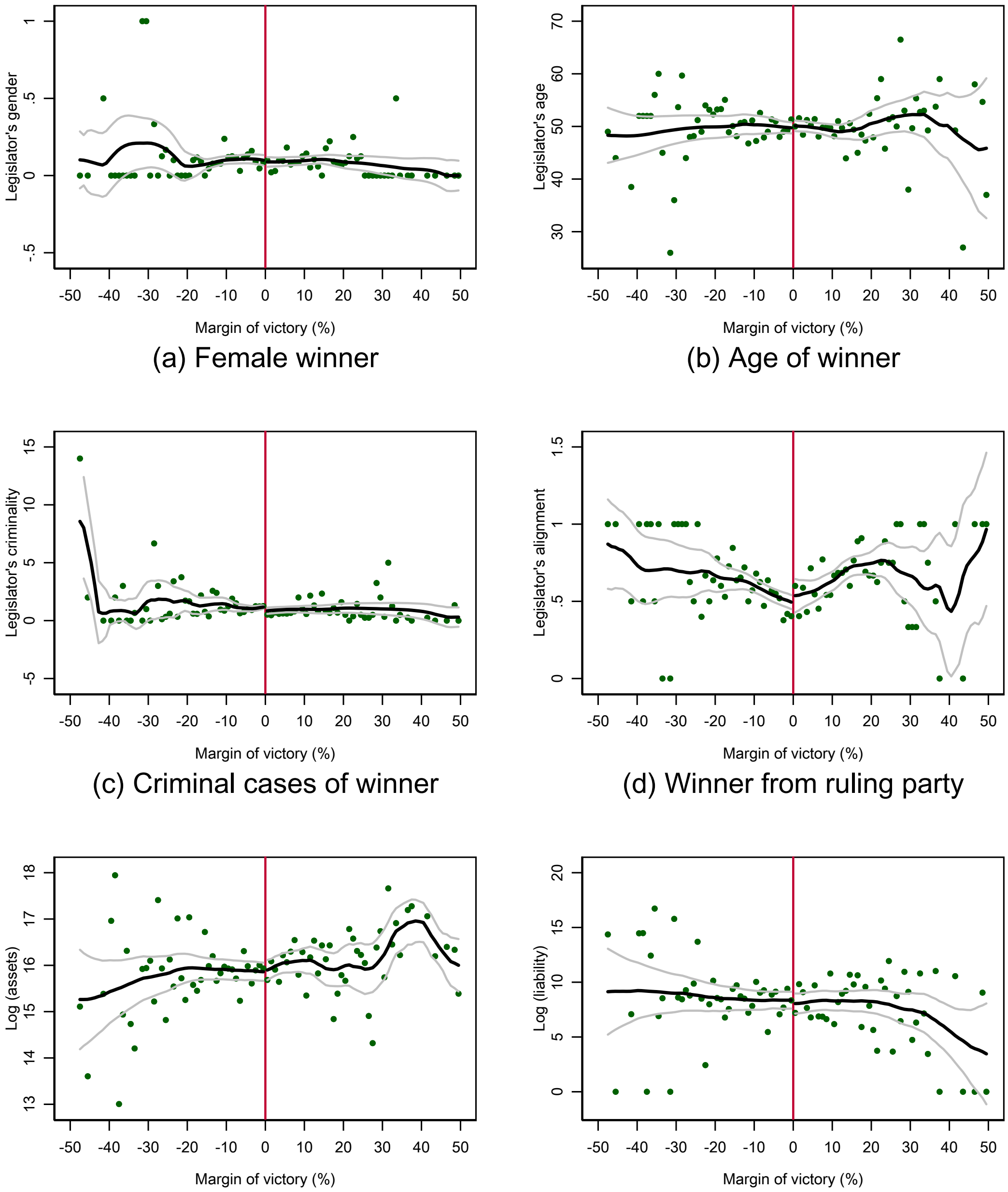
(b) SC/ST constituency

(c) Total number of candidates

Note: the figure plots the continuity checks for constituency characteristics. Each variable is plotted against margin of victory, which is the difference in vote share between the graduate and the non-graduate candidates in mixed elections. By construction, margin of victory is positive for graduate leaders and negative for non-graduate leaders. Each dot in the figure depicts the averages over successive intervals of 0.5% of margin of victory. The curves are local polynomial regressions (with 95% confidence intervals) fitted separately for positive and negative parts of the margin of victory variable.

Source: authors' illustration.

Figure 3: Continuity checks of candidate characteristics

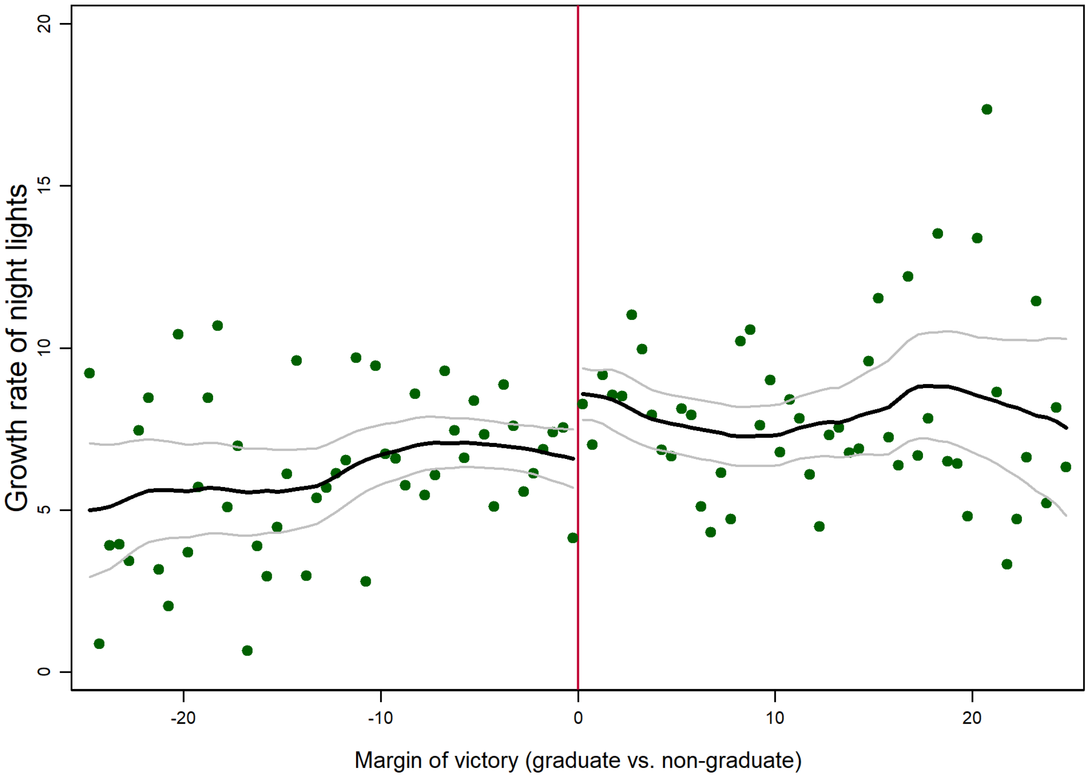


(a) Female winner
(b) Age of winner
(c) Criminal cases of winner
(d) Winner from ruling party
(e) Total assets of winner
(f) Total liabilities of winner

Note: the figure plots the continuity checks for legislator characteristics. Each variable is plotted against margin of victory, which is the difference in vote share between the graduate and the non-graduate candidates in mixed elections. By construction, margin of victory is positive for graduate leaders and negative for non-graduate leaders. Each dot in the figure depicts the averages over successive intervals of 0.5% of margin of victory. The curves are local polynomial regressions (with 95% confidence intervals) fitted separately for positive and negative parts of the margin of victory variable.

Source: authors' illustration.

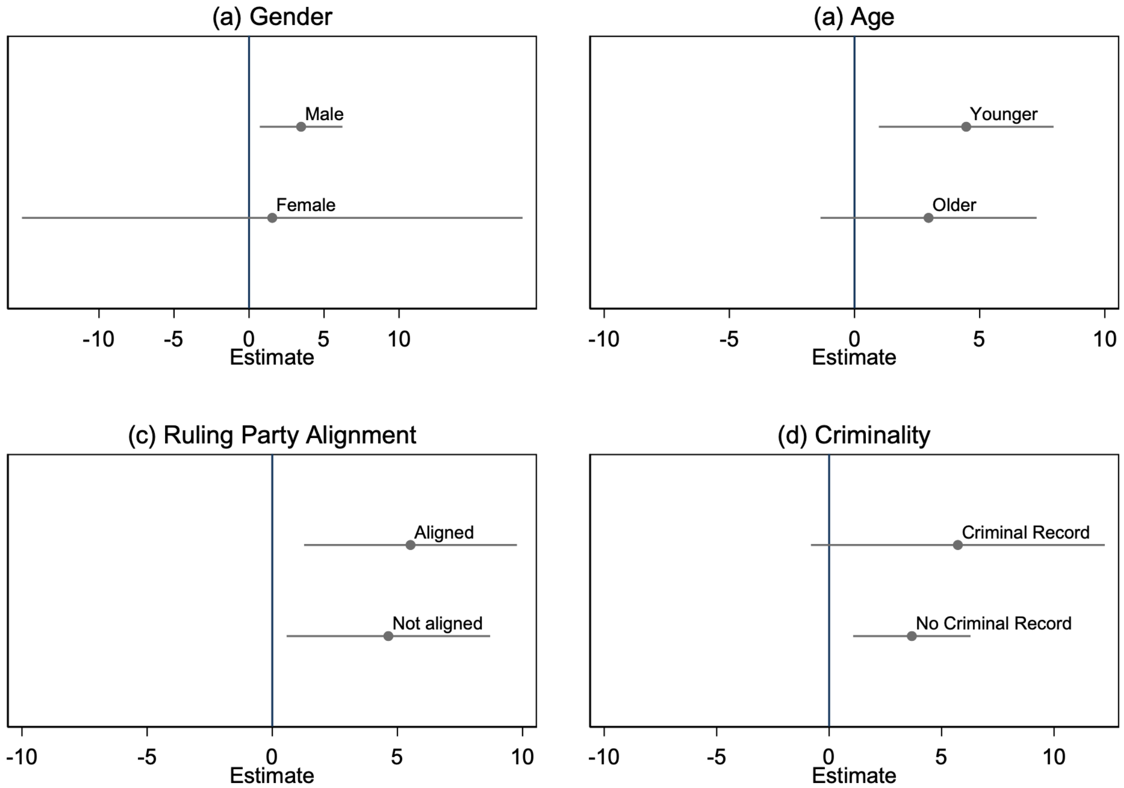
Figure 4: Leader's education and effect on growth in night-time lights



Note: the figure plots the dependent variable, annual growth rate of night-time lights, against margin of victory, which is the difference in vote share between the graduate and the non-graduate candidates in mixed elections. By construction, margin of victory is positive for graduate leaders and negative for non-graduate leaders. Each dot in the figure depicts the averages over successive intervals of 0.5% of margin of victory. The curves are local polynomial regressions (with 95% confidence intervals) fitted separately for positive and negative parts of the margin of victory variable.

Source: authors' illustration.

Figure 5: Heterogeneity in the effect of graduate leader on night-time lights by leader's gender, age, party alignment with state government, and criminal record



Note: the above figure plots the RD estimates of the effect of graduate leaders using IK(h) bandwidth on different sub-samples. We consider different sub-samples based on (a) gender, (b) age of leader, (c) party alignment of elected leader with state government, and (d) the criminal record of the leader. All regressions control for district and year fixed effects and RD estimates are calculated with a local linear regression using a triangular kernel.

Source: authors' illustration.

Table 1: Summary statistics

	Full sample			Mixed sample		
	N	Mean	SD	N	Mean	SD
Growth rate of night-time lights	13909	7.15	24.85	6081	7.13	24.63
Constituency characteristics						
Electorate size (log)	13909	11.88	0.79	6081	11.91	0.75
Voters (log)	13909	11.48	0.71	6081	11.50	0.68
Turnout	13903	68.35	13.80	6078	68.16	13.65
General constituency	13909	0.72	0.45	6081	0.72	0.45
SC constituency	13909	0.15	0.36	6081	0.15	0.35
ST constituency	13909	0.14	0.34	6081	0.14	0.34
Candidate characteristics						
Graduate leader	13909	0.60	0.49	6081	0.48	0.50
Graduate runner-up	13278	0.62	0.49	6081	0.52	0.50
Winner's assets (log)	13889	15.60	1.66	6064	15.61	1.63
Runner-up's assets (log)	13838	15.36	1.73	6065	15.41	1.72
Winner's liabilities (log)	13909	7.81	6.76	6081	7.77	6.72
Runner-up's liabilities (log)	13906	7.79	6.68	6081	8.11	6.65
Winner's criminality	13909	0.88	2.55	6081	0.89	2.48
Runner-up's criminality	13906	0.69	1.90	6081	0.72	1.89
Winner's age	13597	49.80	10.08	5948	49.60	10.28
Runner-up's age	13597	49.62	10.44	5948	49.49	10.50
Male winner	13909	0.92	0.27	6081	0.92	0.27
Male runner-up	13909	0.91	0.28	6081	0.91	0.29
Winner's education	13909	13.47	3.52	6081	12.83	3.57
Runner-up's education	13278	13.52	3.61	6081	12.95	3.61
Female winner	13909	0.08	0.27	6081	0.08	0.27

Note: full sample includes all constituencies and mixed sample includes constituencies where one of the top two candidates is a graduate while the other candidate is a non-graduate. The summary statistics on all variables are estimated using constituency-year level panel data spanning from 2009 to 2012, with data on nightlights spanning until 2013 to allow the calculation of growth rate for this period. Criminality measures the number of criminal cases against a candidate at the time of contesting the election.

Source: authors' elaboration based on data described in Section 3.

Table 2: Effect of electing a graduate leader on growth in night-time lights

	(1)	(2)	(3)	(4)
Panel A				
Graduate leader	3.012**	2.830**	2.203**	2.096
	(1.276)	(1.115)	(0.942)	(1.765)
Observations	2,883	3,589	4,517	1,520
R-squared	0.002	0.001	0.001	0.003
Bandwidth	7.427	10.16	14.85	3.714
Year fixed effects	No	No	No	No
State fixed effects	No	No	No	No
District fixed effects	No	No	No	No
Panel B				
Graduate leader	3.055***	2.872***	2.256***	2.554
	(1.149)	(1.004)	(0.822)	(1.697)
Observations	2,883	3,506	4,517	1,520
R-squared	0.505	0.497	0.488	0.507
Bandwidth	7.427	9.782	14.85	3.714
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No
Panel C				
Graduate leader	3.277***	3.137***	2.123**	3.612*
	(1.263)	(1.068)	(0.854)	(2.076)
Observations	2,876	3,499	4,502	1,516
R-squared	0.579	0.565	0.549	0.599
Bandwidth	7.427	9.782	14.85	3.714
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. Panel A, B, C estimate the model using different fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 3: Effect of electing a 12th pass (completed schooling) leader on growth in night-time lights

	(1)	(2)	(3)	(4)
12th pass leader	3.107**	3.295***	2.415***	4.227*
	(1.562)	(1.235)	(0.888)	(2.468)
Observations	2,168	2,588	3,358	1,216
R-squared	0.604	0.592	0.570	0.631
Bandwidth	7.807	10.19	15.61	3.903
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. Twelfth pass leader is a dummy variable that is 1 if a candidate who has completed schooling wins against a candidate who has not completed schooling and is 0 if the candidate who has not completed schooling wins against a candidate who has completed schooling. The model controls for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 4: Comparison effects of different levels of a leader's education on growth in night-time lights

	(1)	(2)	(3)	(4)
Panel A: Graduate (compared to 'not 12th pass')				
Graduate leader	3.941**	3.354**	2.627***	4.682*
	(1.774)	(1.379)	(1.011)	(2.609)
Observations	1,841	2,220	2,810	1,050
R-squared	0.608	0.591	0.571	0.643
Bandwidth	8.262	10.91	16.52	4.131
Panel B: Graduate (compared to '12th pass non-graduate')				
Graduate leader	3.971*	4.004**	2.891*	0.459
	(2.021)	(2.030)	(1.598)	(2.693)
Observations	1,370	1,351	2,038	728
R-squared	0.569	0.570	0.556	0.590
Bandwidth	9.093	8.971	18.19	4.547
Panel C: '12th pass non-graduate' (compared to 'not 12th pass')				
'12th pass non-graduate' leader	0.805	0.589	-0.361	4.134
	(3.356)	(3.416)	(2.257)	(7.894)
Observations	484	482	737	277
R-squared	0.630	0.630	0.596	0.626
Bandwidth	10.63	10.38	21.27	5.317
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In Panel A, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a candidate who has not completed school education (i.e., 'not 12th pass'), and 0 if the candidate who has not completed school education wins against a graduate candidate. In Panel B, 'graduate leader' dummy is 1 if a graduate candidate wins against a candidate who has completed school education but has not graduated from college (i.e., '12th pass non-graduate'), and 0 if the '12th pass non-graduate' wins against a graduate candidate. In Panel C, '12th pass non-graduate' leader is 1 if a candidate who has completed school education but is not a graduate wins against a candidate who has not completed school education, and 0 if a 'not 12th pass' candidate wins against a '12th pass non-graduate'. Models in all panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 5: Heterogeneous effect of a leader's education on growth in night-time lights by level of development

	(1)	(2)	(3)	(4)
Low HDI states				
Graduate leader	5.736**	5.505**	3.640**	3.839
	(2.661)	(2.488)	(1.724)	(3.911)
Observations	1,420	1,512	2,078	775
R-squared	0.517	0.513	0.487	0.550
Bandwidth	8.789	9.517	17.58	4.394
Medium HDI states				
Graduate leader	2.025	1.908	0.469	2.166
	(1.508)	(1.383)	(0.971)	(2.487)
Observations	1,028	1,111	1,513	597
R-squared	0.621	0.620	0.621	0.638
Bandwidth	8.812	10.26	17.62	4.406
High HDI states				
Graduate leader	1.959*	1.704*	1.404*	2.463
	(1.113)	(0.949)	(0.790)	(1.604)
Observations	792	983	1,307	428
R-squared	0.784	0.774	0.761	0.818
Bandwidth	8.623	11.42	17.25	4.311
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. Models in all panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 6: Leader's education and effect on provision of public goods

	(1)	(2)	(3)	(4)
Graduate leader	0.071	0.049	0.043	0.037
	(0.109)	(0.067)	(0.081)	(0.126)
Observations	297	698	539	156
R-squared	0.834	0.806	0.817	0.846
Bandwidth	3.577	10.47	7.154	1.788
Election-year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. Graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. The outcome variable is constructed by a principal component analysis of availability of various public goods at the constituency level, as described in the first sub-section of Section 3.4. The model controls for fixed effects at the levels of state and the year when the leader was elected. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 7: Leader's education and access to road, electricity, and power

	(1)	(2)	(3)	(4)
Roads				
Graduate leader	0.355***	0.222**	0.256**	0.528***
	(0.135)	(0.090)	(0.101)	(0.164)
Observations	332	720	605	169
R-squared	0.656	0.622	0.629	0.695
Bandwidth	3.643	9.493	7.286	1.821
Power				
Graduate leader	0.345***	0.190**	0.151*	0.422***
	(0.116)	(0.096)	(0.089)	(0.145)
Observations	353	550	624	179
R-squared	0.729	0.697	0.695	0.733
Bandwidth	3.997	6.669	7.993	1.998
Electricity				
Graduate leader	5.368**	5.316**	3.279*	5.497*
	(2.373)	(2.336)	(1.713)	(3.314)
Observations	649	663	998	370
R-squared	0.821	0.821	0.818	0.817
Bandwidth	8.175	8.477	16.35	4.088
Election-year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. The outcome variable is constructed using variables mentioned in the first sub-section of Section 3.4 by principal component analysis. Roads include state highways, national highways, paved roads, unpaved roads, all-weather roads, and other major roads. Power includes access to electricity for domestic, agricultural, and industrial use. Electricity measures the percentage of households in a constituency that have access to electricity as the main source of lighting. All panels control for fixed effects at the levels of state and the year when the leader was elected. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 8: Leader's education and effect on total crimes

	(1)	(2)	(3)	(4)
Fraction of seats won by graduates	-856.5*	-725.1**	-603.4**	-546.9**
	(494.0)	(343.1)	(261.5)	(243.4)
Observations	6396	6396	6396	6396
First Stage Fstat	179.2	289.1	427.3	540.8
Bandwidth	3	5	9	14
Mean of Total Crime	4439	4439	4439	4439
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes

Note: robust standard errors clustered at the district level are given in parentheses. The regressions are estimated on district level annual panel data for the period from 2005 to 2018. The outcome variable measures the district level total number of crimes recorded by the police, including any crime mentioned in the Indian Penal Code. The data was obtained from the National Crime Records Bureau. The district level fraction of seats won by a graduate politician is instrumented by the fraction of seats won by graduates in close elections. Additional controls include the proportion of close elections and linear function of the vote margin between graduate and non-graduate contestants. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table 9: Leader's education and effect on net asset growth

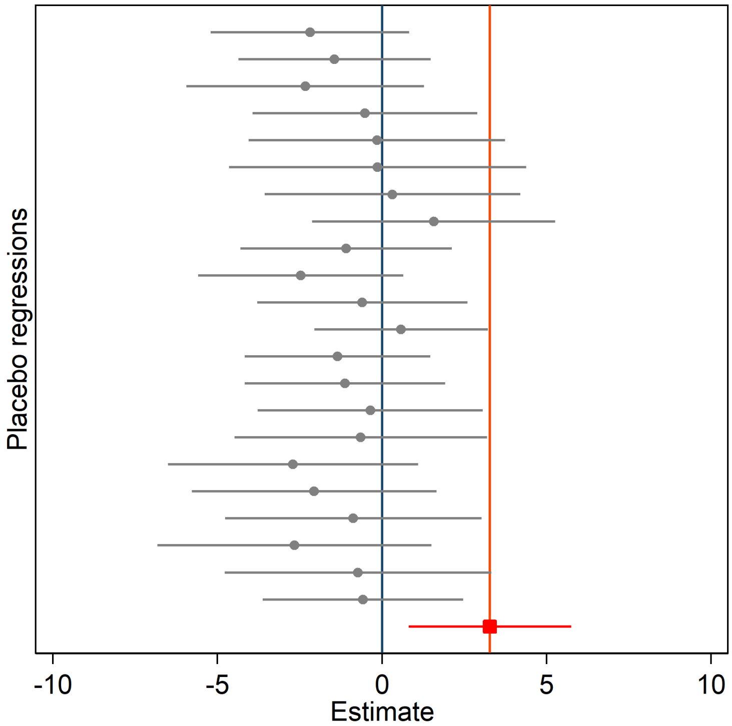
	(1)	(2)	(3)	(4)
Graduate leader	1.134	1.432	-0.466	1.847
	(1.940)	(1.951)	(1.421)	(2.622)
Observations	764	804	1,228	411
R-squared	0.049	0.040	0.026	0.070
Bandwidth	7.234	7.721	14.47	3.617
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. Graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. Net asset is defined as difference between assets and liabilities of leaders. Here, the main dependent variable is the growth in net asset for the duration that a leader holds office. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Appendix

Figure A1: Placebo regressions with artificial thresholds for estimating effect of graduate leader on night-time lights



Note: the figure displays 22 placebo coefficient estimates for the graduate leader dummy with 95% confidence intervals. We estimate 11 of these placebo coefficients on a sub-sample of graduate leaders by redefining the threshold of discontinuity in steps of 0.5 within the interval $[5, 10]$ on the margin of victory variable. Similarly, taking the sub-sample of non-graduate leaders, we redefine the cut-off in steps of 0.5 within the interval $[-10, -5]$ on the margin of victory variable. We plot the coefficients based on these placebo cut-offs, along with the true coefficient estimate from the main specification in Equation (3) that is highlighted in red. The plot shows that the placebo coefficients are not significantly different from zero and they are smaller than the true coefficient.

Source: authors' illustration.

Table A1: Summary statistics for public goods from census data

	Public goods					
	Full sample			Mixed sample		
	N	Mean	SD	N	Mean	SD
Constituency-level infrastructure						
Pre-primary school	2822	0.44	0.37	1196	0.44	0.36
Primary school	2822	0.89	0.14	1196	0.88	0.14
Middle school	2822	0.52	0.24	1196	0.52	0.24
Secondary school	2822	0.28	0.22	1196	0.26	0.21
Senior secondary school	2822	0.13	0.16	1196	0.13	0.15
College	2822	0.04	0.09	1196	0.03	0.07
Vocational training	2822	0.09	0.23	1196	0.08	0.22
Community health centre	2822	0.02	0.05	1196	0.02	0.05
Primary health centre	2822	0.08	0.13	1196	0.07	0.12
Primary health sub-centre	2822	0.28	0.24	1196	0.27	0.23
Dispensary	2822	0.10	0.18	1196	0.09	0.16
Hospital	2822	0.01	0.04	1196	0.01	0.05
Maternity & child welfare centre	2822	0.10	0.17	1196	0.10	0.16
Family welfare centres	2822	0.08	0.16	1196	0.08	0.15
Maternity health centre	2822	0.04	0.13	1196	0.04	0.12
Nutritional centre	2822	0.95	0.11	1196	0.95	0.11
Asha centre	2822	0.82	0.23	1196	0.81	0.23
Tap	2822	0.58	0.38	1196	0.56	0.38
Well	2822	0.57	0.36	1196	0.58	0.36
Handpump	2822	0.80	0.27	1196	0.80	0.28
Tank/pond/lake	2822	0.37	0.30	1196	0.36	0.29
Community toilet complex	2822	0.11	0.19	1196	0.10	0.17
Post office	2822	0.36	0.25	1196	0.35	0.25
Bank/agricultural credit/SHG	2822	0.75	0.26	1196	0.73	0.27
Public distribution dystem	2822	0.66	0.26	1196	0.65	0.26
Total sanitation campaign	2822	0.34	0.36	1196	0.34	0.36
Drainage	2822	0.69	0.30	1196	0.68	0.30
Bus	2822	0.57	0.32	1196	0.56	0.32
Auto	2822	0.48	0.29	1196	0.47	0.29
Mandis/regular market	2822	0.30	0.27	1196	0.29	0.27
Paved roads	2822	0.74	0.26	1196	0.73	0.26
Unpaved roads	2822	0.85	0.20	1196	0.84	0.20
Other Major Roads	2822	0.31	0.22	1196	0.30	0.21
All-weather roads	2822	0.77	0.24	1196	0.76	0.25
National highways	2822	0.08	0.11	1196	0.08	0.10
State highways	2822	0.19	0.16	1196	0.19	0.16
Power (domestic)	2822	0.92	0.20	1196	0.91	0.21
Power (agricultural)	2822	0.75	0.34	1196	0.76	0.34
Power (commercial)	2822	0.62	0.35	1196	0.62	0.34
Electricity	2822	64.11	30.95	1196	63.55	31.04

Note: each of the above variables is an indicator variable, taking a value of 1 if the infrastructure is available in a village and 0 otherwise. The mean of variables above is at the constituency level, indicating the proportion of villages in a constituency that have access to the mentioned infrastructure. Electricity indicates the percentage of households in a constituency that have access to electricity as the main source of lighting. Using the above variables, we construct an index for public goods at the constituency level through a principal component analysis. Full sample includes all constituencies. Mixed sample includes constituencies where among the top two candidates, candidates with graduate degree contest against non-graduate candidates. The table reports mean and standard deviation for a cross section of assembly constituencies for both full and mixed sample.

Source: authors' elaboration.

Table A2: Election years for leaders holding office in 2010

	Year		
	Num	Col %	Cum %
2006	287	25.88	25.88
2007	230	20.74	46.62
2008	69	6.22	52.84
2009	349	31.47	84.31
2010	174	15.69	100

Note: this table summarizes the frequency distribution for the election years in which leaders holding office in 2011 were elected. This represents the estimation sample for the cross-sectional analysis of leaders' impact on constituency level infrastructure.

Source: authors' elaboration.

Table A3: Educational qualification of candidates contesting in elections

	Sex of the candidate								
Education of	Male			Female			Total		
the candidate	Num	Col %	Cum %	Num	Col %	Cum %	Num	Col %	Cum %
Illiterate	28	0.28	0.28	6	0.66	0.66	34	0.31	0.31
Literate	146	1.46	1.74	39	4.31	4.97	185	1.7	2.01
5th pass	180	1.8	3.54	22	2.43	7.40	202	1.85	3.86
8th pass	506	5.06	8.61	60	6.63	14.03	566	5.19	9.06
10th pass	1510	15.11	23.72	143	15.8	29.83	1653	15.17	24.22
12th pass	1520	15.21	38.93	119	13.15	42.98	1639	15.04	39.26
Graduate	4217	42.2	81.13	275	30.39	73.37	4492	41.22	80.48
Post graduate	1649	16.5	97.63	204	22.54	95.91	1853	17	97.49
Doctorate	237	2.37	100	37	4.09	100	274	2.51	100
Total	9993	100		905	100		10898	100	

Note: this table is based on information about winner and runner-up candidates. A person who has completed college and holds a bachelor's degree is considered a graduate leader.

Source: authors' elaboration.

Table A4: Balance test of constituency and candidate characteristics between elections with graduate and non-graduate winners in mixed sample within 7% margin of victory

	Mixed sample within 7% margin		
	Graduate winner	Non-graduate winner	Difference
Constituency characteristics			
Electorate size (log)	11.988	12.007	0.018
	(0.672)	(0.677)	(0.041)
Voters (log)	11.599	11.605	0.006
	(0.604)	(0.621)	(0.037)
Turnout	69.216	68.292	-0.923
	(13.801)	(13.154)	(0.821)
SC constituency	0.135	0.157	0.022
	(0.342)	(0.364)	(0.022)
ST constituency	0.091	0.114	0.022
	(0.288)	(0.318)	(0.018)
Total number of candidates	8.112	8.157	0.044
	(5.607)	(5.999)	(0.354)
Average age of candidates	46.204	46.033	-0.171
	(5.734)	(5.652)	(0.347)
Average years of education of candidates	12.087	12.072	-0.014
	(1.691)	(1.684)	(0.103)
Average assets (log) of candidates	14.176	14.266	0.089
	(1.133)	(1.219)	(0.072)
Average liabilities (log) of candidates	12.649	12.623	-0.026
	(1.345)	(1.270)	(0.083)
Proportion of female candidates	0.065	0.068	0.004
	(0.111)	(0.119)	(0.007)
Proportion of criminal candidates	0.221	0.211	-0.010
	(0.238)	(0.234)	(0.014)
Candidate characteristics			
Female winner	0.061	0.092	0.031*
	(0.239)	(0.289)	(0.016)
Winner's age	49.808	49.414	-0.394
	(10.383)	(10.295)	(0.645)
Winner's criminality	0.291	0.344	0.053*
	(0.455)	(0.476)	(0.028)
Winner's alignment with state political party	0.537	0.510	-0.027
	(0.499)	(0.500)	(0.031)
Winner's assets (log)	15.588	15.472	-0.116
	(1.542)	(1.626)	(0.097)
Winner's liabilities (log)	7.633	7.793	0.160
	(6.695)	(6.634)	(0.406)
Observations	525	555	

Note: standard errors are given in parentheses. Mixed sample includes elections where one of the top two candidates is a graduate while the other is a non-graduate. We further restrict the sample to observations within 7% margin of victory to consider close elections. As seen from the table, 49% of mixed close elections are won by a graduate. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A5: Probability that graduate candidate wins in close elections

	(1)	(2)	(3)
Electorate size (log)	0.612	1.107*	0.276
	(0.406)	(0.566)	(1.183)
Voter (log)	-0.616	-1.184**	-0.422
	(0.398)	(0.563)	(1.159)
Turnout	0.011*	0.023**	0.012
	(0.007)	(0.010)	(0.020)
SC constituency	-0.045	-0.033	-0.067
	(0.048)	(0.051)	(0.067)
ST constituency	-0.078	-0.111	-0.068
	(0.062)	(0.072)	(0.139)
Total number of candidates	0.000	-0.005	-0.004
	(0.003)	(0.005)	(0.007)
Average age of candidates	0.000	0.002	0.006
	(0.003)	(0.003)	(0.005)
Average years of education of candidates	0.007	0.007	0.003
	(0.010)	(0.011)	(0.015)
Average assets (log) of candidates	-0.017	-0.019	-0.014
	(0.017)	(0.020)	(0.029)
Average liabilities (log) of candidates	0.009	0.017	0.016
	(0.013)	(0.014)	(0.019)
Proportion of female candidates	-0.023	-0.013	0.159
	(0.147)	(0.150)	(0.207)
Proportion of criminal candidates	0.029	0.006	0.106
	(0.074)	(0.080)	(0.104)
Observations	1,030	1,030	1030
R-squared	0.007	0.035	0.328
F statistics	0.682	1.159	0.813
P-value for joint significance	0.770	0.309	0.637
Year fixed effects	No	Yes	Yes
State fixed effects	No	Yes	Yes
District fixed effects	No	No	Yes

Note: robust standard errors clustered at the constituency level are given in parentheses. We consider the dependent variable to be a dummy indicating whether winner in close election is a graduate. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A6: OLS estimates for effect of electing a graduate leader on growth in night-time lights

	Full Sample			Mixed Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Graduate leader	0.504*	0.297	0.149	1.146***	0.895**	0.291
	(0.265)	(0.233)	(0.210)	(0.401)	(0.351)	(0.340)
Observations	13,909	13,909	13,865	6,081	6,081	6,064
R-squared	0.000	0.472	0.511	0.001	0.466	0.520
Year fixed effects	No	Yes	Yes	No	Yes	Yes
State fixed effects	No	Yes	Yes	No	Yes	Yes
District fixed effects	No	No	Yes	No	No	Yes

Note: robust standard errors clustered at the constituency level are given in parentheses. Graduate leader is a dummy variable that is equal to 1 if a graduate candidate wins against any other candidate, and 0 if it is vice versa. Full sample includes all constituencies. Mixed sample includes constituencies where a graduate candidate wins in an election against a non-graduate candidate or vice versa. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A7: Heterogeneous effect of leader's education on growth rate of night-time lights by leader's gender

	(1)	(2)	(3)	(4)
Male leader				
Graduate Leader	3.471**	3.167***	2.039**	4.503**
	(1.403)	(1.178)	(0.941)	(2.214)
Observations	2,627	3,202	4,093	1,385
R-squared	0.580	0.566	0.550	0.595
Bandwidth	7.273	9.648	14.55	3.637
Female leader				
Graduate leader	1.549	-5.642	3.603	-3.884
	(8.429)	(9.793)	(3.763)	(12.025)
Observations	319	246	423	185
R-squared	0.621	0.641	0.602	0.690
Bandwidth	11.48	8.649	22.95	5.738
Bandwidth type	IK(h)	CCT	2h	h/2
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. The upper panel only includes male leaders and the lower panel only includes female leaders. Both panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A8: Heterogeneous effect of leader's education on growth rate of night-time lights by leader's age

	(1)	(2)	(3)	(4)
Younger leaders				
Graduate leader	4.463**	4.475**	2.499**	5.107*
	(1.777)	(1.748)	(1.187)	(2.948)
Observations	1,588	1,626	2,437	874
R-squared	0.570	0.570	0.546	0.589
Bandwidth	8.775	9.039	17.55	4.387
Older leaders				
Graduate leader	2.959	3.452**	2.630**	-0.736
	(2.199)	(1.744)	(1.243)	(4.403)
Observations	1,566	1,905	2,389	852
R-squared	0.624	0.609	0.597	0.645
Bandwidth	8.135	10.81	16.27	4.067
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK (h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. We consider sub-samples of leaders by their age, defining younger leaders as those below the median age and older leaders as those above it. All panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A9: Heterogeneous effect of graduate leader by leader's alignment with political party in state legislature

	(1)	(2)	(3)	(4)
Aligned with state ruling party				
Graduate Leader	5.523**	3.301*	1.458	4.188
	(2.165)	(1.751)	(1.315)	(2.714)
Observations	1,701	2,118	2,772	874
R-squared	0.553	0.539	0.530	0.576
Bandwidth	8.744	11.96	17.49	4.372
Not aligned with state ruling party				
Graduate leader	4.639**	4.225***	3.654**	5.661
	(2.070)	(1.538)	(1.434)	(3.669)
Observations	1,365	1,816	1,931	801
R-squared	0.623	0.607	0.598	0.644
Bandwidth	7.622	12.49	15.24	3.811
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. The upper panel includes constituencies where the leader belongs to the ruling party in power at the state level and the lower panel includes constituencies where the leader does not belong to the party in power at the state level. All panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A10: Heterogeneous effect of leader's education on growth rate of night-time lights by leader's criminal record

	(1)	(2)	(3)	(4)
Leader without criminal record				
Graduate leader	3.677***	3.305***	2.432**	4.360*
	(1.329)	(1.127)	(0.944)	(2.492)
Observations	1,995	2,398	3,084	1,027
R-squared	0.610	0.594	0.578	0.632
Bandwidth	7.210	9.428	14.42	3.605
Leader with criminal record				
Graduate leader	5.723*	3.169	2.910	7.207*
	(3.324)	(2.398)	(2.163)	(4.303)
Observations	975	1,359	1,494	535
R-squared	0.566	0.547	0.542	0.575
Bandwidth	8.849	14.59	17.70	4.424
Bandwidth type	IK(h)	CCT	2h	h/2
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. The upper panel includes leaders who do not have any criminal cases registered against them, while the lower panel includes leaders with criminal accusations. All panels control for district and year fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A11: Robustness check—sensitivity analysis of RD specification to higher order polynomials

	Quadratic				Cubic			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Graduate leader	2.139	3.288**	3.269**	3.625	2.518	3.626***	2.956*	5.058*
	(1.828)	(1.325)	(1.357)	(2.392)	(2.299)	(1.387)	(1.748)	(2.842)
Observations	2,883	4,655	4,517	1,520	2,883	5,627	4,517	1,520
R-squared	0.002	0.001	0.001	0.003	0.002	0.001	0.001	0.003
Bandwidth	7.427	15.72	14.85	3.714	7.427	26.10	14.85	3.714
Year fixed effects	No	No	No	No	No	No	No	No
State fixed effects	No	No	No	No	No	No	No	No
District fixed effects	No	No	No	No	No	No	No	No
Panel B								
Graduate leader	2.378	3.190***	3.150***	3.177	2.674	3.275**	3.153**	3.717
	(1.697)	(1.217)	(1.204)	(2.393)	(2.199)	(1.302)	(1.587)	(2.829)
Observations	2,883	4,463	4,517	1,520	2,883	5,388	4,517	1,520
R-squared	0.506	0.488	0.488	0.507	0.506	0.482	0.488	0.508
Bandwidth	7.427	14.54	14.85	3.714	7.427	22.42	14.85	3.714
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	No	No	No
Panel C								
Graduate leader	1.411	3.106**	3.067**	3.650	0.221	2.841**	2.439	2.958
	(1.796)	(1.229)	(1.219)	(3.211)	(2.423)	(1.275)	(1.508)	(4.409)
Observations	2,876	4,448	4,502	1,516	2,876	5,373	4,502	1,516
R-squared	0.580	0.550	0.550	0.599	0.580	0.540	0.550	0.600
Bandwidth	7.427	14.54	14.85	3.714	7.427	22.42	14.85	3.714
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate, and 0 if it is vice versa. Columns (1)–(4) include the quadratic of control function, and Columns (5)–(8) include the cubic control function. Panels A, B, and C present models that include different fixed effects. In Columns (1)–(8), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A12: Robustness check—sensitivity of estimates to extreme values

	Constituency-year				Constituency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Graduate leader	3.139**	2.897***	2.349**	2.156	2.949**	2.716**	2.243**	1.839
	(1.295)	(1.120)	(0.960)	(1.790)	(1.309)	(1.126)	(0.970)	(1.808)
Observations	2,786	3,541	4,367	1,460	2,756	3,534	4,314	1,441
R-squared	0.002	0.001	0.001	0.003	0.001	0.001	0.001	0.003
Bandwidth	7.360	10.37	14.72	3.680	7.350	10.49	14.70	3.675
Year fixed effects	No	No	No	No	No	No	No	No
State fixed effects	No	No	No	No	No	No	No	No
District fixed effects	No	No	No	No	No	No	No	No
Panel B								
Graduate leader	3.048***	2.828***	2.255***	2.551	2.951**	2.694***	2.159**	2.510
	(1.171)	(1.013)	(0.841)	(1.726)	(1.189)	(1.029)	(0.852)	(1.755)
Observations	2,786	3,452	4,367	1,460	2,756	3,417	4,314	1,441
R-squared	0.517	0.509	0.501	0.516	0.518	0.510	0.502	0.516
Bandwidth	7.360	9.938	14.72	3.680				
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	No	No	No
Panel C								
Graduate leader	3.452***	3.188***	2.098**	3.933*	3.314**	3.055***	2.038**	3.814*
	(1.292)	(1.072)	(0.873)	(2.121)	(1.316)	(1.089)	(0.887)	(2.148)
Observations	2,779	3,445	4,352	1,456	2,749	3,410	4,299	1,437
R-squared	0.591	0.577	0.562	0.608	0.592	0.578	0.563	0.609
Bandwidth	7.360	9.938	14.72	3.680	7.350	9.928	14.70	3.675
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. In Columns (1)–(4), we drop any observation where the constituency-year pixel mean is 63. In Columns (5)–(8), we drop any constituency in which the mean pixel intensity for any year is 63. Panels A, B, and C present models that include different fixed effects. In Columns (1)–(8), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A13: Robustness check—sensitivity to extreme values

	Constituency-year				Constituency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Graduate leader	3.239**	3.021***	2.417**	2.149	2.946**	2.690**	2.134**	1.966
	(1.331)	(1.150)	(0.983)	(1.842)	(1.335)	(1.152)	(0.988)	(1.852)
Observations	2,723	3,462	4,279	1,442	2,622	3,355	4,146	1,375
R-squared	0.002	0.001	0.001	0.003	0.002	0.001	0.001	0.003
Bandwidth	7.347	10.28	14.69	3.673	7.347	10.28	14.69	3.673
Year fixed effects	No	No	No	No	No	No	No	No
State fixed effects	No	No	No	No	No	No	No	No
District fixed effects	No	No	No	No	No	No	No	No
Panel B								
Graduate leader	2.836**	2.680***	2.114**	2.371	2.836**	2.596**	1.968**	2.490
	(1.188)	(1.023)	(0.852)	(1.748)	(1.206)	(1.049)	(0.864)	(1.779)
Observations	2,723	3,384	4,279	1,442	2,622	3,236	4,146	1,375
R-squared	0.523	0.517	0.510	0.515	0.527	0.523	0.517	0.518
Bandwidth	7.347	9.952	14.69	3.673	7.213	10.15	14.43	3.606
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	No	No	No
Panel C								
Graduate leader	3.419***	3.055***	1.928**	3.984*	3.170**	2.945***	1.895**	3.793*
	(1.297)	(1.076)	(0.884)	(2.090)	(1.342)	(1.124)	(0.898)	(2.148)
Observations	2,716	3,377	4,264	1,438	2,615	3,229	4,131	1,371
R-squared	0.599	0.587	0.573	0.610	0.601	0.591	0.578	0.609
Bandwidth	7.347	9.952	14.69	3.673	7.213	9.583	14.43	3.606
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. In Columns (1)–(4), we drop any observation where the constituency-year pixel mean is 63 or 0. In Columns (5)–(8), we drop any constituency in which the mean pixel intensity for any year is 63 or 0. Panel A, B, and C estimate RD by varying the degree of fixed effects. In Columns (1)–(8), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A14: Robustness check—effect of graduate leader on the past growth rate in night lights

	(1)	(2)	(3)	(4)
Panel A				
Graduate leader	0.053	0.305	0.306	-0.156
	(1.237)	(0.933)	(0.932)	(1.617)
Observations	2,773	4,397	4,402	1,450
R-squared	0.000	0.000	0.000	0.000
Bandwidth	7.083	14.13	14.17	3.541
Year fixed effects	No	No	No	No
State fixed effects	No	No	No	No
District fixed effects	No	No	No	No
Panel B				
Graduate leader	-1.102	-0.895	-0.186	-1.494
	(1.041)	(0.923)	(0.761)	(1.385)
Observations	2,773	3,339	4,402	1,450
R-squared	0.520	0.517	0.514	0.524
Bandwidth	7.083	9.103	14.17	3.541
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No
Panel C				
Graduate leader	-0.814	-0.463	0.225	-2.073
	(1.174)	(1.013)	(0.803)	(1.733)
Observations	2,766	3,332	4,387	1,446
R-squared	0.619	0.607	0.595	0.655
Bandwidth	7.083	9.103	14.17	3.541
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate and 0 if a non-graduate candidate wins against a graduate candidate. In Columns (1)–(4), we estimate the effect of electing a graduate leader on lagged growth rate in night lights in the previous election term. In Panels A, B, and C, we vary the level of fixed effects. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A15: Robustness check—estimates after controlling for additional characteristics of leader

	(1)	(2)	(3)	(4)
Graduate leader	3.179**	3.098***	2.131**	3.623*
	(1.263)	(1.069)	(0.854)	(2.014)
Criminally accused leader	1.162	0.983	0.543	2.001
	(0.794)	(0.695)	(0.587)	(1.259)
Female leader	0.784	0.356	0.003	1.924
	(0.902)	(0.805)	(0.700)	(1.671)
Observations	2,854	3,480	4,487	1,484
R-squared	0.578	0.565	0.549	0.594
Bandwidth	7.427	9.782	14.85	3.714
Year fixed effects	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Bandwidth type	IK(h)	CCT	2h	h/2

Note: robust standard errors clustered at the constituency level are given in parentheses. In all panels, graduate leader is a dummy variable that is 1 if a graduate candidate wins against a non-graduate candidate, and 0 if it is vice versa. Criminally accused leader is a dummy variable that is equal to 1 when a leader has one or more criminal charges filed against them and equal to 0 otherwise. Female leader refers to a woman candidate who wins the election. In Columns (1)–(4), the RD model is estimated by a local linear regression using a triangular kernel. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A16: Balance test on candidate and constituency characteristics—mixed and non-mixed samples

	Mixed sample	Non-mixed sample	Difference
Electorate size (log)	11.963	11.912	-0.051**
	(0.722)	(0.795)	(0.021)
Voters (log)	11.557	11.509	-0.048**
	(0.659)	(0.723)	(0.019)
Turnout	68.129	68.362	0.233
	(13.582)	(13.771)	(0.368)
SC constituency	0.151	0.153	0.001
	(0.359)	(0.360)	(0.010)
ST constituency	0.112	0.115	0.003
	(0.315)	(0.319)	(0.009)
Total number of candidates	7.789	7.438	-0.352**
	(5.615)	(5.555)	(0.150)
Average age of candidates	46.193	46.734	0.541***
	(5.564)	(5.795)	(0.153)
Average years of education of candidates	12.074	12.609	0.535***
	(1.685)	(2.607)	(0.061)
Average assets (log) of candidates	14.246	14.236	-0.010
	(1.209)	(1.196)	(0.032)
Average liabilities (log) of candidates	12.671	12.715	0.044
	(1.369)	(1.420)	(0.040)
Proportion of female candidates	0.073	0.078	0.004
	(0.123)	(0.137)	(0.004)
Proportion of criminal candidates	0.210	0.198	-0.012**
	(0.230)	(0.223)	(0.006)
Female winner	0.078	0.080	0.002
	(0.268)	(0.271)	(0.007)
Winner's age	49.768	50.092	0.323
	(10.361)	(10.104)	(0.283)
Winner's criminality	0.316	0.302	-0.014
	(0.465)	(0.459)	(0.012)
Winner's alignment with state political party	0.605	0.609	0.004
	(0.489)	(0.488)	(0.014)
Winner's assets (log)	15.588	15.552	-0.036
	(1.673)	(1.702)	(0.045)
Winner's liabilities (log)	7.813	7.802	-0.011
	(6.710)	(6.791)	(0.182)
Observations	2390	3305	

Note: standard errors are given in parentheses. The full sample is divided into two mutually exclusive sub-samples, i.e. mixed and non-mixed samples. Mixed sample consists of elections where, among the top two candidates, one was graduate while the other was non-graduate. Non-mixed sample consists of elections where, among the top two candidates, either both were graduates or both were non-graduates. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.

Table A17: Balance test on candidate and constituency characteristics—RD and non-RD samples

	RD sample	Non-RD sample	Difference
Electorate size (log)	12.005	11.926	-0.079***
	(0.667)	(0.766)	(0.030)
Voters (log)	11.606	11.514	-0.092***
	(0.607)	(0.699)	(0.027)
Turnout	68.587	67.723	-0.864
	(13.633)	(13.529)	(0.557)
SC constituency	0.145	0.157	0.012
	(0.352)	(0.364)	(0.015)
ST constituency	0.105	0.118	0.013
	(0.307)	(0.322)	(0.013)
Total number of candidates	8.147	7.472	-0.675***
	(5.835)	(5.396)	(0.230)
Average age of candidates	46.154	46.228	0.075
	(5.690)	(5.452)	(0.228)
Average years of education of candidates	12.083	12.065	-0.019
	(1.678)	(1.692)	(0.069)
Average assets (log) of candidates	14.224	14.266	0.042
	(1.180)	(1.235)	(0.050)
Average liabilities (log) of candidates	12.627	12.710	0.083
	(1.307)	(1.422)	(0.059)
Proportion of female candidates	0.066	0.080	0.014***
	(0.114)	(0.131)	(0.005)
Proportion of criminal candidates	0.216	0.205	-0.010
	(0.235)	(0.226)	(0.009)
Female winner	0.076	0.080	0.004
	(0.265)	(0.271)	(0.011)
Winner's age	49.563	49.953	0.390
	(10.256)	(10.456)	(0.437)
Winner's criminality	0.320	0.313	-0.007
	(0.467)	(0.464)	(0.019)
Winner's alignment with state political party	0.526	0.676	0.150***
	(0.500)	(0.468)	(0.020)
Winner's assets (log)	15.528	15.642	0.114*
	(1.608)	(1.727)	(0.069)
Winner's liabilities (log)	7.739	7.878	0.139
	(6.660)	(6.757)	(0.275)
Observations	1124	1266	

Note: standard errors are given in parentheses. RD sample and non-RD sample are mutually exclusive. After restricting to mixed sample (i.e. elections where one of the top two candidates is a graduate while the other is a non-graduate), RD sample includes close election observations included in estimation of Equation (3), and non-RD sample includes observations that are excluded from Equation (3). The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: authors' calculations.