Essays on Political Economy and Institutions

Rolly Kukreja

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

INDIAN STATISTICAL INSTITUTE
New Delhi

Essays on Political Economy and Institutions

Rolly Kukreja

Thesis Supervisor: Prof. Abhiroop Mukhopadhyay

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

INDIAN STATISTICAL INSTITUTE
New Delhi

March 2022

Acknowledgments

I would like to express my heartfelt gratitude to my advisor, Prof. Abhiroop Mukhopadhyay for his unwavering support and excellent guidance through the years of my research. Not only has he been the source of countless interesting discussions about research ideas but his constant emphasis on academic rigour has always pushed me to improve upon the more technical qualities of my work. I would like to thank him for always stressing the importance of integrity and ethics in research, these are principles that will continue to guide my work for the rest of my life. I am especially grateful to him for being extremely understanding and cognizant of the mental health issues that can sometimes accompany a period of prolonged doctoral research but are often not spoken about.

I would like to extend my gratitude to all faculty members at the Economics and Planning Unit, Indian Statistical Institute, Delhi for fostering an environment that allows one to learn and ask questions without feeling intimidated. I would like to thank Prof. Farzana Afridi whose courses formed the basis of the first steps to my journey in empirical research. I also thank Prof. Tridip Ray and Prof. Chetan Ghate for their constant encouragement and support through the period of my research. Outside Indian Statistical Institute, I would like to convey special thanks to Prof. Sabyasachi Das for always providing valuable comments and feedback on my research.

I am grateful to all my peers and colleagues for being the source of many interesting discussions and many joyful times spent together. I would especially like to mention my senior, Prachi Singh who was a source of support and help in many a time of need. I would like to thank the Indian Statistical Institute for providing an environment that is conducive to research and to the administrative and support staff for always being helpful.

The journey to the completion of a doctoral dissertation is fraught with challenges, both of the intellectual and psychological kind. I cannot thank my parents enough for

being my biggest source of strength, for always being patient, for always trusting me, for always being kind even when circumstances were not. The desire to make them proud has been the utmost source of motivation to overcome any hurdles through this journey. I am grateful to them for giving me the freedom to pursue the career I wished for without any questions or doubts. I would also like to thank my in-laws for being extremely gracious with their patience and support in view of the time and space that research work demands.

I would also like to thank my grandparents, uncles and aunts, some of whom aren't around anymore, but whom I know would have been proud to see me finish. Special thanks to my cousins–brothers and sisters who have always been a source of love, support and fun times. My friends– Ira and Gunjan for offering words of encouragement whenever I hit a roadblock.

Finally, I would like to thank my partner, Saurabh, for being my biggest cheerleader throughout the period of our relationship, for truly believing in me when I did not, and for bringing pure joy to my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis consists of three empirical essays that explore distinct themes in the realm of political economy and economics of institutions. All, in essence, are investigations into the role of distinct attributes of political institutions and political actors induced by these institutions, in shaping varied development outcomes. The first chapter begins by asking whether a change in political representation leads to a change in expressed trust in political institutions. This study shows that increased representation leads to higher political trust and further suggests that such improvement in political trust is accompanied by increase in voter turnout and mediated by an improvement in infrastructure and public goods provision. The second chapter shifts focus from features of political institutions to characteristics of political actors and their consequences for human capital. It investigates the impact of criminality of elected representatives on the provision of publicly financed higher education institutions and finds that electoral districts that elect criminally accused leaders witness reduced provision of such institutions. The third chapter continues with the broad narrative of characteristics of politicians and evaluates a policy that introduced minimum education requirements for candidates contesting local elections in the context of a large state in India. Findings from this study show that while education of leaders in itself might be perceived as a desirable outcome, restricting candidature based on education qualifications can lead to lower representation of historically disadvantaged communities and women while not leading to any demonstrable improvement in leader performance.

The sections that follow provide an overview of each of the essays, outlining the research questions, empirical strategies and main results.

1.1 Does Political Representation affect Trust in Political Institutions?: Evidence from India

The first chapter lies at the confluence of two important strands of literature– one focusing on the origins and consequences of trust in political institutions and the second establishing the importance of representation for distribution of material resources. Most industrialized economies have reported a decline in reported confidence in political institutions especially since the 1960s. This observed decline has given rise to a considerable body of literature devoted to understanding the determinants and consequences of political trust (Hetherington, 1998; Newton and Norris, 2000; Catterberg and Moreno, 2006; Mishler and Rose, 2001). For instance, evidence has shown that higher political trust is correlated with greater law compliance and tax morale, and greater support for redistributive policies (Hetherington et al., 2005; Rudolph and Evans, 2005; Torgler, 2005; Scholz and Lubell, 1998). On the other hand, studies on the impact of inequality in representation have unraveled a negative relationship between per capita representation and share of resources allocated to an electoral district (Ansolabehere et al., 2002; Atlas et al., 1995; Horiuchi and Saito, 2003; Knight, 2008; Bhavnani, 2018). Combining these two notions together, this chapter seeks to answer whether a change in political representation in the context of state legislative assemblies, has an impact on self reported confidence (trust) in two political institutions–namely, the state government and politicians.

This study uses self reported confidence in political institutions from two waves of a household level panel data set–the IHDS (India Human Development Survey) held in 2004-05 and 2011-12. To identify a source of variation in representation, this paper uses the delimitation exercise of 2008 which essentially reallocated legislative assembly constituencies among administrative districts within a state, while keeping the total number of constituencies fixed at the pre-2008 level, generating plausibly exogenous variation in representation. This study employs a difference-in-differences estimation strategy with household fixed effects, and defines two categories of treatment for districts–gaining seats and losing seats. Estimates from the difference-in-differences exercise show that individuals in districts that gained seats due to the delimitation show an improvement in confidence in both the state government and politicians by about 6 and 8 percentage points respectively, whereas there is no change in confidence for districts that lost seats.

Given the evidence that higher political trust might be correlated with higher political participation and therefore, voter turnout (Hooghe and Marien, 2013; Grönlund and Setälä, 2007), this paper further investigates if the change in representation is associated with a change in voter turnout and finds that voter turnout indeed increases in districts that gained seats. Finally, using an index of rural development constructed as a composite of availability of rural infrastructure and amenities, this paper provides suggestive evidence that the increase in confidence is mediated by an improvement in development of districts that gained seats, lending credibility to theories that posit that political trust originates in the actual performance of political institutions.

1.2 The Political Economy of Higher Education: Criminality and Institutions

The role of various attributes of political actors is widely acknowledged for being instrumental in explaining economic growth and development outcomes. The criminality of politicians is one such attribute that has been the subject matter of empirical investigations in recent literature. For example, it has been shown that criminal politicians lead to lower economic growth (Prakash et al., 2019), lower private investment and returns (Nanda and Pareek, 2016), lower welfare (Chemin, 2012) and promote crime (Prakash et al., 2021). This chapter augments this literature by investigating the impact of politician criminality on a hitherto unexplored area, that of human capital by looking at the provision of publicly financed higher educational institutions. The role of higher education attainment in the population for promoting economic growth has been emphasized in empirical literature (Castelló-Climent et al., 2018; Castelló-Climent and Mukhopadhyay, 2013). At the same time, very little is known about the role of political factors in provision of tertiary education in India. This chapter seeks to bridge this gap by asking if the election of candidates accused of serious crimes has a deleterious impact on the provision of state government funded and managed colleges in the context of state legislative assembly elections in India.

This study uses the location, year of establishment and management type of state government funded colleges obtained from the All India Survey on Higher Education (AISHE) 2016-17 and maps them to state legislative assembly constituencies to construct

a constituency year level panel of new and existing number of colleges in a constituency. This is combined with data on candidates' criminal records obtained from their affidavits filed at the time of nomination. Focusing on elections held between 2004 and 2008, this study employs a regression discontinuity design limiting the sample to close elections between candidates accused of serious crimes and candidates not accused of serious crimes to identify the impact of criminality of politicians on new state government funded colleges. The regression discontinuity estimate shows that the probability of a new state government funded college being opened in a year in constituencies that elect a candidate accused of serious crimes is lower by 4 percentage points compared to constituencies that did not elect a candidate accused of serious crimes. This translates into a 19 percentage point lesser likelihood of witnessing a new state government college opening in an entire electoral term.

Further, this chapter shines light on interesting heterogeneity in the impact of criminal politicians based on other politician characteristics such as whether the elected candidate is affiliated with the Chief Minister's party and if the candidate is a re-elected incumbent. In a departure from previous evidence, this study reveals that the detrimental impact of criminality is only manifested when the elected candidate is affiliated with the Chief Minister's party emphasizing the importance of access to power (through alignment) in influencing exclusively state supplied outcomes like state government funded colleges. Finally, using rates of completion of higher education from the 71st round of the National Sample Survey, this study provides suggestive evidence that such lower provision of state government colleges is most likely to impact the lowest income groups in society.

1.3 Educated Leaders through Legislation but at what Cost?

Higher education is widely perceived as a credible proxy for the quality of leaders (Besley et al., 2005; Baltrunaite et al., 2014). This is substantiated by studies which show that educated leaders lead to better outcomes in terms of economic growth (Besley et al., 2011) and educational attainment of citizens (Diaz-Serrano and Pérez, 2013). Given the emphasis on leader education as a desirable attribute, there have been attempts to design policies that promote election of educated leaders. This chapter investigates the impact

of such a policy introduced in the context of a large state in India- Rajasthan, which mandated minimum education qualifications for candidates contesting local elections in 2014. While the objective of such minimum education restrictions is to improve efficiency and accountability of leaders, they can be criticized on grounds of being potentially exclusionary, discriminatory and undemocratic. Given this backdrop, this chapter asks whether the introduction of minimum education requirements had an impact on other characteristics of elected leaders, namely their caste category and gender.

This study uses data from two rounds of village council elections held in Rajasthan– one before the policy was introduced in 2010 and the other held after the policy in 2015. To identify variation in the impact of this policy, we use whether the incumbent elected in 2010, as being representative of the most viable candidate, met the education criteria instituted before the 2015 election. Combining the data on the existing incumbent's education with data on caste category and gender of the elected leader in 2010 and 2015, this study employs a difference-in-differences estimation strategy. Estimates from this exercise show that there was a decline in the probability of the elected leader belonging to Scheduled Tribes and Other Backward Castes along with a decline in the probability of election of a woman leader, where the policy hit the hardest in terms of the existing incumbent not meeting the minimum education criteria. These results are supported by checks for pre-existing trends and a gamut of robustness checks.

Further, this study investigates whether the minimum education criteria led to an improvement in leader performance as measured in terms of actual outcomes. Specifically, we look at per capita work days generated under the NREGS (National Rural Employment Guarantee Scheme), economic activity as measured by night time luminosity and at human capital outcomes as measured by new school openings. We find no improvement in any of these outcomes in the aftermath of introduction of the minimum education restriction. This study highlights the unintended consequences of a policy that seeks to improve leader quality by forcing minimum education qualifications in a context of unequal access to education and paints a note of caution for policy makers.

Chapter 2

Does Political Representation affect Trust in Political Institutions?: Evidence from India

2.1 Introduction

Public trust in political institutions is increasingly been acknowledged as being crucial to democracy as it provides legitimacy and stability to political institutions, thus enabling them to perform their role in society. The notion of political trust has received considerable attention in the past couple of decades, particularly against the backdrop of the declining political trust being reported in industrialized economies (especially the United States) since the 1960s. This has spawned a sizeable literature devoted to determining the possible causes and implications of this decline (Hetherington, 1998; Newton and Norris, 2000; Catterberg and Moreno, 2006)[1], the understanding of which could, in turn, help in arresting such a fall.

On the other hand, the implications of political trust for citizens' attitudes and behaviour have been empirically established. For instance, higher political trust has been documented to be positively associated with greater support for redistributive policies (Hetherington et al., 2005; Rudolph and Evans, 2005), greater law compliance and tax

[1]Most of these studies have relied on self reported confidence in public institutions from periodic surveys such as the World Values Survey (WVS), the American National Election Studies (NES) and the General Social Surveys (GSS) for measuring political trust. That are typically repeated cross sections that can provide valuable insights as to the trends in reported confidence in a host of political institutions like the parliament, politicians, political parties, support for democracy etc.

morale (Torgler, 2005; Scholz and Lubell, 1998). Given the fact that observed trust in political institutions is a fluid characteristic that has shown variation over time, combined with the evidence that such trust has consequences for citizens' behaviour and attitudes, it is of interest to understand the factors that determine its evolution. In particular, it becomes important to understand if characteristics of political institutions, such as the way they allocate representatives, can move citizens' political trust. In this paper, I study the causal impact of political representation as measured by the number of legislators representing an electorate, on trust in political institutions and examine how this impact is mediated by economic outcomes.

The manner of allocation of political representatives within a population is an important feature of political institutions. Empirical evidence points to the existence of a *big constituency disadvantage* where overrepresented constituencies[2] are favoured over underrepresented[3] ones by political constituents in distribution of resources as the value of a single vote is higher in overrepresented constituencies (Ansolabehere et al., 2002; Atlas et al., 1995; Horiuchi and Saito, 2003; Bhavnani, 2018)[4]. This is supported by theoretical literature that points to an inverse relationship between constituency size and per capita economic outcomes (Knight, 2008; Atlas et al., 1997). However, it remains unknown if the impact of such variation in representation extends beyond economic outcomes and also translates to disparities in citizens' political trust.

In this paper, I estimate the causal impact of political representation on two measures of political trust– confidence in politicians and confidence in government in the context of state level politics in India. I use the change in representation generated by a redistricting exercise in the setting of state legislative assembly constituencies in India. The redrawing of electoral boundaries was suspended in India from 1976 until 2008[5]. When boundaries were redrawn in 2008, the objective of equalization of representation across regions necessitated a reallocation of seats of the state legislative assembly among the administrative districts in a state. As a consequence, some districts gained seats in the legislative assembly, some lost seats and others remained unchanged. This exercise pro-

[2]Constituencies with smaller than average population size

[3]Constituencies with greater than average population size

[4]A related strand of literature focuses on the impact of representation by highlighting concerns associated with *gerrymandering* which is the act of manipulating electoral boundaries so as to accord an unfair advantage to a particular party or group (Gul and Pesendorfer, 2010; Coate and Knight, 2007)

[5]See section on delimitation for details

vides us with plausibly exogenous[6] variation in representation that can be exploited to estimate the impact of change in representation on political trust.

Using this change in representation generated by the redistricting exercise along with respondents' assessments of their confidence in politicians and state government from two waves of a household level panel dataset– the IHDS (India Human Development Survey) held in 2004-05 and 2011-12 as measures of political trust, this paper first implements a difference-in-differences strategy with two treatment groups to estimate the causal impact of increase in representation on political trust. Estimates from this exercise show that in districts that gained seats, households that report having confidence in politicians increase by 8 percentage points and households that report having confidence in the state government increase by about 6 percentage points, relative to households in districts that experienced no change in seat allocation. There is no negative impact of losing representation on political trust for households living in districts that lost seats- the other treatment group.

Next, I investigate if the observed improvement in political trust is also associated with changes in citizens' behaviour, in particular, citizens' turning out to vote. Political participation of citizens remains one of the most widely discussed behavioural aspects in the context of political trust (Citrin, 1974; Hetherington, 1999) and voter turnout provides a direct measure of such political participation. Several political scientists have shown that trust stimulates voter turnout (Hooghe and Marien, 2013; Grönlund and Setälä, 2007). This is based on the hypothesis that a positive perception or judgement of political institutions is necessary to induce people to participate politically.

Estimates from the same difference-in-differences strategy with the district level voter turnout as dependent variable show that voter turnout was higher by about 1.2 percentage points in the gaining districts relative to the unchanged districts in the post delimitation period. Mirroring the results for political trust, there was no significant result for the districts losing seats. This also serves as an additional robustness check to verify pre-trends as voter turnout data is available for earlier periods even though the unavailability of the political trust data for earlier periods makes it impossible to check for pre-trends in political trust.

I suggest three possible channels that could be mediating this change in political

[6]More on this in the data section

trust. The first channel derives from the literature highlighting the relationship between under-representation and distribution of resources (Horiuchi and Saito, 2003; Knight, 2008; Bhavnani, 2018). For instance, Ansolabehere et al. (2002) find that state transfers in the United States were significantly skewed in favour of over-represented counties before court mandated redistricting took place. If this mechanism held true for the Indian delimitation context, then the observed results follow from the theory of institutional performance that says that improved transfers from the government would lead to improvement in political trust. Moreover, this improvement in performance should reflect in better provision of public goods and infrastructure. Using an index that combines indicators for availability of amenities like electricity, schools, health centres, transport, communication, banking and drinking water, I find that districts that gained seats show an improvement in availability of public goods and infrastructure in rural areas.

The second channel proposes that an improvement in politician quality could lead to higher political trust. This hypothesis hinges on the idea that addition of seats in a district gives opportunities for new candidates to enter, who could presumably have more desirable characteristics. I find a decrease in age and criminal cases of candidates in districts that gained seats relative to unchanged districts, but no change in number of criminal cases for actual winners.

The third channel posits that a smaller constituency size improves channels for communication, elicits better responsiveness from representatives and enables greater mobilization of voters that in turn improves political trust. This theory also draws from the literature on decentralization that argues that political trust should be higher for lower levels of government as well as for smaller size constituencies due to better responsiveness of and higher interaction with political representatives (Denters, 2002; Hansen, 2013; Rahn and Rudolph, 2005). While the lack of data does not allow a direct test for the third channel, results for voter turnout show that this improvement in political trust is concomitant with an increase in political participation as measured by voter turnout in state assembly elections. This provides some suggestive evidence for the mobilization channel.

A host of robustness checks are employed to strengthen the results. Firstly, no similar effects are observed for reported confidence in other institutions like the military, police, media and courts for which there is no a priori reason to expect any changes on account

of change in representation. This gives confidence that the results are not a product of a general improvement in confidence in all institutions and are infact rooted in political considerations. Secondly, addition of state specific time trends and district demographic variables yields more or less similar results. Thirdly, limiting the sample to districts that gained or lost at most one seat does not affect the results, showing that the findings are not driven by districts at the extreme of the delimitation exercise. Fourthly, employing the change in seats as the treatment variable as opposed to dichotomous treatment variable gives the similar results. Finally, a randomization exercise is carried out by randomly reassigning the two treatments (gaining or losing seats) to districts in a state and then estimating placebo treatment effect. This process is repeated to arrive at a distribution of simulated beta coefficients. The p-values from this exercise further lend support to the estimates.

One concern that stems from the construction of the delimitation exercise is that since the reallocation of seats was based on population of the district relative to state, the districts that gained seats would be the ones that had higher rate of population growth during the boundary freeze and presumably different trends. To address this potential source of endogeneity, I use the population of the district relative to the state from the 2001 and 2011 censuses to control for trends in population.

This paper is situated at the confluence of two different strands of literature–one relating to political trust and the other focusing on the role of political representation. The discourse on political trust began by documenting various theories of origins of political trust (Mishler and Rose, 2001; Newton and Norris, 2000) and have gone on to establish the consequences of trust for a whole host of outcomes such as tax morale of citizens, their support for redistributive programs and political participation (Hetherington et al., 2005; Scholz and Lubell, 1998; Rudolph and Evans, 2005). This paper seeks to further augment the understanding of the notion of political trust by establishing causality with another key characteristic of political institutions– political representation.

On the representation aspect, this paper ties in with the literature focusing on the distributive consequences of under-representation. In the case of the United States, Ansolabehere et al. (2002) use an index of representation as an explanatory variable and find that state fiscal transfers were highly skewed in favour of overrepresented counties and that court mandated redistricting led to substantial equalization of distribution of

public funds within states. Horiuchi and Saito (2003) find similar implications for under-representation using the 1994 electoral reforms in Japan.

Providing evidence on the impact of representation on distribution in the Indian context, Bhavnani (2018) finds that bigger (underrepresented) constituencies in legislatures also lose out on representation in the executive and argues that this is due to large political parties focusing on winning the more numerous relatively small constituencies. Bhavnani (2016) uses night time luminosity to show that representation positively impacts economic development in India. Investigating the impact of representation outside the sphere of distribution, Daxecker (2019) uses data from Indian parliamentary elections from 1991 to 2009 and shows that electoral violence is less prevalent in over-represented constituencies.

This paper departs from the aforementioned studies by focusing on a novel outcome of interest, i.e. political trust. At the same time, it reaffirms support for the distributive consequences of representation by showing that districts gaining representation show improvement on an index of publics goods and infrastructure.

The main contribution of this paper lies in documenting an hitherto unexplored causal relationship between political trust and representation. No other causal study has so far looked at these aspects in connection with each other. Secondly, it contributes to the scant literature on confidence in institutions as well as the consequences of political representation in the Indian context.

Thirdly, the use of the IHDS dataset, being a household level panel affords numerous advantages over previous studies which have typically employed surveys like the World Value Surveys (WVS) which are repeated cross sections. Having a household level panel with rich demographic and socio-economic variables allows one to difference out the impact of any time invariant socio-cultural factors that might influence political trust. Hence, these estimates are not confounded by any regional or household specific time invariant factors. Also, the IHDS data provides a plethora of detailed household and individual specific information that allows use of plenty of controls to improve precision of estimates.

Fourthly, it contributes to the literature relating voter participation and representation by providing suggestive evidence that an increase in representation is accompanied by increase in voter turnout. And finally, it provides further evidence in support of the

distributive impact of representation.

The remainder of the paper is organized as follows. Section 2 gives background on political trust and the delimitation exercise in India. Section 3 details the data and descriptive statistics. Section 4 outlines empirical strategy. Section 5 presents the main results. Section 6 describes various falsification tests and robustness checks. Section 7 describes results for voter turnout. Section 8 delves into possible mechanisms. Section 9 discusses observed results in relation to the timing of elections. Section 11 concludes the paper.

2.2 Background

2.2.1 Political Trust

Recent work in economics has emphasized the importance of interpersonal and generalized trust for growth and development outcomes (Zak and Knack, 2001; Tabellini, 2010; Algan and Cahuc, 2010; Nunn et al., 2018), however the determinants and significance of political trust remain relatively unexplored. On the other hand, the notion of political trust has attracted considerable attention in the sphere of political science especially against the backdrop of the declining political trust being reported in industrialized economies since the 1960s.

Most of this literature has focused on developed economies and relied on sources such as the self reported confidence in public institutions from periodic surveys like the various rounds of the World Value Surveys (WVS), the American National Election Studies (NES) and the General Social Surveys(GSS) for measuring political trust. The discourse surrounding political trust began by documenting declining political trust in developed economies (Putnam, 2000; Listhaug and Wiberg, 1995) and has focused on explaining the origins and importance of political trust. For instance, Mishler and Rose (2001) use data from 10 post-Communist countries to test the competing cultural and institutional theories of origins of political trust and find evidence in support of the micro-institutional theory that argues that political trust is determined by institutional performance and further that citizens' evaluations of institutional performance are conditioned by their own experiences with the government rather than aggregate performance. Newton and Norris (2000) also provide evidence in support of the institutional performance theory.

Jain et al. (2015) document differences in political trust between the residents of the states of Andhra Pradesh and Telangana and suggest that this difference maybe correlated with poor governance in Telangana relative to Andhra Pradesh.

A related strand of literature has documented a negative relationship between political trust and perceptions of corruption (Clausen et al., 2011; Seligson, 2002). For instance, Pharr and Putnam (2000) using newspaper reports of corruption in Japan, finds that *officials' misconduct has been by far the single best predictor of citizen citizen confidence in government.* This suggests that political trust is not a fixed characteristic but a fluid one that is ultimately rooted in the actual performance of the government. The motivation of this study lies in exploring whether representation is a factor that can move political trust.

While it is clear that political trust originates in the performance of political institutions, one might be compelled to ask why political trust itself matters at all. The notion of political trust has been considered to be deeply related to the ideas of legitimacy and stability of the government. It's not hard to understand on an intuitive level that political trust is vital for the effectiveness of a wide range of government policies that entail some contribution or sacrifice from the public and evidence points to the same direction.

Studies focusing on the consequences of political trust have documented positive associations between political trust and tax morale (Torgler, 2005), tax compliance (Scholz and Lubell, 1998) as well as support for redistributive policies. Hetherington et al. (2005) for instance, attributes the decline in support for liberal policies in the United States to decreasing trust in the government to implement such policies effectively, repudiating the common narrative of the American public turning more conservative. He uses data on political trust from the National Election Studies (NES) in the United States and argues that the level of political trust affects the support for redistributive programs that are based on targeting racial minorities or the economically disadvantaged because they imply a material sacrifice on part of all but benefits only a few.

Rudolph and Evans (2005) augment this account further and show that political trust matters for policies that entail an ideological sacrifice on part of citizens. Using data from the United States, they show that political trust matters more for conservatives than liberals when it comes to support for government spending. When it comes to political participation, evidence is mixed with regard to the relationship with political

trust. While some political scientists have not found any relationship between voting and political trust (Citrin, 1974), others have found trust to be positively related to institutionalised political participation such as voting and party activity (Hooghe and Marien, 2013). Political trust can also have an impact on voter behavior extending beyond the decision to vote or not. For instance, using data from American National Election Studies spanning over a three decade period, Hetherington (1999) shows that distrustful individuals are more likely to vote for non-incumbents in two party races and for third party candidates in case there are more than two parties.

Therefore, trust in political institutions is a fluid characteristic that can be modified by how political institutions behave towards people. It also matters for a whole host of outcomes such as how effectively the government can garner support for its redistributive efforts or how effectively it can administer its tax policy. At the same time, no study has so far looked at the impact of this particular aspect of political institutions– the allocation of representatives across populations, on political trust. This paper attempts to offer novel evidence in support of this particular channel of representation as another potential force that can move political trust.

2.2.2 Delimitation in India

Delimitation refers to the act or process of fixing limits on boundaries of territorial constituencies[7]. The task of delimitation is assigned to a body called the Delimitation Commission established under the Delimitation Act that is responsible for redrawing boundaries of assembly and parliamentary constituencies[8]. In India, such Delimitation Commissions have been set up four times- in 1952, 1963, 1973 and 2002. In 1976, a constitutional amendment passed by the Indira Gandhi government during the Emergency froze the boundaries of assembly and parliamentary constituencies until the 2001 Census. The reason given for this was that reapportionment of parliamentary seats would reward states with higher population growth rates with greater representation in the parliament thereby reducing the incentive to implement family planning programmes.

As a result, electoral boundaries in India remained unchanged for a period of about three decades. This led to wide inequality in representation with regions with higher

[7]from the website of the Delimitation Commission- http://eci.nic.in/delim/index.asp
[8]constituencies are electoral districts

population growth rates being underrepresented in both the state assemblies and the Parliament and those with relatively lower population growth being overrepresented. Against this backdrop, the Delimitation Commission 2002 was set up with the responsibility of redrawing boundaries and reallocating seats so as to equalize population across constituencies. An amendment to the Delimitation Act in 2003, while allowing for redistribution of parliamentary and assembly constituencies within states, froze the total number of both at the level determined on the basis of the 1971 census until 2026, i.e., the total number of seats in the State Legislative Assemblies as well as the total number of seats assigned to a state in the Lok Sabha would remain unaltered through this delimitation exercise and until the first census after 2026.

The delimitation exercise was carried out based on population figures from the 2001 census. While the total number of seats remained fixed, the number of seats reserved for Scheduled Castes (SCs) and Scheduled Tribes (STs) was to be re-worked based on the 2001 census. The constituencies were to be delimited in a way that population (on the basis of 2001 census) of each parliamentary and assembly constituency in a state shall, so far as practicable, be the same throughout the State[9].

For the purpose of delimitation of the assembly constituencies, the entire population of the state was first divided by the total number of seats in the State Legislative Assembly. The average population per seat arrived at in the first step was to serve as the guiding factor in delimiting constituencies. However, for practical considerations the Delimitation Commission allowed a deviation of 10 percent above or below the state average[10]. In the next step, the seat entitlement of each administrative district was determined by dividing the population of the district by the average population per constituency determined in the first step. If the calculated seat entitlement of the district contained a fraction, fractions greater than one half were counted as one and less than one half were ignored[11]. The change in the number of seats for a district was therefore equal to the determined entitlement minus the existing seats in the district. For example, if a district with a

[9]Delimitation procedure as outlined on the website of the Delimitation Commission of India

[10]"The Delimitation Commission has, however, taken an internal decision that as constituencies cannot be delimited having exactly equal population in all cases, a deviation to the extent of 10 percent plus or minus from the State/district average would be acceptable to the Commission, if the geographical features, means of communication, public convenience, contiguity of the areas and necessity to avoid breaking of administrative units so demand."-Guidelines and Methodology from the website of the Delimitation Commission of India

[11]In practice, there were some deviations from this rule necessitated by the fixing of number of seats at the pre-delimitation level

determined entitlement of 10 seats had 8 constituencies before delimitation, the district would gain 2 seats as a result of the delimitation exercise. Similarly, a district with an entitlement that is less than it's pre-delimitation allocation would end up losing seats in the state assembly. Moreover, the Delimitation Commission decided that, so far as possible, all assembly constituencies in a district should be confined within the territorial limits of that district. This change in the number of constituencies within the unit of an administrative district as a consequence of delimitation, gives us the variation in representation that we can exploit to establish causality with political trust.

Next, the total population of the district was divided by its seat entitlement to arrive at the average population per constituency in the district. The areas of the district were to be divided into the requisite number of assembly constituencies (the entitlement determined in the second step) taking into account the average population per constituency in the district with a permissible deviation of 10 percent plus or minus from the district average. Seats were reserved for SCs and STs in proportion to their population to the total population of the state. The recommendations of the Delimitation Commission were approved by the president in February, 2008. The first election under the newly delimited boundaries was held in the state of Karnataka in May 2008.

While a significant amount of literature focusing on redistricting in the West has concentrated on the presence and implications of *gerrymandering*[12], Iyer and Reddy (2013) find that the redistricting process in India was politically neutral for the most part. This is important because it provides confidence that the results in this paper do not follow from an exercise that was rigged to be advantageous to particular parties. Delimitation was not carried out in Jammu & Kashmir and was deferred for the states of Arunachal Pradesh, Assam, Jharkhand, Manipur and Nagaland[13]. These states are excluded from the analysis as they might have completely different dynamics and would therefore be inappropriate for inclusion in the control group. A total of 492 districts were subjected to delimitation, out of which 110 districts gained seats, 150 districts lost

[12]Gerrymandering refers to the practice of manipulating constituency boundaries to confer political advantage to a particular political party or group

[13]Delimitation in Jammu & Kashmir would require amendment to the state constitution. States of Assam, Nagaland, Manipur and Arunachal Pradesh were exempt from delimitation because of concerns that the census figures did not reflect correct population shares due to large scale influx of illegal migrants who have settled in these states. Delimitation in Jharkhand was stayed because reworking reservation on basis of 2001 census would have led to a decrease in the seats reserved for STs, which led to questions about declining ST population and the veracity of the census figures.

seats and and 232 districts neither lost nor gained seats due to delimitation. Figure 2.1 shows the distribution of the districts by the number of seats gained/lost as a result of delimitation. Households residing in districts that neither gained or lost constitute the control group, whereas households in districts that gained or lost seats make up the two treatment groups.

2.3 Data

2.3.1 Data sources

The primary data source are the two waves of the IHDS Surveys- IHDS I[14] and IHDS II[15]. The first round of the IHDS was carried out in 2004-2005 and provides a nationally representative sample of 41554 households. The second round was carried out in 2011-2012 and surveyed 42152 households which included about 40000 households interviewed in the previous round. Linking IHDS-I and IHDS-II provides a rich household level panel dataset with a wide range of socio-economic characteristics.

The analysis is restricted to the rural sample of the IHDS in order to use data about availability of rural amenities from the census of 2001 & 2011. I also include only those respondents who report living in the same district for a period of at least 8 years. This ensures that results are not confounded by migration. This leaves us with a sample of about 26000 households after excluding states that did not undergo delimitation.

In order to collect data about various components of social capital in the tradition of international surveys like the World Value Surveys and General Social Surveys, both rounds of the IHDS include a section on "Confidence in Institutions" which asks respondents to report their confidence in institutions like politicians, state government, army, police, panchayats/municipalities, newspapers/media, courts, hospitals, schools and banks on a scale of 1 to 3, with 1 corresponding to the highest level of confidence[16]. Of these, I focus on the self reported confidence in politicians and the state government as appropriate measures of political trust. More specifically the questionnaire asks respondents if they have confidence in politicians to fulfill promises and in the state government

[14] Desai et al. (2018)
[15] Desai and Vanneman (2018)
[16] These questions were posed to one individual from each household–a knowledgeable informant who was typically the head of the household

to look after people. Responses are coded as 1 if the respondent reports having a *"great deal of confidence"*, 2 if the respondent reports having *"only some confidence"* and as 3 if the respondent reports having *"hardly any confidence at all"*.

The use of self-reported political trust in literature has been associated with debates about the interpretation of such measures. Studies have been focused on trying to understand whether such measures capture dissatisfaction with the political system or with the existing incumbents, whether they are influenced by short term fluctuations or are stable over time. However it has been generally agreed upon that such measures do capture citizens' political perceptions and evaluations of their political contexts (Levi and Stoker, 2000). Another issue associated with self reported confidence in government is that of social desirability bias– which refers to respondents' propensity to answer questions in a way that they deem would be socially acceptable. However, in this case it would unlikely that such a bias would systematically differ for respondents in districts that gained seats or lost seats as opposed to respondents in districts that neither gained or lost seats.

Figure 2.2 & Figure 2.3 show the break up of the responses for confidence in politicians and the state government for the rural sample for both rounds of the IHDS. For the purpose of my analysis, I transform these variables into dichotomous variables that take value 1 if the respondent reports having either a great deal of confidence or only some confidence and 0 if they report having hardly any confidence. Apart from the advantages conferred by using a household level panel that makes sure that the estimates are not confounded by time invariant cultural or social factors, the IHDS also provides a wide range of socio-economic indicators like education, assets and consumption to be used as controls to improve precision of estimates.

The data source for the main independent variable of interest, i.e , change in representation are the reports and papers available on the website of the Delimitation Commission of India. The delimitation procedure required the publication of working papers detailing the delimitation plans for each state. These working papers after incorporating suggestions from the Associate Members of the Delimitation Commission were worked into draft proposals which were then subject to public sittings and then made into final orders. These working papers, proposals, final papers and orders have all been made available on this website and serve as the source of the data for changes in seat allocation of districts in state legislative assemblies. Since the IHDS reports the location of respon-

dents to the level of the 2001 census districts, the districts as of the 2001 census are the relevant administrative units in this analysis.

The data for candidate and winner characteristics have been made available online by the Association for Democratic Reforms (ADR). This data is only available for 2004 onwards, when it was made compulsory for candidates to file an affidavit declaring their assets, details of criminal cases etc along with their nomination papers. It is to be noted that the dataset does not provide information about all candidates that stood in elections and does not cover all assembly elections as of now.

The village level amenities and public goods data as well as demographic data has been sourced from the Census 2001 and 2011. The voter turnout and other electoral data was obtained from the Lokdhaba website maintained by the Trivedi Centre for Political Data, Ashoka University[17].

2.3.2 Descriptive Statistics

Table 2.1 presents the summary of individual as well as household characteristics of respondents in the total sample and separately for the treatment and control groups from the first wave of the IHDS. Table 2.2 compares the two treatment groups with the control in terms of baseline characteristics. This helps to understand variation in respondents characteristics in treatment and control groups in the pre-delimitation (pre-treatment) phase. Table 2.1 shows that about 39% respondents residing in districts that later gained seats express having confidence in politicians compared with about 40% in districts that remained unchanged and 45% in districts that lost. From Table 2.2, one can see that the differences between these are also statistically significant. Similarly, about 75% respondents in districts that gained seats report having confidence in the state government as opposed to about 78% in areas that lost seats or remained unchanged.

Table 2.1 shows that about 76% respondents in the entire sample are male The average age of respondents in the total sample is 43 years, with the respondents in gaining districts being slightly younger than those in districts that remain unchanged and those in losing districts being slightly older than in the unchanged districts. Table 2.2 further shows that these differences are statistically significant. From Table 2.2, one can see that respondents from unchanged and gaining districts are similar in terms of household size, number of

[17] Bhogale et al. (2019)

children and per capita consumption. The average household size is lower in districts that lost seats. The average years of education is higher for households in losing districts compared with those in unchanged districts, while households in gaining districts report lower average education and highest adult education. Average agricultural land holding is 2.5 acres and is larger in unchanged districts compared with both losing and gaining districts for our sample of households.

It is to be noted that while there are statistically significant differences in the baseline characteristics of respondents in the treatment and control groups, these differences are for the most part small in magnitude. Given the possibility that personal and social characteristics could have an impact on reported trust, I explicitly control for these variables in all regressions.

Table 2.3 shows descriptive statistics for all district level variables employed in the paper while Table 2.4 reports baseline differences in these variables by treatment status. The 2001 Census is used as the source of all demographic data as well as village level amenities and public goods. As is expected, Table 2.3 shows that the gaining districts have a higher population than unchanged districts. The losing districts also have a greater population than the unchanged districts. Table 2.4 shows that these differences are statistically significant. This might seem curious at first but one needs to keep in mind that losing/gaining seats depended on population of a district relative to the state and not the absolute population. The gaining districts also have a higher share of urban population but are similar to unchanged districts in terms of SC, ST population proportions as well as proportion of literate population.

Table 2.4 reports that the difference between gaining and unchanged districts in terms of average household size from the 2001 Census is not significant, mirroring the statistics from IHDS 2004-2005 in Table 2.2. Table 2.4 also shows that the population of a gaining district relative to the state was higher than a district that was unchanged, however the population of a losing district relative to state was not statistically different from the unchanged. To understand this, one should note that the change in seats was determined by seat entitlement (which was calculated based on the population of the district relative to state) as well as the existing number of seats that the district already had. While the relative populations of losing and unchanged districts were similar, it was the difference in the existing number of seats– which were presumably based on relative populations

during the previous delimitation exercise, that determined the change in seats. The rest of the district characteristics are included as controls in robustness checks with the exception of the development index which is a proposed mechanism.

The population weighted development index is calculated by first combining 18 village level indicators of public goods using principal component analysis into an index and constructing a population weighted average of the same index. Table 2.4 shows that this development index is not significantly different for gaining and unchanged districts, but losing districts have better amenities than unchanged districts suggestive of a small constituency bias.

The election data covers state legislative assembly elections from 1998 to 2008. District level voter turnout is not statistically different between unchanged and gaining districts but turnout is greater for losing districts indicating the inverse relationship between constituency size and turnout. The differences in relative representation index[18] by treatment status have expected signs and are statistically significant.

The final two panels in Table 2.3 and Table 2.4 show means and differences in means for candidates and winners in state assembly elections. These were based on candidate affidavit data from the ADR website from 2004 to 2008. Both candidates and winners were younger in gaining and losing districts compared to unchanged. Winners in losing districts are less corrupt in terms of number of criminal cases relative to unchanged districts while there is no significant difference in total assets for winners in either category.

2.4 Empirical Strategy

Given that the first round of the IHDS took place in 2004-05 with the second in 2011-12 and delimitation took place in the intervening years (2008), this setting lends itself to the implementation of a difference-in-differences strategy. As described earlier, respondents in districts that gained seats ended up gaining additional MLAs[19] in the state legislative assembly whereas respondents in the districts that lost seats lost MLAs as a consequence of delimitation. I define respondents living in districts that gained seats as the first treatment group and those living in districts that lost seats as the second treatment group. Respondents living in districts whose seat allocation remained unchanged serve

[18]This is calculated at the village level– exact calculation explained in the mechanism section
[19]Member of Legislative Assembly

as the control group. The following specification is then used to estimate the impact of change in representation:

$$Y_{idt} = \alpha_i + \gamma Post_t + \beta_1 Post_t * Treat1_d + \beta_2 Post_t * Treat2_d + \theta \mathbf{X_{idt}} + \tau Relativepopulation_{dt} + \rho Electionyear_t + \epsilon_{idt} \quad (2.1)$$

This is the usual difference-in-differences specification with multiple treatments. Y_{idt} is the main outcome variable of interest which takes value 1 if the respondent belonging to household i in district d at time t reports having either *"a great deal of confidence"* or *"only some confidence"* in political institutions (politicians or state government), and 0 if respondent reports having *"hardly any confidence at all"*. $Post_t$ takes value 1 for the second wave of the survey. This term captures temporal changes in confidence that do not vary across treatment and control groups. $Treat1_d$ is equal to 1 for all respondents residing in districts that gain seats due to delimitation and 0 otherwise. $Treat2_d$ equals 1 for respondents residing in districts that lost seats and 0 otherwise. The main coefficients of interest are β_1 & β_2. α_i refers to household fixed effects. X_{idt} is a vector of individual and household level controls and includes age, sex, marital status, years of education, household size, number of children, main income source and agricultural landholding. I also allow for differential trends by baseline income quintiles and caste and religion groups[20]. The standard errors are clustered at the district level for all reported regressions.

Since the delimitation exercise reallocated seats based on population shares of districts in a state, districts that grew at a faster rate during the three decade boundary freeze were more likely to gain seats. To assuage concerns that the coefficient of treatment variable would be capturing just differential trends in population growth of districts, I explicitly control for the population of the district relative to state[21]. This effect is captured by the variable $Relativepopulation_{dt}$ In absence of any other data, I use the census 2001 population figures as proxy for population during the first round of the IHDS and census

[20]The IHDS 2004-05 reports income quintiles of households and categorizes them into seven caste-religious groups like Brahmins, OBCs, Adivasis, Muslims etc. The inclusion of interaction of baseline income groups with the Post variable as well as interaction of religion-caste group with the post variable ensures that the observed results are confounded by differential trends across income or religion groups that might be concentrated differently across districts

[21]This is simply population of district divided by population of state

2011 for the second round.

Finally, since there is reason to believe that politicians and governments behave favorably during election years which might affect reports of confidence, I include $Electionyear_t$ as a control which is equal to 1 if the respondent was interviewed in the year of state elections and 0 otherwise. I employ this specification separately for both confidence in politicians and confidence in state government as outcome variables.

2.5 Main Results

Table 2.5 reports the results from the main regressions. The first and the second columns show the results for the differences-in-difference specification without any controls or household fixed effects. The coefficient of $Post*Treat1$ is positive and significant for both confidence in politicians and the state government showing that there was an improvement in reported confidence for respondents from districts that gained representation. The coefficient of $Post*Treat2$ is negative but insignificant in case of both reported confidence for politicians state government suggesting no symmetric negative effects for respondents in districts that lost seats.

Columns 3 and 4 report the same results with household fixed effects. This does not change the sign or significance of our main coefficients of interest. Not only that, even the magnitude of the coefficients remains more or less the same. This is not surprising, as there is little reason to believe that household time invariant characteristics would be correlated to the treatment in question.

Columns 5 and 6 report the results of the regression with the full set of controls and household fixed effects. Again the coefficient of $Post * Treat1$ for the regression of confidence in politicians does not change much and remains significant, whereas that for confidence in state government slightly increases in magnitude and precision. The results suggest that the percentage of respondents reporting having confidence in politicians in districts that gained seats is about 8% points higher in post delimitation phase than in the pre-delimitation phase compared with respondents living in districts with unchanged representation. Similarly, the improvement for confidence in the state government is about 6.5% for districts that gained seats. Again there is no robust symmetric impact for respondents in districts that lost seats. Interestingly, the coefficient of *Elections* is

positive and significant at 0.01 % level for both the regressions, confirming the suspicion that reported political trust is higher during state election years. Overall, these results consistently suggest that political trust improves with an increase in representation. The next section strengthens these results by using falsification tests and robustness checks.

2.6 Falsification Tests and Robustness Checks

2.6.1 Falsification tests

This section discusses a range of robustness checks that were employed to establish greater confidence in the results discussed previously. One concern with the results discussed in the previous section is that they might be reflective of an environment of a generally increasing trust in institutions due to some other changes coinciding with the delimitation exercise. To assuage this concern, I use questions in the IHDS interviews that asked about confidence in other institutions like the military, police, newspapers/media, panchayats/nagar panchayats and courts[22]. I estimate the main specification (1) with the dependent variable being the aforementioned confidence reports. If these regressions render insignificant treatment coefficients, then the observed improvement in political trust is more credible since confidence in institutions like military or newspapers and media is unlikely to be affected by having more representatives in the state assembly.

The results of these falsification are shown in Table 2.6. The coefficients of $Post * Treat1$ & $Post * Treat2$ are insignificant for confidence in military, police, newspapers/media and courts. The coefficient of $Post * Treat2$ is significant and negative for confidence in panchayats suggesting some impact of losing representation and this merits further investigation. However, it does not mirror the observed results on political trust variables.

[22]The IHDS also includes questions on confidence in hospitals and schools. However the first round asked about confidence in hospitals and schools in general, whereas the second round separately asked for confidence in government and private schools and similarly for hospitals. This renders these measures incomparable across the two surveys, which is why they are not included in the falsification test. For another variable, confidence in bank, the necessity of transforming the report from three point scale to a dichotomous one does not leave usable variation across the years

2.6.2 Summary Index

To address the issue of multiple inference arising due to the use of multiple confidence measures in the previous section, this section employs the summary index approach discussed in Anderson (2008). I construct two separate summary indices– one combining the two confidence measures for political institutions and the second summary index combining all measures for confidence in other institutions used for the falsification tests. These summary indices are essentially weighted averages of standardized confidence variables constructed in a way that maximises the amount of information captured by the index. Using these indices instead of individual confidence measures not only mitigates the multiple inference problem by reducing the number of tests to be conducted, but also answers the question as to whether change in representation had an impact on political trust in general, while not having an overall general impact on non-political trust.

Table 2.7 shows the estimates from the difference-in-differences strategy employing these summary indices as dependent variables. Column 1 shows that there was a general improvement in political trust for people in districts that gained seats, but no change for people in districts that lost seats. Confirming the results in the previous section, column 2 shows that the change in representation had no impact on the index of non-political trust in either the gaining or losing districts.

2.6.3 Robustness to alternative specifications

Next, I modify the original specification and show that the observed treatment effects are robust to alternative specifications. First, I introduce state specific time trends to take into account state specific changes in confidence over time. These results are shown in column 1 of Table 2.8 for politicians and column 1 of Table 2.9 for state government. The treatment coefficient for districts that gained seats remains significant and roughly of the same magnitude for confidence in politicians. For confidence in state government, the treatment effect is slightly smaller and but still significant at the 10% level.

Secondly, I estimate the original specification (1) with added district level demographic controls which include log of district population, proportion of urban population, proportion of SC population, proportion of ST population and the proportion of literate population. These results are shown in column 2 of Table 2.8 & Table 2.9 for confi-

dence in politicians and confidence in state government respectively. The coefficient of $Post * Treat1$ remains roughly of the same magnitude in both the regressions and still significant at 10%. Controlling for the proportion of urban population especially helps address concerns that the results might be indicative of districts that grew at a faster rate in terms of population due to greater urbanization. The third columns in Table 2.8 & Table 2.9 add controls for both district level demographics as well as state specific time trends. The coefficient for $Post * Treat1$ remains significant for confidence in politicians, it is insignificant for confidence in state government (p *value*-0.13) but it should be emphasized that this specification includes a rather stringent set of controls.

Thirdly, I estimate the original specification (1) with a sample limited to districts that gained at most one seat or lost at most one seat. Column 4 of Table 2.8 & Table 2.9 present the results for these regressions. The coefficient of $Post * Treat1$ remains significant for the regressions for confidence in politicians and the state government for this limited sample, in fact they are slightly higher in magnitude. This quells concerns that the results might be driven by districts that gained a lot or lost a lot of seats.

Fourthly, instead of using a dichotomous treatment variable I use the $Changeinseats$ as the independent variable of interest in the original specification. The results are shown in Column 5 of Tables 2.8 & 2.9. The coefficient of $Changeinseats * Post$ is significant for regressions for both confidence in politicians and state government, so redefining the treatment as a continuous variable does not alter results. Lastly, introducing $Changeinreservedseats$ as an additional treatment does not alter the coefficient of $Changeinseats * Post$ suggesting that these results are not driven by changes in seats reserved for STs and SCs (Column 6 of Tables 2.8 & 2.9).

2.6.4 Randomization inference

Lastly, I use randomization inference as an alternative strategy to claim causality between representation and political trust. This exercise entails randomly reassigning the two treatments to districts in a state and then estimating a placebo treatment effect[23] [24]. Replicating this procedure multiple times provides a distribution of these placebo treatment effects which can then be used to arrive at an estimate of the probability of

[23] See Gupta and Spears (2017)
[24] See Heß et al. (2017) ritest for implementation in Stata

obtaining given results by chance alone. To do this, the treatments of gaining or losing seats are randomly reassigned to districts within a state while keeping the number of districts that gained and lost fixed. The placebo treatment effect is estimated by running the original regression of confidence in politicians (state government). A 1000 replications are carried out which yield a distribution of the simulated beta coefficients.

The kernel density of the distribution so obtained is shown in Figure 2.4 & Figure 2.5. Given this distribution, the percentage of beta coefficients at least as high as our original estimated coefficient gives an estimate of the probability of observing our results by chance alone. In this case, for the regression of confidence in politicians, about 78 simulated treatment effects exceeded our observed treatment effect. This means that the probability of obtaining a treatment effect at least as high as the observed one by chance alone is 0.078. Similarly, for confidence in state government the one sided p-value is 0.064, meaning that only 6.4 percent of the simulated beta coefficients were greater than the observed treatment effect.

2.7 Voter Turnout

One of the most important and widely discussed aspects of political trust for citizens' behaviour is their political participation. Two contradicting theories have been highlighted in literature in this respect. The first argues for a positive relationship between political trust and political participation such as turning out to vote. This is based on the hypothesis that positive perceptions or judgements of political systems are necessary to induce people to vote, as a result of which distrusting citizens are less motivated to cast a vote. The second competing theory claims that distrust could also be motivating factor in political participation as citizens feel compelled to intervene in a political system that they find untrustworthy (Marien and Hooghe, 2011).

Nevertheless, empirical evidence has found some support for the former theory by showing a positive association between political trust and voting behavior (Marien and Hooghe, 2011). For instance, Grönlund and Setälä (2007) use the European Social Survey and establish that certain types of political trust– trust in politicians and parliament is associated with a higher propensity to vote. Here I provide some evidence in support of this claim.

Drawing from the literature linking political trust to political participation, I provide evidence that improvement in political trust is accompanied by an increase in voter turnout. This also serves as an additional robustness check because if the positive association between political trust and participation is true and if as we have observed, political trust improves with increase in representation, then increase in representation should also be accompanied by more voter participation.

Further, looking at changes in voter turnout also allows for the construction of false breaks in the pre-delimitation era as another robustness check. It is not possible to implement this exercise with the confidence variables as the IHDS data is only available for two rounds- 2004-05 & 2011-12. Using the district level voter turnout[25] as the dependent variable, I implement a difference-in-differences strategy with the districts that gained seats constituting one treatment group and the districts that lost as another treatment group. Unchanged districts serve as the control group. Analogous to the relative population variable used in earlier regressions, I explicitly control for the proportion of district electors to the state electors in these regressions.

The results of this exercise are shown in Table 2.10. Column 1 shows the aforementioned specification run for the period of 2000-2018. This specification also controls for district fixed effects, election year fixed effects and state-year fixed effects. The results from the turnout regressions exactly mirror the results from the confidence regressions. The coefficient of $Post * Treat1$ is positive and significant while that of $Post * Treat2$ is insignificant. These results show that on an average, voter turnout increased by about 1.5 percentage points in districts that gained seats as compared to districts that remained unchanged in the post delimitation period. At the same time there is no negative impact of losing seats on voter turnout, just like the results for the confidence outcome.

Next, I create a false break in the year 2003 and look at the district level voter turnout for the period of 1998-2008. Column 3 of Table 2.9 shows the result for this regression. Coefficients of both $Post * Treat1$ $Post * Treat2$ are insignificant. Since three new states of Uttarakhand, Jharkhand and Chhattisgarh were created in 2000 and people in these states voted for governments of different states before and after 2000, I exclude these states[26] from this regression. Column 2 of Table 2.10 uses this limited sample for original specification in order to enable a comparison with Column 3.

[25]Source: Lokdhaba website maintained by TCPD, Ashoka University

[26]Jharkhand is automatically dropped as it did not undergo delimitation

To further establish the robustness of the voter turnout result, Column 4 shows results from difference-in-differences specification with a false break created in 1998 and elections spanning 1988 to 2008, incorporating a ten year window (with two state elections) before and after the false break. The coefficient of $Post * Treat1$ is actually negative and marginally significant for this specification. This result provides confidence that the turnout results shown in the first column are not driven by any pre-existing trends which would have been captured by this false break regression. This also lends credibility to the original results for confidence, because even though the data is unavailable to test for pre-existing trends in political trust- an associated phenomenon,i.e., turnout exhibits no such worrisome patterns.

While providing evidence in support of a link between political trust and voter turnout, this result also appeals to the literature that calls for a direct relationship between electorate size and voter turnout (Hansen et al., 1987; Oliver, 2000; Levine and Palfrey, 2007). For instance, Geys (2006) through meta-analysis of 83 aggregate-level studies, concludes that there is a negative relationship between electorate size and voter turnout and hypothesizes that this on account of the negative association between electorate size and probability that one single voter will make a difference to the outcome, drawing from the Downsian voting model.

Increase in voter turnout can also be a consequence of increased voter mobilization on account of reduction in constituency size. Numerous studies have suggested a negative relationship between voter mobilization and electorate size. Shachar and Nalebuff (1999), for instance, use state level data for US Presidential elections and show that both voter turnout and parties' efforts are higher for more closely contested elections and when the voting population is lower. Mori (2015) uses data for national elections in India from 1977-2007 and provides evidence for negative relationship between constituency size and voter turnout and argues that this due to national political parties directing efforts to gain vote share in smaller constituencies. While the paucity of data does not allow to test for the mobilization hypothesis directly, given the fact that more personal mobilization stimulates turnout (Gerber and Green, 2000), the improvement in voter turnout does provide suggestive evidence in its favour.

2.8 Possible Channels

2.8.1 Public goods and rural development

This section explores possible mechanisms through which increase in representation could have an impact on political trust. The first hypothesis hinges on the evidence pointing to a negative relationship between representation and distribution of economic resources. Ansolabehere et al. (2002) for the United States and Horiuchi and Saito (2003) for Japan show that fiscal transfers are skewed in favour of over-represented regions. Knight (2008) documents the same effect for representation of states in the US Senate and proposes two channels through which the small state advantage works– the first being that increased representation means more chances of being represented by the proposer in committees and the second being the voter cost channel which suggests that increase in representation make smaller constituencies more attractive from the perspective of a proposer looking to form the cheapest possible coalition.

Unfortunately, data for state transfers to districts are not available in the Indian context to directly test for this mechanism. However, if it was true that increased representation led to redistribution of resources in favor of previously underrepresented constituencies, one would expect to find an improvement in development outcomes for the districts that gained seats. To check for this mechanism, I use the village amenities from census 2001 & 2011. These contain village level availability of public goods and infrastructure like schools, health centres, hospitals, electricity, transport, access by paved roads, post offices, drinking water etc. By linking the village level data to the district level constituency changes, one can check if increase in representation is associated with improvement in provision of public goods and infrastructure which in turn is leading to improvement in political trust following from the theory of institutional performance.

I check for this channel in two different ways. First I use principal component analysis to combine village level binary indicators for availability of 18 public goods including schools, health facilities, transport, communication, banking, electricity and drinking water facilities[27]. The indicators for drinking water and electricity take value 1 if there

[27]These include availability of primary school, middle school, secondary school, senior secondary school, drinking water, bus facility, railway station, primary health centre, primary health sub centre, dispensary, telephone, post office, post telegraph office, approach by paved road, electricity, commercial bank and agricultural credit society

was any source of drinking water and electricity in the village for any purpose respectively. Then I take the population weighted average of this village level index to arrive at a district level development index.

I run the original difference-in-differences regression with this district level development index as a dependent variable and check for the coefficients of $Year2011 * Treat1$ & $Year2011 * Treat2$ [28]. The results for this regression are shown in column 1 of Table 2.11. The estimates show that the districts that gained seats show an improvement in the development index by 0.1139. This is against a mean of 1.441 in 2001. Interestingly, the coefficient for the districts that lost seats is insignificant mirroring the results for confidence in political institutions.

Next, I use the village amenities from the 1991 census to check if similar results are obtained for the treatment groups in the 1991-2001 period when there was no delimitation. Since the 1991 census did not collect information about availability of banks, I use indicators for 15 non-banking public goods in this case. To generate a comparable counterpart for this, I carry out the same exercise as before omitting banking variables for 2001-2011. These results are shown in the second column 2 of Table 2.11. The coefficient of $Year2011 * Treat1$ remains similar in magnitude to column 1. Column 3 shows results from regression over 1991-2001. The sample size is smaller in this regression as it was not possible to match all villages over the two censuses. The coefficient of $Year2001 * Treat1$ is insignificant lending credibility to results in Column 1 by ruling out pre-existing trends.

Next, I construct a measure of change in representation at the village level. The problem with looking at district level outcomes is that it does not allow us to precisely measure the change in representation at a granular level and as result what we observe is in fact an average of outcomes in some sense. For instance, even though a given district gains seats the actual increase in representation for a locality/village would in fact depend on how populations are allocated across these higher number of constituencies meaning that some regions within a district are more likely to gain than others.

Since assembly constituencies do not remain the same after delimitation, the way to more precisely measure changes in representation is to look at the village level. The IHDS does not permit analysis at a finer level as households can not be mapped to administrative units smaller than the district. But census data allows us to test for this

[28] $Post$ in this case represents 2011 census round

mechanism at the village level. For this purpose, I consider villages that remained the same across the 2001 and 2011 censuses.

I construct a Relative Representation Index (RRI)[29] for a village by dividing the number of representatives per elector in the constituency where the village is situated by the number of representatives per elector in the state. Since any given village lies inside a unique assembly constituency, the number of representatives in the numerator is one. Also, since the total size of the state legislative assembly was held constant through delimitation, the number of seats in the denominator remains the same. I use the number of electors in the election immediately before and immediately after delimitation for the purpose of this calculation. An RRI greater than 1 indicates over-representation relative to state and RRI lower than 1 indicates under-representation. A positive change in the RRI indicates increase in representation while a negative change in the RRI is indicative of a decrease in representation.

Column 2 of Table 2.12 shows the changes in RRI for the two treatment groups relative to the unchanged. As expected, villages in districts that gained seats show an increase in RRI relative to districts that were unchanged and villages in district that lost seats show a decrease in the RRI. Next I use the change in the RRI as the main independent variable to look at the changes in the village level development index as constructed previously, controlling for changes in village population and population of district relative to state. Column 1 of Table 2.12 shows results from this regression. The coefficient of change in RRI is positive and significant corroborating the results from the district level development regression.

2.8.2 Winner and candidate quality

The second channel proposes that an improvement in politician characteristics such as age, quality as measured by criminal cases or education led to an increase in political trust in the gaining districts. This would be true if increase in representation encouraged political parties to assign better quality candidates to the gaining districts as they would be worth more seats in the state legislature after the delimitation exercise. This could also happen if allotment of more seats facilitated the entry of newer candidates who could

[29]This is a commonly used measure of representation, see Ansolabehere et al. (2002) and Bhavnani (2018)

not enter otherwise because of the dominance of existing seasoned politicians.

I find mixed evidence in support of this channel using candidate data sourced from affidavits filed at time of nomination. To illustrate this, I focus on two sets of analysis– one with all the candidates that contested and the second including just the winners. Since candidate data is only available from 2004 onwards when the Supreme Court made filing of these affidavits mandatory, I exclude states for which pre-delimitation candidate data remains unavailable (these would be states that had their first election after Supreme Court judgement after delimitation).

Tables 2.13 & 2.14 show the main results from the original specification with year and state year fixed effects using various candidate(winner) characteristics as the dependent variable of interest. Analogous to the relative population variable in the original regression, the proportion of district electors to state electors is used to control for trends in population. Table 2.13 shows that not only are candidates younger by about 1.2 years in gaining districts relative to no change districts in the period after delimitation, but they also have lower number of criminal cases. Table 2.14 shows that even winners are about 1.6 years younger in gaining districts compared with unchanged districts, although the same is not true for criminal cases. There are no significant differences in assets or education for both candidates and winners. Interestingly, both these sets of results show no negative effects for the districts losing representation, mirroring the results for political trust.

While an improvement in candidate characteristics in terms of criminal cases and age can explain the increased trust in politicians in general, it is less likely to be driving the results for increased trust in the state government which alludes to the actual winners. Secondly, better candidate and winner quality would only manifest after an election is held under the new boundaries and as discussed in the next section, the observed improvement in political trust is not contingent on an election being held after delimitation. This compels us to fall back on the idea of improved development as the more credible channel behind the increased political trust.

2.8.3 Improved responsiveness and mobilization

This leaves us with the third possibility of improvement in responsiveness, interaction or improved voter mobilization due to reduction in constituency size as possible drivers of

the results. The hypothesis of improved responsiveness and interaction is based on the idea that smaller constituencies (as a result of gaining seats) make it easier for politicians to access and to respond to needs of a more homogeneous electorate. At the same time, a smaller population finds it easier to hold politicians accountable. For instance, Hansen (2013) uses municipal mergers in Denmark as a quasi experiment to establish the negative impact of polity size on political trust and attributes this partly to decline in perceived responsiveness of officeholders.

Denters (2002) using survey data from Denmark, the Netherlands, Norway and the United Kingdom, documents a negative relationship between political trust and population size and suggests that this is partly on account of greater satisfaction with government responsiveness and attentiveness in less populous units. He also documents higher trust in local officeholders as compared with national officeholders. This is notionally similar to the arguments cited in economics literature in favour of decentralisation based on improved ability to respond effectively to people's preferences (Bardhan, 2002; Bardhan and Mookherjee, 2000). Improved responsiveness and interaction could also be a product of the fact that the districts gaining representation become more valuable from the perspective of votes, for they now entail paying attention to a smaller population for a legislative seat.

While the IHDS does not have data on perceived or actual responsiveness of officeholders, it is not hard to imagine that this could be one of the potential reasons for improved political trust. A post poll survey conducted by Lokniti [30] in the state of Karnataka[31] in 2008 posed the question to respondents to Bangalore region (which ended up gaining 12 additional seats after delimitation)- " *Will the increase of the number of assembly seats in Bangalore city make any difference*?" to which 51.3% respondents replied that it will make a difference as against 32.9% stating that it will make no difference. With respect to the expected change, 38.1% respondents agreed that Bangalore would be given greater importance by the government, 13.1% believed that MLAs will be more accessible as they would be greater in number and 18.8% respondents agreed that their issues will be raised

[30]The Lokniti Programme for Comparative Democracy is a research programme of the Centre for the Study of Developing Societies (CSDS), Delhi. They regularly carry out election studies for both national and state elections in India

[31]Karnataka was the first state to undergo elections under new boundaries in India in May 2008

more in the assembly because of larger number of MLAs [32][33].

Similar arguments would suggest an increase in voter mobilization for the gaining districts, particularly the kind that would involve direct and more personal contact. Mori (2015) used national elections data for India reported a positive relationship between constituency size and turnout and suggested that this is because national parties concentrate their efforts in smaller constituencies as they provide more bang for the buck. While there is no district level data for party expenditure in case of state elections in India, the enhanced voter turnout in gaining districts does provide suggestive evidence in support of this mechanism.

2.9 Timing of elections

A natural question that arises as a consequence of the observed increase in political trust is whether this effect should manifest only after the conduct of an election after delimitation when the new boundaries actually come into effect or if the drivers of this improvement emerge at the very announcement of the new boundaries. To understand this, one must keep in mind the non synchronised nature of state elections in India.

When the recommendations of the Delimitation Commission were approved by the president in February 2008, the states of Meghalaya, Nagaland and Tripura had already undergone elections under old boundaries in early 2008. Karnataka, Chhattisgarh, Madhya Pradesh, Delhi, Mizoram and Rajasthan were scheduled to under go elections later that year while Sikkim, Maharashtra, Andhra Pradesh and Odisha would have had elections under new boundaries in 2009 and so on.

Considering that state elections are held every five years and that the second round of the IHDS surveys was held in 2011-12, we would have states that had not undergone elections under the newly delimited boundaries at the time of the second survey in our sample. These would have been states that had last held elections under old boundaries in 2007 or early 2008 and therefore had the next round scheduled for 2012 or 2013. These states were Uttar Pradesh, Uttarakhand, Meghalaya, Tripura, Mizoram and Gujarat[34].

[32]Source is the post poll survey report for the state election study for Karnataka in 2008 carried out by Lokniti

[33]Second question posed to the 51.3% respondents who said it will make a difference in answer to the first question

[34]While Punjab had its first election under new boundaries in January 2012, the interviews for IHDS II were held in the same year form February until September

While one might feel inclined to argue that the increase in confidence should only be observed after the conduct of elections under new boundaries, it is necessary to consider the underlying mechanism that is proposed as the driver. If improvement in performance of politicians and the state government towards the people in gaining districts through better provision of public goods and village amenities leads to greater confidence, the question to ask is why would political actors wait until the next election under new boundaries to act. The information about the new boundaries of electoral constituencies was made available to all in 2008 when the delimitation order was passed by the President, irrespective of when a state would have the next election.

Assuming that politicians and MLAs seek re-election and targeting public goods to gaining districts makes more sense now that they are *worth more* seats, the improvement in their performance should be immediate and not contingent on holding an election with new boundaries. To test for this, I create another indicator $Noelection$ that equals 1 if respondent belonged to a state[35] that at the time of the second survey had not held an election under the newly delimited boundaries. I then ran the original specification including an interaction of $Noelection$ with $Post * Treat1$ and $Post * Treat2$. This is in fact a triple differences estimation exercise to check whether the impact of change in representation varied according to the state election timing.

The results from this regression are shown in Table 2.15. Column 1 shows the triple differences exercise for confidence in politicians and column 2 for confidence in state government. As expected from the previously proposed hypothesis, the coefficients of the triple difference $Post * Treat1 * Noelection$ is insignificant for both dependent variables, suggesting no variation in impact of gaining seats based on whether states had an election under new boundaries.

2.10 Conclusion

Using a nationally representative panel dataset with self reported political trust and a redistricting exercise that led to change in representation of districts in the state legislatures, this paper shows that an increase in representation improves political trust for rural India. This observed effect is robust to alternative specifications and is not driven

[35]Uttar Pradesh, Uttarakhand, Meghalaya, Tripura, Mizoram and Gujarat

by districts that either gained or lost a lot of seats. Moreover, this improvement in political trust is not accompanied by an increase in trust in other institutions that are likely to be unaffected by the increase in representation. It contributes to the understanding of the notion of political trust by tying it up with an important aspect of institutional behavior,i.e., the distribution of representation.

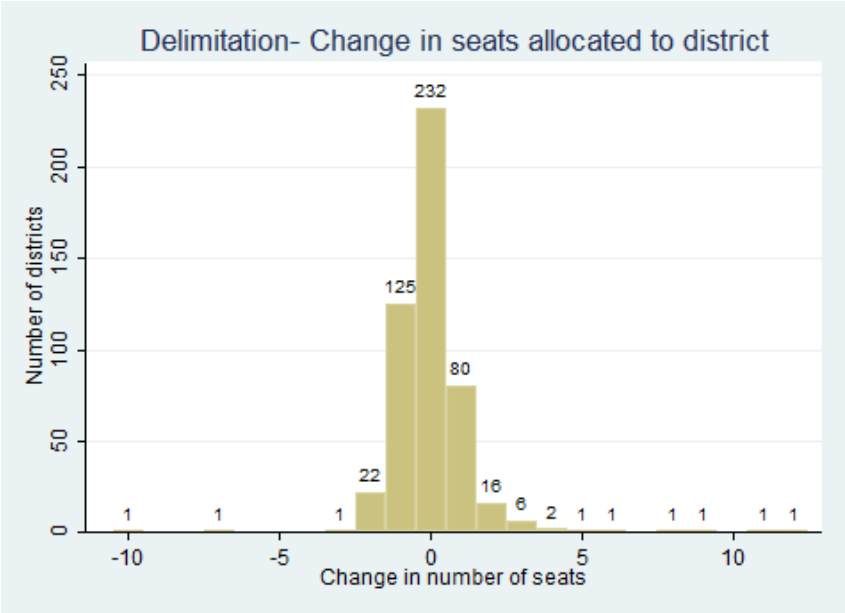
The observed improvement in political trust is concomitant with an increase in voter turnout in districts that gained seats. While providing evidence for political participation as a correlate of political trust, this result also provides suggestive evidence for the link between representation and voter mobilization. Improvement in political trust stimulates people to vote and this could be mediated by improvement in the behaviour of political actors through better delivery of public goods as shown and better responsiveness and mobilization. While the lack of previous data does not allow for a check of parallel trends, an analysis of the voter turnout confirm parallel trends for gaining and unchanged districts in the pre-delimitation period, thereby adding to the robustness of the original results for political trust.

Thirdly, the increased trust is accompanied by an improvement in rural development as measured by index of availability of public goods. This result on one hand sits right in with the literature documenting distributive consequences of representation. On the other hand, it also provides evidence in support of the theory of institutional performance for origins of political trust that argues that trust moves in response to how institutions behave towards them. In this sense, while more representation enables gaining districts to get more attention from political actors and the consequent improvement in public goods delivery leads to an increase in political trust.

This finding is crucial as it contributes to our understanding of the importance of representation on one hand and that of the nature of political trust on the other by tying them up together. The consequences of representation for distribution remain well acknowledged in literature. This paper additionally highlights another channel through which representation can have even wider ranging ramifications than previously thought, through its impact on level of political trust. At the same time, it sheds light on the role of political trust in stimulating political participation which can further enhance accountability for political representatives.

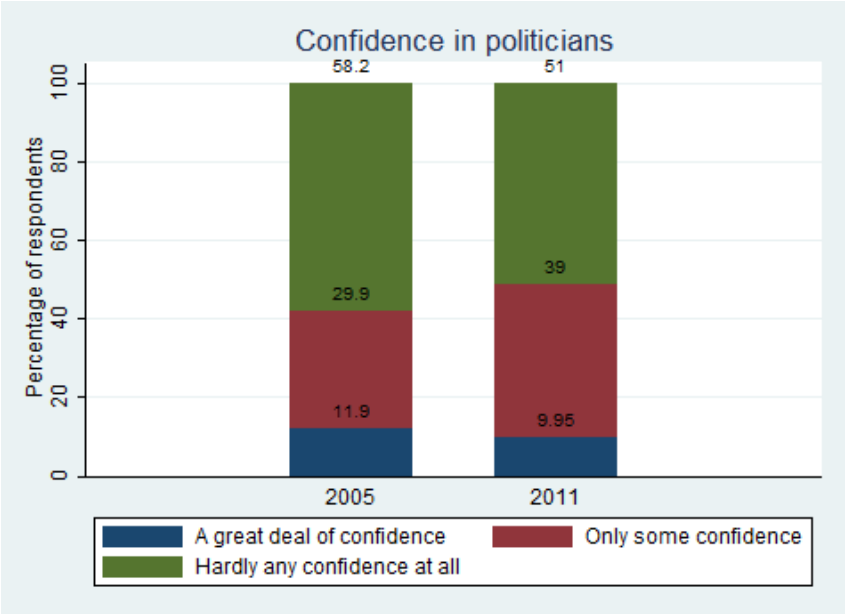
Figures and Tables for Chapter 2

Figure 2.1: Distribution of change in seats due to delimitation



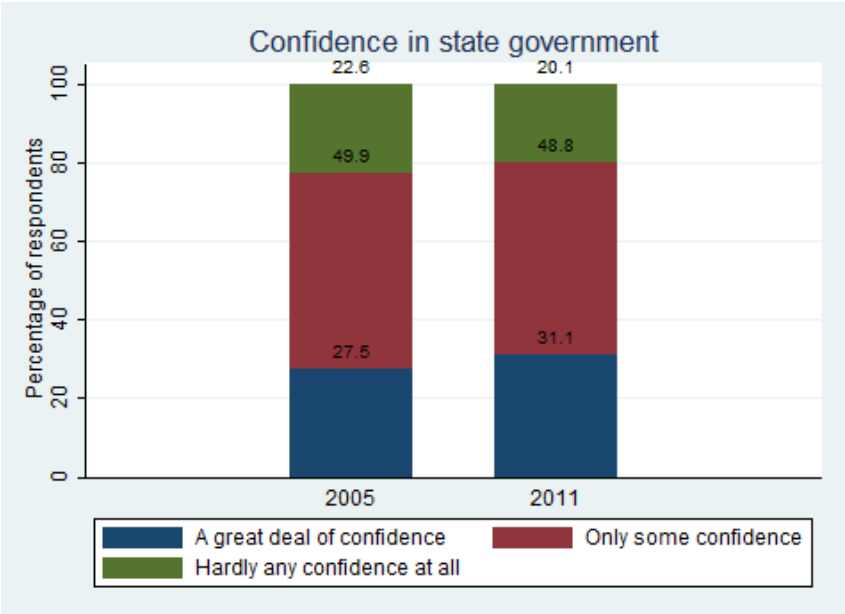
Source: Papers of the Delimitation Commission

Figure 2.2: Confidence in politicians



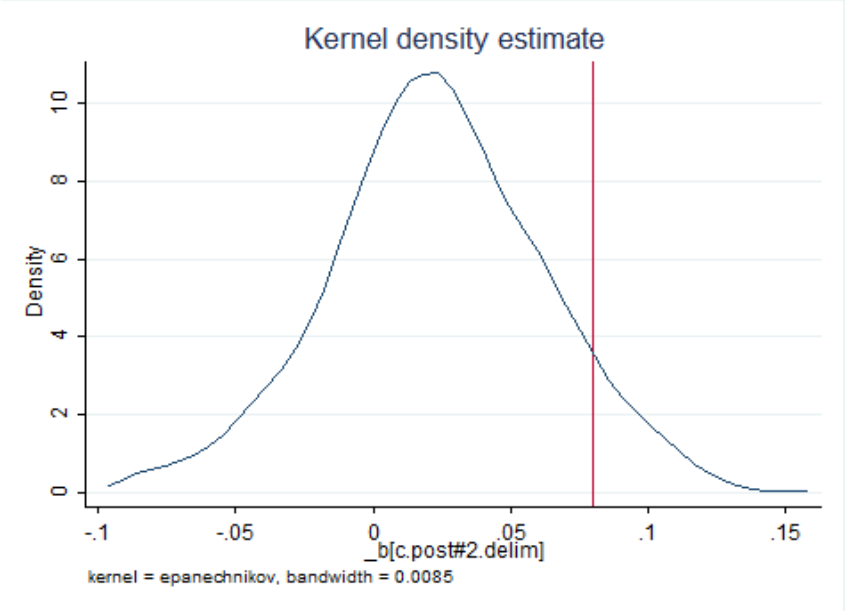
Source: IHDS I & II

Figure 2.3: Confidence in state government



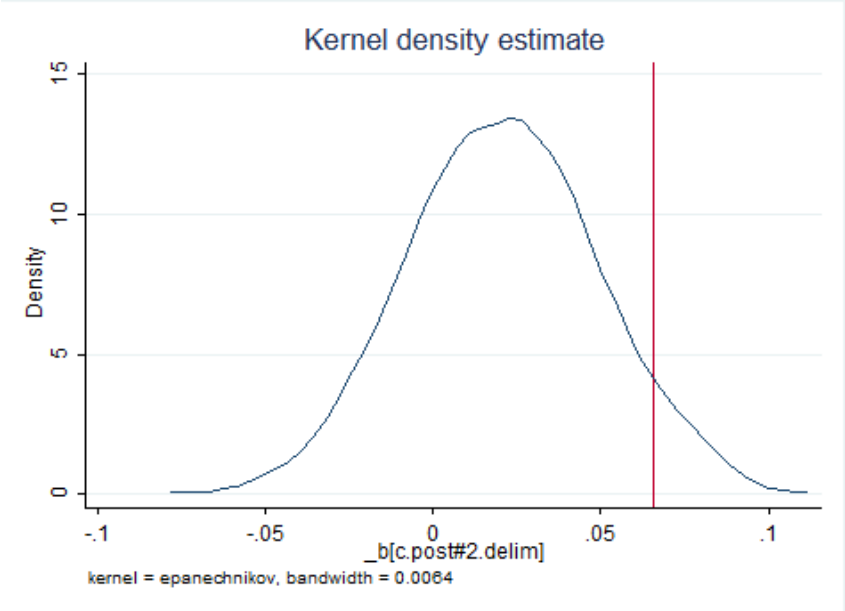
Source: IHDS I & II

Figure 2.4: Randomization exercise-Confidence in Politicians



Distribution of simulated coefficients of $Post*Treat1$ in regression of confidence in politicians. One sided p-value is 0.078

Figure 2.5: Randomization exercise-Confidence in State Government



Distribution of simulated coefficients of $Post*Treat1$ in regression of confidence in state government. One sided p-value is 0.064

Table 2.1: Baseline respondent characteristics by treatment status

	Gained seats	Lost seats	No change in seats	Total
Confidence in politicians	0.394	0.445	0.407	0.417
	(0.489)	(0.497)	(0.491)	(0.493)
Confidence in state govt	0.749	0.781	0.781	0.774
	(0.433)	(0.414)	(0.413)	(0.418)
Age	42.19	43.97	43.12	43.21
	(14.34)	(14.40)	(14.36)	(14.38)
Sex (male=1)	0.754	0.754	0.782	0.766
	(0.431)	(0.430)	(0.413)	(0.423)
Years of education	4.107	4.669	4.518	4.478
	(4.443)	(4.556)	(4.583)	(4.547)
Whether married	0.871	0.847	0.853	0.855
	(0.335)	(0.360)	(0.354)	(0.352)
Household size	6.173	5.701	6.219	6.029
	(3.447)	(2.963)	(3.168)	(3.173)
Number of children	2.166	1.775	2.171	2.033
	(1.970)	(1.678)	(1.902)	(1.853)
Log of per capita consumption	6.352	6.384	6.339	6.357
	(0.620)	(0.641)	(0.684)	(0.656)
Agricultural land owned(in acres)	2.274	2.418	2.742	2.524
	(6.035)	(5.030)	(6.074)	(5.726)
Observations	5813	9021	11128	25962

Mean coefficients; sd in parantheses. Source- IHDS 2004-2005. Sample includes respondents belonging to rural areas in states that underwent delimitation

Table 2.2: Baseline respondent differences by treatment status

	No change-Gained	No change-Lost
Confidence in politicians	0.0134*	-0.0375***
	(0.00796)	(0.00702)
Confidence in state govt	0.0319***	0.000423
	(0.00687)	(0.00590)
Age	0.928***	-0.846***
	(0.232)	(0.204)
Sex (male=1)	0.0274***	0.0273***
	(0.00678)	(0.00596)
Years of education	0.411***	-0.151**
	(0.0736)	(0.0649)
Whether married	-0.0180***	0.00629
	(0.00563)	(0.00506)
Household size	0.0451	0.517***
	(0.0529)	(0.0436)
Number of children	0.00519	0.396***
	(0.0312)	(0.0256)
Log of per capita consumption	-0.0127	-0.0450***
	(0.0107)	(0.00943)
Agricultural land owned(in acres)	0.468***	0.325***
	(0.0985)	(0.0799)
Observations	16941	20149

Difference in means; standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ Source- IHDS 2004-2005. First column shows results from t-tests for difference in characteristics of respondents in districts that gained seats against districts that didn't gain or lose. Second column shows results from t-test between respondents from districts that lost seats against unchanged districts.

Table 2.3: Baseline district characteristics by treatment status

	Gained seats	Lost seats	No change in seats	Total
Panel A: Demographics				
Population (Ten thousands)	210.8	184.3	163.7	181.0
	(183.5)	(109.6)	(99.26)	(128.4)
Proportion of urban population	0.307	0.212	0.214	0.235
	(0.214)	(0.116)	(0.184)	(0.180)
Population of district relative to state	0.0935	0.0429	0.0446	0.0558
	(0.180)	(0.0422)	(0.0693)	(0.105)
Proportion of SC population	0.146	0.165	0.167	0.161
	(0.0892)	(0.0749)	(0.0806)	(0.0815)
Proportion of ST population	0.178	0.100	0.130	0.133
	(0.284)	(0.193)	(0.224)	(0.233)
Average household size	5.372	5.185	5.527	5.391
	(0.670)	(0.721)	(0.839)	(0.780)
Proportion of literate population	0.647	0.670	0.633	0.647
	(0.132)	(0.113)	(0.130)	(0.127)
Population weighted development index	1.441	2.111	1.313	1.572
	(1.789)	(1.845)	(1.599)	(1.749)
Observations	116	138	228	482
Panel B: Election				
District voter turnout	0.633	0.654	0.630	0.638
	(0.107)	(0.0933)	(0.103)	(0.102)
Average no. of candidates per constituency	8.979	7.918	9.025	8.680
	(3.630)	(3.075)	(3.988)	(3.685)
District electors/State electors	0.0610	0.0450	0.0434	0.0478
	(0.0835)	(0.0450)	(0.0719)	(0.0683)
Observations	222	304	479	1005
Panel C: Representation index				
Relative representation index 2001	0.989	1.133	1.032	1.052
	(0.184)	(0.172)	(0.137)	(0.168)
Observations	74785	98340	164656	337781
Panel D: Candidate characteristics				
Candidate age	46.83	46.72	44.36	45.57
	(11.99)	(11.69)	(11.59)	(11.77)
No. of criminal cases	0.628	0.520	0.553	0.561
	(1.807)	(1.442)	(1.957)	(1.798)
Years of education	12.18	12.67	11.88	12.16
	(4.608)	(4.290)	(4.796)	(4.635)
Total assets (in lakhs)	83.30	92.13	64.70	76.37

Continued on next page

Table 2.3 – *Continued from previous page*

	Gained seats	Lost seats	No change in seats	Total
	(544.4)	(558.5)	(589.8)	(571.3)
Observations	2339	2685	5017	10041
Panel E: Winner characteristics				
Winner age	50.43	49.55	47.86	49.15
	(10.69)	(10.51)	(10.63)	(10.66)
No. of criminal cases	0.931	0.817	1.057	0.941
	(2.198)	(1.827)	(2.876)	(2.376)
Years of education	13.36	13.40	13.28	13.34
	(3.735)	(3.749)	(3.775)	(3.754)
Total assets (in lakhs)	109.1	123.0	115.1	116.1
	(284.7)	(363.9)	(396.2)	(356.9)
Observations	662	799	905	2366

Mean coefficients, sd in parentheses. Panel A shows baseline summary statistics of district level demographics by treatment status (Source: Census 2001). Panel B shows treatment wise baseline turnout and other electoral variables at district level using election data from period 1998 to 2008 (Source: Election data from Lokdhaba website). Panel C gives treatment wise representation index calculated at village level. Panel D shows pre-delimitation summary statistics for candidates characteristics in state legislative assembly elections. Panel E gives pre-delimitation summary statistics for winners in state legislative assembly elections. (Source: candidate affidavits data from website of Association of Democratic Reforms (ADR)). Data used from elections spanning 2004 to 2008 included. In all cases, sample includes districts in all states that underwent delimitation not limited to districts covered in the IHDS sample.

Table 2.4: Baseline district differences by treatment status

	No change-Gained	No change-Lost
Panel A: Demographics		
Population (Ten thousands)	-78.58***	-20.57*
	(15.41)	(11.14)
Proportion of urban population	-0.0975***	0.00209
	(0.0232)	(0.0175)
Population of district relative to state	-0.0173*	0.0017
	(0.0090)	(0.0065)
Proportion of SC population	0.000803	0.00156
	(0.0096)	(0.0085)
Proportion of ST population	0.0182	0.0300
	(0.0266)	(0.0229)
Average household size	0.0984	0.342***
	(0.0954)	(0.0859)
Proportion of literate population	-0.0196	-0.0372***
	(0.0156)	(0.0134)
Population weighted development index	-0.213	-0.798***
	(0.195)	(0.183)
Observations	328	366
Panel B: Election		
District voter turnout	-0.0029	-0.0238***
	(0.0085)	(0.0073)
Average no. of candidates per constituency	0.0462	1.107***
	(0.315)	(0.268)
District electors/State electors	-0.0176***	-0.0016
	(0.0062)	(0.0046)
Observations	701	783
Panel C: Representation Index		
Relative representation index 2001	0.0429***	-0.102***
	(0.0007)	(0.0006)
Observations	239441	262996
Panel D: Candidate characteristics		
Candidate age	-2.470***	-2.362***
	(0.293)	(0.278)
No. of criminal cases	-0.0760	0.0330
	(0.0478)	(0.0429)
Years of education	-0.301**	-0.791***
	(0.123)	(0.115)
Total assets (in lakhs)	-18.60	-27.42**

Continued on next page

Table 2.4 – *Continued from previous page*

	No change-Gained	No change-Lost
	(14.41)	(13.85)
Observations	7356	7702
Panel E: Winner characteristics		
Winner age	-2.571***	-1.689***
	(0.545)	(0.513)
No. of criminal cases	0.127	0.240**
	(0.134)	(0.118)
Years of education	-0.0717	-0.118
	(0.197)	(0.187)
Total assets (in lakhs)	6.009	-7.816
	(18.07)	(18.51)
Observations	1567	1704

Columns show results from t-tests for difference in baseline means according to treatment status. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Panel A shows difference in means for district devel demographics by treatment group (Source: Census 2001). Panel B shows difference in means for district voter turnout and other election variables using elections from 1998 to 2008 (Source: Election data from Lokdhaba webite). Panel C gives t-tests for representation index calculated for villages. Panel D shows difference in baseline means for candidates in state assembly elections by treatment group. (Source: candidate affidavits data from website of Association of Democratic Reforms (ADR)). Panel E shows difference in baseline means for winners in state assembly elections. Data used from assembly elections from 2004 to 2008. Sample includes districts in all states that underwent delimitation not limited to districts covered by IHDS.

Table 2.5: Main Results

	(1) Confidence in politicians	(2) Confidence in state govt	(3) Confidence in politicians	(4) Confidence in state govt	(5) Confidence in politicians	(6) Confidence in state govt
Post	0.0615**	0.0231	0.0622**	0.0232	0.1007**	0.0848**
	(0.0270)	(0.0186)	(0.0269)	(0.0185)	(0.0489)	(0.0398)
Treat1	-0.0134	-0.0319				
	(0.0317)	(0.0292)				
Treat2	0.0375	-0.0004				
	(0.0337)	(0.0270)				
Post*Treat1	0.0809**	0.0560*	0.0799*	0.0558*	0.0811**	0.0655**
	(0.0408)	(0.0311)	(0.0407)	(0.0311)	(0.0409)	(0.0303)
Post*Treat2	-0.0212	-0.0316	-0.0214	-0.0297	-0.0062	-0.0217
	(0.0426)	(0.0347)	(0.0427)	(0.0347)	(0.0422)	(0.0324)
Relative population					2.4548	-0.9052
					(2.0972)	(1.5233)
Election year					0.0709***	0.0750***
					(0.0256)	(0.0245)
Observations	51697	51403	51697	51403	50976	50683
Household fixed effects	No	No	Yes	Yes	Yes	Yes
Household level controls	No	No	No	No	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression. The first and second columns show results from the regression without household fixed effects or controls. The third and fourth columns add household fixed effects. The fifth and sixth columns show results from the main specification with both household fixed effects and controls. Controls include age, gender, marital status, years of education, household size, number of children, extent of agricultural land owned and main source of income. We also include differential trends by religion and caste groups and baseline income quintiles as part of our controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Falsification tests

	(1) Confidence in military	(2) Confidence in police	(3) Confidence in newspapers/media	(4) Confidence in panchayats/nagar palika	(5) Confidence in courts
Post	-0.0108	0.1704***	0.0259	0.0710*	0.0514*
	(0.0096)	(0.0382)	(0.0276)	(0.0395)	(0.0285)
Post*Treat1	-0.0036	0.0345	-0.0176	-0.0028	0.0111
	(0.0095)	(0.0320)	(0.0203)	(0.0242)	(0.0202)
Post*Treat2	0.0014	-0.0239	-0.0043	-0.0711***	-0.0257
	(0.0072)	(0.0312)	(0.0141)	(0.0238)	(0.0224)
Relative population	-1.2677*	-0.5505	-1.1214	0.0417	-1.6742*
	(0.6473)	(1.3700)	(0.7870)	(1.1075)	(0.8982)
Election year	0.0098**	0.0884***	0.0196	-0.0001	0.0529**
	(0.0045)	(0.0275)	(0.0120)	(0.0166)	(0.0216)
Observations	50744	50912	48679	50827	49545
Household fixed effects	Yes	Yes	Yes	Yes	Yes
Household level controls	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes results from a regression with a different dependent variable. Controls include age, gender, marital status, years of education, number of children, extent of agricultural land owned and main source of income. We also include differential trends by religion and caste groups and baseline income quintiles as part of our controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Summary Index

	(1) Political trust index	(2) Non political trust index
Post	0.2472**	0.2829***
	(0.1007)	(0.1005)
Post*Treat1	0.2246***	0.0184
	(0.0837)	(0.0796)
Post*Treat2	-0.0562	-0.1126
	(0.0896)	(0.0784)
Relative population	3.0057	-7.1329
	(4.7793)	(4.4979)
Election year	0.1823***	0.1914***
	(0.0559)	(0.0636)
Observations	47883	47883
Household fixed effects	Yes	Yes
Household level controls	Yes	Yes

Table 2.8: Robustness checks- Confidence in politicians

	(1) State specific time trends	(2) District specific controls	(3) District controls with trends	(4) Change in seats ≤ 1	(5) Changes in seats	(6) Changes in reserved seats
Post	0.1206	0.0674	0.0888	0.1159**	0.1165**	0.1197**
	(0.0817)	(0.0745)	(0.0958)	(0.0502)	(0.0468)	(0.0472)
Post*Treat1	0.0772*	0.0771*	0.0688*	0.0935**		
	(0.0397)	(0.0408)	(0.0396)	(0.0447)		
Post*Treat2	0.0307	0.0115	0.0361	0.0031		
	(0.0379)	(0.0446)	(0.0387)	(0.0431)		
Relative population	-1.5983	2.9938	-2.6765	2.3604	2.3733	2.4277
	(3.5404)	(2.1561)	(6.4842)	(2.3015)	(2.0864)	(2.0856)
Election year	-0.0030	0.0735***	0.0010	0.0779***	0.0682***	0.0671**
	(0.0327)	(0.0261)	(0.0315)	(0.0274)	(0.0261)	(0.0259)
Change in seats*Post					0.0302*	0.0323*
					(0.0160)	(0.0167)
Change in reserved seats*Post						-0.0087
						(0.0154)
Observations	50976	50976	50976	45538	50976	50976
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household level controls	Yes	Yes	Yes	Yes	Yes	Yes
State specific time trends	Yes	No	Yes	No	No	No
District specific controls	No	Yes	Yes	No	No	No

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression with the same dependent variable- confidence in politicians. The first column shows results by adding state specific time trends to the main regression. The second column introduces the log of district population, proportion of urban population, proportion of SC population, proportion of ST population and proportion of literate population in district as additional controls. The third column shows results with both district controls and state time trends. The fourth column shows result with sample limited to districts that at most gained or lost one seat. The fifth column shows results from regression with change in seats as the main independent variable and sixth column adds changes in reserved seats as an additional control. Usual household controls include age, gender, marital status, years of education, household size, number of children, extent of agricultural land owned and main source of income.We also include differential trends by religion and caste groups and baseline income quintiles as part of our controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.9: Robustness checks- Confidence in state government

	(1) State specific time trends	(2) District specific controls	(3) District controls with trends	(4) Change in seats ≤ 1	(5) Changes in seats	(6) Changes in reserved seats
Post	-0.0410	0.0265	-0.0848	0.0907**	0.0952**	0.0921**
	(0.0469)	(0.0585)	(0.0582)	(0.0413)	(0.0381)	(0.0386)
Post*Treat1	0.0486*	0.0518*	0.0424	0.0746**		
	(0.0294)	(0.0304)	(0.0284)	(0.0314)		
Post*Treat2	-0.0115	-0.0183	-0.0059	-0.0205		
	(0.0270)	(0.0321)	(0.0273)	(0.0342)		
Relative population	-1.3786	-1.9634	-3.2735	-0.5133	-0.9678	-1.0172
	(3.0032)	(1.4692)	(4.4507)	(1.6663)	(1.5816)	(1.5993)
Election year	0.0777***	0.0747***	0.0800***	0.0760***	0.0741***	0.0751***
	(0.0277)	(0.0221)	(0.0265)	(0.0260)	(0.0244)	(0.0243)
Change in seats*Post					0.0285**	0.0264*
					(0.0139)	(0.0145)
Change in reserved seats*Post						0.0084
						(0.0130)
Observations	50683	50683	50683	45284	50683	50683
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household level controls	Yes	Yes	Yes	Yes	Yes	Yes
State specific time trends	Yes	No	Yes	No	No	No
District specific controls	No	Yes	Yes	No	No	No

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression with the same dependent variable- confidence in state government. The first column shows results by adding state specific time trends to the main regression. The second column introduces the log of district population, proportion of urban population, proportion of SC population, proportion of ST population and proportion of literate population in district as additional controls. The third column shows results with both district controls and state time trends. The fourth column shows result with sample limited to districts that at most gained or lost one seat. The fifth column shows results from regression with change in seats as the main independent variable and sixth column adds changes in reserved seats as an additional control. Usual household controls include age, gender, marital status, years of education, household size, number of children, extent of agricultural land owned and main source of income. We also include differential trends by religion and caste groups and baseline income quintiles as part of our controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: Delimitation and voter turnout

	(1) Period 2000-2018	(2) Period 2000-2018 (limited sample)	(3) False break 1998-2008 (limited sample)	(4) False break 1988-2008 (limited sample)
Post*Treat1	0.0148***	0.0128***		
	(0.0042)	(0.0043)		
Post*Treat2	-0.0023	-0.0010		
	(0.0034)	(0.0035)		
Average no. of candidates per constituency	-0.0005	-0.0003	-0.0000	0.0003
	(0.0005)	(0.0005)	(0.0013)	(0.0004)
District electors/State electors	-0.1862**	-0.2018**	-0.5652	-1.0981***
	(0.0906)	(0.0952)	(0.4730)	(0.3755)
Post 2003*Treat1			-0.0072	
			(0.0056)	
Post 2003*Treat2			0.0005	
			(0.0044)	
Post 1998*Treat1				-0.0082*
				(0.0043)
Post 1998*Treat2				-0.0087**
				(0.0039)
Observations	1896	1780	963	2060
Year fixed effects	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression with district level voter turnout as the dependent variable. The first column gives the result from the main specification for period 2000-2018. The second column omits the states of Chhattisgarh and Uttarakhand which were created in 2000 to make the results comparable to third column. The third column shows results from difference in differences strategy with a false break created in 2003 and covering a 5 year window pre and post delimitation. The fourth column shows results with sample extended to include 10 year window and false break in year 1998. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: Mechanism-Rural development

	(1) Population weighted development index 2001-2011	(2) Population weighted development index 2001-2011(no banks)	(3) Population weighted development index 1991-2001(no banks)
Year 2011 (= 1)	1.0049***	0.8364***	
	(0.1379)	(0.1286)	
Year 2011*Treat1	0.1139**	0.1028**	
	(0.0501)	(0.0454)	
Year 2011*Treat2	-0.0325	-0.0165	
	(0.0493)	(0.0462)	
Relative population	-4.5822	-3.9267	-0.3572***
	(8.4432)	(7.1122)	(0.1308)
Log of district population	0.2151	0.2841	0.0115
	(0.2503)	(0.2123)	(0.0421)
Year 2001 (=1)			0.3043
			(0.2823)
Year 2001*Treat1			-0.0399
			(0.0671)
Year 2001*Treat2			-0.1257*
			(0.0710)
Observations	902	902	720
District fixed effects	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression. The first column shows the results for population weighted development index that is a composite of 18 indicators of public good availability in district for 2001-2011 period. The second column covers 2001-2011 but excludes indicators for banks and credit societies from the development index to enable comparison with 1991 census. The third column shows diff in diff results as robustness check for changes in development index between 1991 and 2001 census. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.12: Changes in representation index

	(1) Change in development index	(2) Change in RRI
Change in RRI	0.1884***	
	(0.0630)	
Change in population of district relative to state	-7.1572	
	(6.4217)	
Change in village population	0.0004***	
	(0.0000)	
=1 if District gained seat		0.0672***
		(0.0149)
=1 if District lost seat		-0.1491***
		(0.0137)
Observations	337781	337781
State fixed effects	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Both columns show results from village level regressions. The first column shows estimated coefficient of change in representation (measured by change in RRI) from a regression of change in village development index on change in representation. The second column shows results from regression of change in representation index on a two separate treatment dummies- the first taking value= 1 if village was located in a district that gained seats and the second taking value= 1 if village was located in a district that lost seats. Villages that were in districts that witnessed no change in seats form the base category. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Mechanism-Candidate quality

	(1) Age	(2) No. of criminal cases	(3) Years of education	(4) Total assets (in lakhs)
Post*Treat1	-1.2191***	-0.1368**	-0.1945	-2.1647
	(0.4465)	(0.0643)	(0.1822)	(21.2613)
Post*Treat2	-0.1535	-0.0053	-0.1468	-7.9599
	(0.3725)	(0.0555)	(0.1745)	(18.0451)
District electors/State electors	-5.0549***	-0.0382	0.6236	-156.9376**
	(1.0000)	(0.2122)	(1.1259)	(71.1096)
Observations	36785	36785	34994	36785
District fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes a separate regression with a different candidate characteristic as a dependent variable. All regressions include district, year and state year fixed effects * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.14: Mechanism-Winner quality

	(1) Age	(2) No. of criminal cases	(3) Years of education	(4) Total assets (in lakhs)
Post*Treat1	-1.6398***	-0.1899	-0.0125	7.0189
	(0.6301)	(0.2255)	(0.2232)	(54.3278)
Post*Treat2	-0.2687	-0.0459	-0.1094	-32.4064
	(0.6009)	(0.2018)	(0.2207)	(68.4549)
District electors/State electors	4.4340	2.0123*	-5.7785***	15.6658
	(4.8943)	(1.1635)	(2.0814)	(763.0547)
Observations	5098	5098	4943	5098
District fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-year fixed effects	Yes	Yes	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes a separate regression with a different winner characteristic as a dependent variable. All regressions include district, year and state year fixed effects * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.15: Timing- Elections with new boundaries

	(1) Confidence in politicians	(2) Confidence in state govt
Post	0.0693	0.1354***
	(0.0597)	(0.0474)
Post*Treat1	0.1009*	0.0717*
	(0.0514)	(0.0369)
Post*Treat2	0.0049	-0.0213
	(0.0499)	(0.0364)
Post*Treat1*No election	-0.0721	-0.0593
	(0.0776)	(0.0587)
Post*Treat2*No election	-0.0175	-0.0891
	(0.0874)	(0.0616)
Relative population	2.5257	-1.2722
	(2.1595)	(1.5102)
Election year	0.0578*	0.1110***
	(0.0299)	(0.0286)
Post* No election	0.0544	-0.0641
	(0.0570)	(0.0402)
Constant	0.2550**	0.7756***
	(0.1142)	(0.0856)
Observations	50976	50683
Household fixed effects	Yes	Yes
Household level controls	Yes	Yes

Standard errors are clustered at district level and given in parentheses. Each column denotes result from a separate regression. "No election" takes value 1 if respondent belongs to a state that did not have elections under newly delimited boundaries until the second round of the IHDS. First column shows heterogeneity of effects by election timing for confidence in politicians. Second column shows the same for confidence in state government. Controls include age, gender, marital status, years of education, household size, number of children, extent of agricultural land owned and main source of income. We also include differential trends by religion and caste groups and baseline income quintiles as part of our controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

The Political Economy of Higher Education: Criminality and Institutions

3.1 Introduction

A profusion of recent work in economics and political science has devoted itself to addressing the question of the impact of politician quality on economic and other outcomes. For instance, educational attainment of leaders has been widely used as a proxy for politician quality in empirical research (Besley et al., 2005; Baltrunaite et al., 2014) where it has been shown that higher education of leaders leads to higher economic growth (Besley et al., 2005) and higher educational attainment of citizens (Diaz-Serrano and Pérez, 2013). Another measure of politician quality is their criminality which has been the specific focus of a substantial part of this recent literature. While the detrimental impact of politician criminality on economic growth (Prakash et al., 2019), welfare and crime (Chemin, 2012), private investment (Nanda and Pareek, 2016) and household poverty (Cheng and Urpelainen, 2019) has been established, less is known about the impact of politician criminality on human capital.

In this paper, I investigate the impact of electing representatives accused of ***serious*** crimes[1] on the provision of higher education institutions, in particular of new public

[1]The definition of serious crimes and the rationale for using serious crime accusations as an indicator of criminality is discussed in detail in the data section

state government colleges in the context of state legislative assembly elections in India. I focus on higher education institutions due to overwhelming evidence that establishes the robust role of higher education in promoting economic growth and development. For instance, not only does higher rate of tertiary education lead to higher economic growth (Castelló-Climent et al., 2018), but Castelló-Climent and Mukhopadhyay (2013) also show that the share of population with tertiary education is positively correlated with state level economic growth while there was no such association for population shares with only primary or middle school education in India. The potential role of higher education institutions in promoting economic growth by way of improving tertiary education attainment in the population provides the rationale for focusing on higher education institutions as an outcome of interest.

Within the tertiary education landscape, government funded and managed colleges play an important role in India accounting for 33% [2] of share in total enrollment. State government funded and managed colleges account for about 52% of the total number of government colleges in India[3]. Members of state legislative assemblies (MLAs) often play a key role in lobbying the state governments for bringing new colleges to their constituencies. Numerous accounts suggest that not only do MLAs initiate proposals for opening new colleges in their constituencies (TheHindu, 2021b[4]; TheTimesofIndia, 2018[5]; TheTribune, 2019[6]; TheHindu, 2021a[7], TNM, 2018[8]), but are also instrumental in obtaining approvals for funds and land for such institutions (TheTribune, 2020[9]). Extensive reporting in the media with MLAs being explicitly credited for these developments also

[2]Source: the final report of the All India Survey on Higher Education 2016-2017. The report does not give a breakdown of the enrollment percentage by centrally managed or state managed government colleges

[3]Source: raw data from the All India Survey on Higher Education 2016-2017

[4]This reports that two MLAs from the state of Tamil Nadu made requests to the state government to set up degree colleges in their respective constituencies during the Question Hour of the State Legislative Assembly. The Higher Education Minister of the state consequently responded by informing the House that their pleas will be taken into consideration.

[5]News report that the UP state government had accepted the demand of Agra district's Bah constituency MLA Pakshalika Singh's demand of setting up a degree college in her area.

[6]Article quoting the MLA of Chabbewal constituency in the state of Punjab as expressing gratitude to the CM and Finance Minister for accepting his request for opening a college in his constituency

[7]Article quotes the MLA of Gurazala constituency in the state of Andhra Pradesh expressing gratitude to the CM for sanctioning a medical college in his constituency during his tenure.

[8]Reports the efforts of the MLA of Nawanshahr in the state of Punjab toward opening a degree college in his constituency where the MLA also states the amount of funds being sanctioned by the state government and land approval obtained from the local panchayat.

[9]Reports the MLA of Ludhiana East in Punjab visiting the construction site of a college in constituency which he describes as his "dream project" while giving an update about the construction status and his efforts to finish the project.

indicates that such attribution matters for future elections and that higher education institutions can be considered as being electorally salient.

To look at the impact of politician criminality on the provision of state government colleges, I exploit the landmark judgement of the Supreme Court of India in 2003 which made it mandatory for all candidates contesting in state or national elections in India to file a sworn affidavit disclosing their criminal record, financial status and education status at the time of filing their nomination papers. This affidavit information not only states the number of charges but also allows us to identify the type of charges. In this paper, I use accusations of serious crimes as the relevant indicator for politician criminality. This is for two reasons– first, focusing on distinction between candidates accused of serious crimes alleviates concerns that criminal charges might be frivolous and politically motivated in nature and not a true indicator of criminality. Second, empirical evidence has emphasized that the negative impact of politician criminality is amplified for politicians with serious charges (Prakash et al., 2019; Prakash et al., 2021).

I use a regression discontinuity design focusing on close elections between candidates accused of serious crimes and candidates not accused of serious crimes contesting for position of members of legislative assembly (MLA) in state elections between 2004 and 2008. Looking at elections with narrow margins between criminal and non-criminal candidates allows one to make a credibly causal claim about the impact of criminality by generating *near* random variation in the criminal identity of the elected leader within this limited sample (Imbens and Lemieux, 2008; Lee, 2008). I support this identification strategy by showing evidence from empirical tests that establishes the validity of this regression discontinuity design in this context.

This paper uses location information and year of establishment of new state government funded colleges obtained from the AISHE 2016-17 (All India Survey on Higher Education) and maps them to state legislative assembly constituencies to create a constituency year level panel. Using this data set and employing a regression discontinuity design, results show that constituencies that elect a candidate accused of serious crimes are 3.8% less likely to receive a new state government college in a year compared with constituencies that elect a non-criminal candidate on an average. This is a large impact considering that the probability of receiving a state government college in a year is only 1.6% on an average for the full sample. This impact translates into a 18.7 percentage

points decrease in the probability of obtaining a new state government college during an electoral term.

One might argue that the welfare loss from such reduced provision of state funded colleges by criminal politicians could be compensated by an accompanying increase in opening of new private colleges in such constituencies. For instance, it might be the case that criminal politicians in their bid to extract rents, encourage the opening of private colleges as opposed to government funded colleges. It is possible that the extensive requirements for licensing and approvals for opening private colleges in the tightly regulated higher education sector offer greater opportunities for rent extraction. However, I find no evidence that this decrease in provision of state government funded colleges is off set by an improvement in opening of new private colleges.

Extending the discussion on welfare implications, I use data about individuals' educational attainment from the 71st round of the National Sample Survey (NSS) conducted in 2014. I show that the reduced provision of state government funded colleges is most likely to affect the educational attainment of the most economically vulnerable sections of society who depend on affordable education provided by state run higher education institutions.

In a notable departure from previous evidence which found that the negative impact of criminal politicians on economic growth was greater for politicians **not aligned** with the state government (Prakash et al., 2019), I find that this detrimental impact on new public colleges is in fact more pronounced for criminal candidates belonging to the Chief Minister's party in the state. This is important as it emphasizes that the nature of public education institutions as a development outcome is distinct from outcomes like aggregate economic growth which are a composite of growth of both public and private sectors. Here the need for access to power is essential in order for politicians to be able to affect the provision of higher educational institutes in the first place. To add to the heterogeneity in effects based on other candidate characteristics, I find that the negative impact is greater for criminal candidates who were not re-elected incumbents vs. criminal candidates who were re-elected. I find no difference in effects between criminal candidates with college education vs. candidates without college education. I also do not find any evidence pointing to heterogeneity in effects depending on the existing level of development or corruption in states.

While existing work suggests that rent seeking behaviour by criminal politicians can dampen economic growth by reducing private investment and consumption (Nanda and Pareek, 2016; Chemin, 2012; Prakash et al., 2019), criminally charged politicians have also been shown to display lack of political effort in terms of attendance rates in legislatures and utilization of local funds (Gehring et al., 2019). I suggest that both of these aspects could explain the lower provision of state colleges in constituencies which elect leaders with serious crimes. First, state government institutes as government undertakings provide education at subsidized costs with little room for politicians to profit, meaning that self-interested criminal politicians would supply less of it. Secondly, lower effort is likely to translate into lesser new government colleges as their establishment would entail lobbying for resources with the state government.

This paper contributes to the understanding of the political economy of the education sector, establishing the first links between politician quality and the provision of higher education in the Indian context. The involvement of the state in education provision at all levels belies any notion that distribution in this respect would be unaffected by politician characteristics. Lahoti and Sahoo (2020a) for instance, show that graduate politicians improve education outcomes and funding but only in states where the level of development is high to begin with. Second, this paper directly contributes to the literature on the impact of criminal politicians by extending the scope of investigation to a hitherto unexplored area of an exclusively publicly provided good. Unlike the impact of politician criminality on aggregate economic activity which could be driven by a deleterious impact on both public and private sectors, this paper establishes how criminal politicians can be specifically detrimental to the provision of a government supplied good. Thirdly, this paper uncovers interesting heterogeneity in the impact of politician criminality by their alignment with the ruling party. While previous work showed evidence that impact on economic growth is worse where the criminal politicians are **not** aligned with the ruling party (Prakash et al., 2019), this paper finds that the opposite is true for the provision of new state government colleges. This goes out to show that the nature of impact of politician criminality depends on the nature of the outcome in question. Where access to power is essential for moving outcomes as is the case of government colleges, a distinction between the performance of criminal vs. non criminal politicians is only manifest when the politicians are aligned with the Chief Minister's party.

This paper is most closely related to works exploring the impact of politician criminality on development outcomes. Prakash et al. (2019) demonstrate using a regression discontinuity design that criminally accused politicians lead to slower economic growth as captured by night time luminosity in constituencies where they are elected, with this negative effect being much larger in cases where politicians were accused of serious crimes in India. Contrary to the results in this paper, they find larger impact in constituencies where the criminal leader is *not aligned* with the state government, which they attribute to oversight from the ruling party. Chemin (2012) looks at the impact of criminal politicians on welfare and finds that criminal politicians lead to lower consumption expenditure by marginalized castes and increase in criminal activity.

Nanda and Pareek (2016) use close election between criminal and non-criminal candidates and show that election of criminal politicians reduces the return on investment and total investment for private firms, however this decline is largely mitigated by an increase in investment in state controlled firms, This is in contrast to the findings for reduction in provision of state government colleges where I find no offsetting increase in provision of private aided or private unaided colleges. Cheng and Urpelainen (2019) make a distinction between different types of policy outcomes while identifying the effects of criminal politicians. They show that election of state legislators with criminal charges leads to worse outcomes in terms of household electrification and literacy rates but no worse outcomes in terms of provision of local infrastructure such as paved roads or power grids. Prakash et al. (2021) look at the impact of politician criminality on the total crime in a district and find that the total crime count is higher by 1055 cases per year for a district where all leaders were accused of serious crime compared with a district with no leaders accused of serious crimes. While all of these papers look at different economic outcomes, this is the first such study to examine the impact of criminal politicians in the arena of higher education provision.

The rest of the paper is organized as follows: Section 2 gives the background on the role of politicians in higher education and criminality in politics in India. Section 3 describes the main data. Section 4 details the regression discontinuity design used to identify causal impact and establishes its validity. Section 5 shows the main results along with sensitivity analysis and robustness checks. Section 6 looks at heterogeneity in the results based on other politician and constituency characteristics. Section 7 provides a

discussion about the possible mechanism, Section 8 discusses the potential implications for human capital, Section 9 puts the findings into perspective and Section 10 concludes the paper.

3.2 Background

3.2.1 Higher education in India

This section describes the higher education system in India and the potential influence of MLAs in provision of state government colleges. The higher education system in India includes both public and private institutions. This comprises public universities and their affiliating colleges as well as private universities. Public universities can be broadly classified as central universities which are promoted, funded and governed by the central government and state universities which are promoted, funded and managed by state governments.[10] While public universities can have both public and private affiliating colleges, private universities in India are not allowed to have affiliating colleges. The higher education system has traditionally been dominated by public sector institutions, however the post liberalisation era has seen a proliferation in the number of private colleges in India, especially after 2000.

State run institutions nevertheless continue to occupy a key position in the tertiary education landscape. As per the AISHE 2016-17, even though government colleges accounted for 22% of all colleges, their share in total enrollment was disproportionately higher standing at 33%. Not only are government institutions considerably more affordable than their private counterparts, even public perception is generally tilted in the favour of government institutions for higher education. According to data from the 71st round of the National Sample Survey (NSS) 2014, of all individuals attending private institutions for graduate degree or diploma or above, 42% stated inability to get admission in government institutions as the reason for attending private with another 17% stating the unavailability of government institutions. Only 7.5% stated the quality of government institutions not being satisfactory as the reason.

[10] AISHE classifies universities based on their type into central universities, state public universities, state private universities, deemed universities-government, deemed universities-government aided, deemed universities-private, institutes of national importance, institutes under state legislature acts, central open university, state open university, state private open university and others. Central, state public and state private universities constitute about 72% of all universities in India.

According to the AISHE classification, government colleges can be further classified on the basis of their management as central government colleges, state government colleges, colleges run by local bodies and university colleges. State government run colleges account for the bulk of government colleges in India with their share being at 52% as per the AISHE 2016[11]. Given the preponderance of state government colleges in the arena of public education institutions and given the fact that they are most likely to fall within the sphere of influence of state level politicians, I focus on the impact of criminality of politicians on additions to this particular category of institutions.

3.2.2 Politicians and higher education

The scope of legislative and executive responsibilities and powers of Members of Legislative Assemblies of states or MLAs is widely known. MLAs are supposed to debate and vote for bills, amend laws, raise issues and to hold the state government accountable for its spending and activities. Recent literature has shed light on another sphere of MLA powers– their control over the bureaucracy (Iyer and Mani, 2012; Gulzar and Pasquale, 2017; Nath, 2015). This originates in their ability to influence bureaucrats' careers through assignment and transfer of jobs. MLAs also lobby the state government to direct resources and funds for projects and schemes towards their own constituencies (Jensenius, 2015). Recognizing these two roles help to understand how MLAs can influence the opening of new state government colleges.

Firstly, MLAs can influence the location of new colleges by lobbying the state government to locate any planned new colleges in their own constituencies. Often, state governments announce plans to open up a fixed number of new colleges in a period, but the location of the said colleges is left to be determined. With their lobbying power and bureaucratic control, MLAs can exercise efforts to make sure such colleges are opened in their constituencies. For instance, TheTimesOfIndia (2020) reports the plans of the government of the state of Haryana to open one additional college at the request of the MLA of the Kalayat constituency. Jagnani and Khanna (2020), in their paper on effects of new elite public colleges, similarly mention that state administrators often lobby the federal government to procure such colleges.

[11]Considering central government, state government, local body run colleges and university colleges as constituting the sum of government colleges. state government colleges account for 52% of the total, local body run colleges are 26% of the total with central government colleges accounting for a mere 2%.

Secondly, MLAs can also propose and create pressure on the state government for setting up new institutions in their constituencies by raising requests in their state legislative assemblies. TheHindu (2021b) reports that two MLAs from Tamil Nadu raised the request for opening colleges in their respective constituencies during the Question Hour in the State Legislative Assembly. Similarly, DeccanHerald (2021) reports the demands of an MLA from the state of Karnataka towards the state government for releasing funds for the Department of Higher Education of the state. This is supported by news reports crediting MLAs for establishment of state government colleges. TheTimesofIndia (2018) reports that the demand by the MLA from constituency of Bah in Uttar Pradesh for opening a college in her area was approved by the UP Government. Media reports also often quote MLAs as thanking the state government for approving their request for new colleges in their constituencies (TheTribune, 2019; TheHindu, 2021a).

Third, bureaucratic access implies that MLAs can potentially manipulate the maze of processes involved in the setting up colleges including land allotment by state authorities, permissions from the Departments of Higher Education of their respective states, as well as granting of affiliation by state universities, to suit their own purpose. For instance, TheTimesofIndia (2013) reports the District Commissioner expediting identification of land for medical college in Hazaribag in the state of Jharkhand after the local MLA who initiated the college took up the issue. TNM (2018) reports the MLA from Nawanshahr in the state of Punjab as being instrumental in obtaining sanction for opening a degree college in the area and quotes the MLA describing the extent of funds released by the state government and the land allotted by local authorities for building the college. Extensive coverage in the news also indicates that politicians seek credit and attribution for these institutions. This along with efforts to make sure the colleges are set up during their tenure (TheTribune, 2020; TribuneNewsline, 2018; TheHindu, 2021a) suggests that public colleges are electorally salient.

To further emphasize the influence of politicians on higher education, Kapur and Mehta (2007) highlight the tight control politicians exercise over the higher education system and the benefits that accrue to them through rent seeking on contracts, appointments, admissions and grades; using colleges as screening mechanisms for politics and for propagating partisan politics. Therefore, politicians remain closely involved in not only the establishment but also in the management and administration of colleges.

3.2.3 Criminality of politicians

The issue of rampant presence of criminally accused politicians in electoral politics in India has captivated the attention of popular media and researchers alike. For instance, about 43% of the elected MPs (Members of Parliament) in 2019 had criminal charges against them, up from 34% of the elected MPs in the 2014 election (TheHindu, 2019). Against this backdrop, the Supreme Court passed a judgement in 2003 making it compulsory for all candidates seeking to contest elections to the Parliament or State Legislative Assemblies to file a sworn affidavit with details of all criminal charges and convictions against them as well as their educational qualifications and assets and liabilities.

The pervasiveness of criminally accused politicians has spawned research dedicated to determining the reasons why political parties field criminal candidates (Aidt et al., 2011; Vaishnav, 2017) and why voters choose to vote for them (Banerjee et al., 2014; George et al., 2018). At the same time, evidence emerging from recent literature has lent credence to the hypothesis that criminal politicians lead to worse outcomes. Prakash et al. (2019) use a regression discontinuity design with close elections between criminally accused and non criminally accused candidates and find that criminally accused MLAs lead to lower economic growth. Chemin (2012) uses data on criminal charges against MPs and finds that criminal politicians reduce private consumption by backward and marginalized castes as measured by monthly per capita expenditure.

Nanda and Pareek (2016) demonstrate a negative impact of criminal politicians on private sector investment which is largely offset by increase in investment in state-owned firms. Cheng and Urpelainen (2019) find that election of MLAs with criminal charges in close elections against non-criminal candidates leads to worse outcomes in terms of electrification and literacy but has no impact on local infrastructure. Criminal politicians also lead to more criminal cases in districts where they are elected (Prakash et al., 2021; Chemin, 2012).

Existing studies on the impact of politician criminality have suggested that the tendency of criminal politicians to indulge in rent seeking behaviour and to promote criminal activity can impact economic outcomes by discouraging private investment (Nanda and Pareek, 2016) and reducing consumption (Chemin, 2012). This thesis is supported by Asher and Novosad (2018) who show that as local rents increase, criminal politicians are more likely to win elections and elected politicians accumulate more charges as well as

gain more wealth. Another reason why criminal politicians are likely to result in worse outcomes is the lack of political effort. Gehring et al. (2019) provide evidence in support of this by showing that criminally charged politicians exhibit lower attendance rates in Parliament and lower rates of utilization of local development funds.

The impact of criminal politicians on the provision of state funded colleges could follow from both these channels. If it was true that new state government colleges would allow for opportunities for rent extraction, we might find that criminal politicians encourage their openings. On the other hand, if such opportunities were limited, one should expect lesser provision of government colleges. Following from the second channel, considering that opening new state government colleges involves lobbying on the part of the MLA, the lack of effort by criminal politicians could translate into lower provision of government colleges by them.

3.3 Data

3.3.1 Data sources

State Government Colleges

The main source of data for new college openings is the AISHE (All India Survey on Higher Education) from the year 2016-2017. This survey is an annual exercise conducted by the Ministry of Education that aims to collect detailed data from all higher educational institutions on a wide range of parameters including their type, management, time of establishment, location, courses offered, student enrollment, teachers and infrastructure. The AISHE 2016-17 dataset reports these variables for 34906 colleges all over India.

The AISHE broadly categorizes higher education institutions into three types– universities, colleges and stand-alone institutions[12]. Colleges are further classified into six types based on their management– central government colleges, state government colleges, colleges run by local bodies, private aided colleges, private unaided colleges and university colleges. As stated above, this paper focuses on public colleges funded and managed by the state governments.

[12]Universities are empowered to award degrees under some Act of Parliament or State Legislature, colleges are not empowered to award degrees on their own and are affiliated to universities and stand-alone institutions are not affiliated with universities and are not empowered to provide degree and therefore run Diploma level programmes (Source: Report on AISHE 2016-17)

The AISHE 2016-17 raw data is a college level data set that reports the name, location, year of establishment, management type, courses and infrastructure of 34906 colleges. These lists of college names and addresses were fed into the Geocoding API available on the Google Maps Platform which returned the geographic coordinates of each of these colleges. Of all the 34906 colleges in the raw dataset, I was able to map 32889 or 94.2% colleges to their coordinates. This mapping was then combined with state legislative assembly constituency shapefiles made publicly available by Sandip Sukhtankar and Manasa Patnam[13] to map colleges to the assembly constituencies within which they are located.

This geocoded college data is the first of its kind constituency level higher education institutions data set to be used in empirical research in the Indian context, to the best of my knowledge. Further, using the location and year of establishment of colleges, I was able to create a constituency level panel of new college openings and number of existing colleges in each year. I use whether a new state government college was opened in a constituency in a given year as the main dependent variable of interest[14]. This constituency level panel data of college openings was combined with MLA election and affidavit data which brings us to the second set of data sources described below.

Election data and criminal charges

The data for state legislative assembly elections comes from the legislative assembly election results from 1980-2018 compiled by Jensenius and Verniers (2017) and made publicly available as a part of the SHRUG open data platform (Asher et al., 2021). These include the constituency wise candidate names, their party affiliation, gender, position, vote share as well as election month, year, total electors, reservation status of constituency and turnout percentage for all state assembly elections held between 1980 and 2018.

The data for criminal charges of candidates is obtained from the data set compiled by Prakash et al. (2019)[15]. This data set includes the number of criminal charges against all candidates as well as their age, gender, education and assets and liabilities. Additionally,

[13]https://uva.theopenscholar.com/sandip-sukhtankar/data

[14]I alternatively use the number of state colleges opened as a dependent variable. The instances of more than one new state college opening up in a constituency in a year are rare, so both these variables are approximately the same.

[15]This data has been made available for public use through the SHRUG open data platform (Asher et al., 2021) and was downloaded from the same

it also codes for whether the candidate was charged for atleast on major crime as per the definition followed by the Association for Democratic Reforms (ADR)[16]. Since the candidate affidavits were available only for the period after the 2003 Supreme Court ruling which made disclosure of criminal charges mandatory, this data is only available for elections held from the year 2004 onwards.

The delimitation exercise in 2008 led to a redrawing of assembly constituency boundaries because of which the constituencies before and after delimitation are essentially incomparable. At the same time, an examination of year wise number of state government colleges reveals an increasing trend from 2005 onwards after flat-ish growth in the preceding years. This increasing trend reaches it's pinnacle in years 2007 and 2008 before falling again (shown in Figure 3.A1 in the appendix). Since this is the period that saw the most robust growth in state government colleges, I look at the impact of criminality on new state colleges during this pre-delimitation period.

This limits the sample to elections held after the affidavits were made mandatory from 2004 but before elections were held under new boundaries in 2008. This 4 year window implies that the sample includes at most 1 election per state. Table 3.1 shows the states included in the final sample and their election years. The regression discontinuity design entails including only those elections where the winner was criminally charged and the runner-up was not or where the runner-up was criminally charged but the winner was not. This leaves us with 497 unique constituencies and 2442 constituency-year observations where the top 2 candidates were of differing criminal status and for which college openings data could be mapped.

The definition of criminality I use is whether a candidate is accused of a serious crime as per the ADR definition. This is on account of two reasons. First, there is a frequently expressed concern that criminal charges against candidates could be politically motivated, being filed by political rivals to tarnish reputations and would therefore, not be a true indicator of criminality. Some of these charges might also be incurred in the process of engaging in political protests which inadvertently turned violent (Prakash et al., 2021).

[16]The ADR lists 8 criteria for coding serious criminal cases– A crime is classified as a serious crime depending on whether the maximum punishment for the offence committed is five years or more, whether the offence is non-bailable or offences pertaining to electoral violation or offence related to the loss to exchequer or offences related in nature to assault, murder, kidnap or rape, offences mentioned in Representation of the People Act(Section 8), offences under Prevention of Corruption Act or offences related to crimes against women. It also gives an exhaustive list of IPC(Indian Penal Code) Sections charges filed under which are considered to be of a serious nature.

This is less likely to be the case for serious criminal charges. Second, evidence from the impact of criminality on outcomes has consistently reported worse outcomes where politicians were accused of serious charges. Prakash et al. (2019) find that the type of charges matter for the impact of criminality on economic growth, they show that politicians with serious charges negatively impact economic growth whereas the results for politicians with non serious charges are insignificant. Similarly, Prakash et al. (2021) find substantially higher increases in crime when the politician is accused of a serious crime than when looking at politicians accused of any crime. It is for these reasons that I focus on the presence of serious criminal charges as a reliable indicator of politician criminality.

3.3.2 Descriptive statistics

Table 3.2 reports the summary statistics of the main dependent variable and other baseline constituency characteristics for the sample. The left panel reports statistics for all the constituency year observations in the included states that had elections between 2004 and 2008 and for which college data could be mapped. The right panel reports the same statistics but for the regression discontinuity sample where either the winner was accused of a serious crime and the runner up was not or the runner up was a serious crime accused and the winner was not.

The probability of a constituency witnessing the opening of a new state government college in a year is about 1.6 percent for the full sample and 1.4 percent for the regression discontinuity sample. Since the opening of more than one state college is a rare occurrence, the statistics for the number of new state college openings are roughly the same as that for any state college opening. On an average, each constituency had less than one state government college while the average number of total colleges in a constituency is 4.7 both in the full sample and the RD sample.

The average turnout percentage is about 64 percent. About 32% constituencies in the full sample and 27% constituencies in the regression discontinuity sample had elected an MLA belonging to the Chief Minister's party in the previous election. Around 32% constituencies had re-elected the incumbent in the previous election in the full sample and the corresponding figure is 30% in the RD sample. An overwhelming 94% constituencies had elected a male leader in the previous election. 10.7% constituencies were reserved for

SCs and 4% were reserved for STs in the RD sample.

3.4 Regression Discontinuity Design

A simple regression of the outcome of interest on whether an MLA was accused of a serious crime would yield biased estimates due to the fact that constituencies that elect criminal politicians are likely to be different from constituencies that do not elect criminal politicians and these unobservable differences could be correlated with opening of new state government colleges. An ideal situation to evaluate impact of criminality would be where criminal and non-criminal politicians are randomly allocated to constituencies. However, the reality of electoral politics is far removed from this ideal research scenario.

Regression discontinuity design (RDD) (Lee and Lemieux, 2010; Imbens and Lemieux, 2008) as an empirical strategy provides an approximation to this ideal scenario and has been widely used in the context of elections. An RDD entails comparing constituencies where a criminal candidate won a close election against a non-criminal candidate to constituencies where a criminal lost a close election against a non-criminal candidate. The underlying intuition is that very close elections, i.e., elections where the winning margin is very narrow generate almost random variation in the criminal identity of the winning candidate (Lee, 2008). The assumption here is that constituencies where criminal candidates narrowly won are likely to be similar to constituencies where criminal candidates narrowly lost in terms of unobservable characteristics.

Formally, an RDD defines a threshold value of a forcing variable based on which treatment (in this case, a criminal MLA) is assigned. The electoral context lends itself easily to this framework. The sample is limited to elections where a candidate accused of a serious crime won against a candidate not accused of a serious crime and elections where a candidate accused of a serious crime lost against a candidate not accused of a serious crime. The vote margin of the criminal candidate defined as the difference in vote shares of the criminal and non criminal candidates acts as the forcing variable with a cut off value of 0. Treatment is assigned, i.e., the elected MLA is criminally accused when the forcing variable exceeds 0 ($margin \geq 0$). Negative values of $margin$ correspond to elections where the criminally accused candidate lost against non criminal candidate.

The idea is to limit the sample to observations to the vicinity of the threshold value

and check for a discontinuous jump in the outcome at the threshold where there is a discontinuity in treatment. In this sample where observations are limited to a certain bandwidth around the threshold, the election of a criminal MLA happens as if it was random. I use local linear regressions with a triangular kernel and use the optimal bandwidth generated by the data driven procedure described in Calonico et al. (2017). More specifically, I limit the observations to the optimal bandwidth and implement the following regression with weighting based on a triangular kernel:

$$Y_{cst} = \alpha + \beta seriouscrime_{cst} + \beta_1 margin_{cst} + \beta_2 margin_{cst} * seriouscrime_{cst} + \epsilon_{cst} \tag{3.1}$$

for $margin \in$ bandwidth

Here, Y_{cst} is a dummy that takes value 1 if a new state government college was established in constituency c in state s in year t. $seriouscrime_{cst}$ is a dummy that takes value 1 if the MLA of constituency c in state s in the year t was accused of a serious crime and 0 if the MLA was not accused of serious crime. For election years, I code $seriouscrime_{cst}$ as equal to 1 if the outgoing MLA was accused of a serious crime implying that attribution of any new govt colleges in an election year should be to the outgoing MLA as the incoming MLA is unlikely to be able influence new college openings so quickly. However, I also show in the robustness checks that the results are invariant to excluding election years from the sample.

β , the coefficient of *seriouscrime* is the main estimate of interest which captures the discontinuity in the outcome variable as the victory margin of criminally accused MLAs turns positive. I also allow for an interaction of *margin* with *seriouscrime* dummy in effect allowing different slopes on different sides of the threshold as is standard in RDD. Standard errors are clustered at the constituency level in all regressions. While the main specification does not control for year and state fixed effects, I introduce them and show that the results still go through as a part of robustness checks.

3.4.1 Validity

There are two main threats to the validity of the RD design. The first is the probability of manipulation of the forcing variable around the threshold and the second is presence of

discontinuities in pre-determined constituency characteristics at the threshold. I address both of them in the proceeding sections.

Manipulation of forcing variable

If candidates with serious criminal charges could manipulate close elections in a way that their probability of winning would be higher, then the assumption of approximately randomized assignment of treatment around the threshold would be violated. If this sort of manipulation was true, it would present itself as a discontinuous jump in the probability of a criminal candidate winning at the treatment threshold value of 0. In this case, we would expect a discontinuous jump in the density of the winning margin of the criminal candidate ($margin$).

To check for this, Figure 3.1 plots the histogram of margin of criminal candidate and shows that there is no evidence of a discontinuity at the threshold of 0. I also formally check for manipulation of the forcing variable by implementing the McCrary test following McCrary (2008). Figure 3.2 shows the result, pointing to no discontinuity. The estimated size of discontinuity in the density of margin is 0.03979 with a standard error of 0.16152 showing that we cannot reject the hypothesis of continuity in the forcing variable at the threshold value of 0.

Balance in pre-determined characteristics

The other threat to the validity is the presence of discontinuity in pre-determined constituency characteristics at the threshold of treatment. If there is a discontinuous change in pre-treatment constituency characteristic, it would imply that the constituencies that elect candidates with serious charges in close elections are systematically different from constituencies that elect non-criminal candidates violating the assumption of successful randomization around the threshold (Lee and Lemieux, 2010).

It must be noted that this does not require balance in characteristics with criminal and no criminal constituencies, it only requires that there is no discontinuity in characteristics around the threshold value. I formally check this by using the pre-determined characteristics as the dependent variable in my main regression discontinuity specification. I do this using two different bandwidths, first I use optimal bandwidths derived separately for each of these dependent variables using the data driven procedue in Calonico et al. (2017).

Secondly, I check for discontinuities in characteristics using the optimal bandwidth calculated using the main RD specification for the main dependent variable– whether a new state government college was opened. Following Calonico et al. (2017) the calculated bandwidth is 7.248.

Table 3.3 shows estimates for discontinuities. Columns 1 and 2 show the means for criminal (serious) and non criminal (non serious) MLA constituencies for the full sample where winner and runner-up have different criminal status, column 3 shows the simple difference of means. Column 4 shows the regression discontinuity estimate for optimal bandwidths derived separately for each of the dependent variables. Column 5 shows the regression discontinuity estimates with the pre-determined optimal bandwidth of 7.248 which is used for the main results.

The difference in new state government college openings in the previous electoral term is not significantly different in constituencies that elected MLAs with serious crimes vs constituencies that did not elect MLAs with serious crimes. The regression discontinuity estimate is insignificant as well, suggesting that there was no pre-existing difference. The same is the case for the number of existing state government colleges indicating that there is no pre-existing differential demand for state colleges between these two types of constituencies.

I also check for discontinuities in electorate size, turnout percentage, alignment with the ruling party, MLA's gender and incumbency status in the previous election and whether the constituency was reserved for STs or SCs. Columns 4 and 5 show that none of these characteristics exhibit discontinuities lending support to the validity of the RDD in this setting.[17]

3.5 Main results

Figure 3.3 shows the regression discontinuity graphically. The scatter plot represents the means of the dependent variable– whether a new state government college was opened within $margin$ bands of width 0.5%. The solid lines are linear fits from specification (1) estimated with a triangular kernel and allow for different intercepts and slopes on either side of the threshold value. The difference in the intercept at $margin$=0 represents the

[17]Table A1 in Appendix shows the balance table for constituencies with criminal and non-criminal MLAs irrespective of the criminal status of the runner-up

effect of having an MLA with serious criminal charges. The graph indicates a discontinuity in probability of new state college opening at the threshold value, indicating a fall in new state college openings as we move from constituencies with leaders not charged of serious crimes to constituencies with winners charged with serious crimes,

Table 3.4 shows the regression discontinuity estimate based on specification (1). The data driven procedure for optimal bandwidth calculation[18] in Calonico et al. (2017) gives a bandwidth of 7.248% on either side of the threshold. I alternatively also show results with 0.5 times and 1.5 times of the optimal bandwidth of 7.248 in line with standard practice.

Column 1 from Table 3.4 shows that constituencies that elect candidates with serious criminal charges are 3.8 percent points less likely to receive a state government college per year compared to constituencies that elect candidates without serious criminal charges. Considering that the average likelihood of getting a state government college in a given year is only 1.4 percent for our sample (and 1.6 percent for the full sample), this is a sizeable effect. Columns 2 & 3 show the specification with 0.5 times and 1.5 times the optimal bandwidths and the estimates remain significant and range between 2.4 to 4.5% in magnitude suggesting robustness across different bandwidths.

3.5.1 Sensitivity Analysis

To establish the robustness of the results, I undertake a host of sensitivity analyses. First, I report the regression discontinuity estimates with weighting based on a triangular kernel varying the bandwidths from 1% to 10% margin on either side of the threshold value of 0. These are shown in Figure 3.4, The RD estimate is consistently negative through out the range of bandwidth and is significant at 10% for bandwidths between 4 to 10%. The detailed results are shown in Table 3.5.

Second, I estimate the regression discontinuity using weighting based on a rectangular or uniform kernel which weights observations within the bandwidth the same irrespective of their distance from the threshold value. I report these discontinuity estimates varying the bandwidth from 1% to 10% of margin on both sides of the threshold. Figure 3.5 plots the RD estimates from these regressions. Again the RD estimate is consistently negative

[18] According to Calonico et al. (2017), this bandwidth choice is an upgraded version of both the IK and the CCT implementations of the MSE-optimal bandwidth selectors discussed in Imbens and Kalyanaraman (2012) and Calonico, Cattaneo, and Titiunik (2014b), respectively

but significant at 10% level for 3 to 7% size of bandwidths. Table 3.6 reports the detailed results from these regressions.

Third, instead of a local linear regression, I allow for higher order polynomials functions of margin on either side of the threshold and estimate the regression discontinuity from this specification, These are shown in Table 3.7. The RD estimates from the quadratic, cubic and quartic polynomial specifications are in fact higher in magnitude to the local linear regression estimates and are significant at 5% level indicating robustness of the results to alternative specifications.

3.5.2 Robustness

I implement a host of robustness checks to strengthen the main results. First, I include state and year fixed effects in the RD specification. The RD estimate from this specification is shown in the first column of Table 3.8. The RD estimate is similar in magnitude to the original result and is still significant at 5% level indicating that the results are not driven by specific states or years within our study period.

Second, I add the existing number of state government colleges in a constituency at the beginning of each year as a covariate in my RD specification. The balance checks already showed that there was no discontinuity in the number of state government colleges at the threshold, adding it as a control further allows us to account for differences in the existing stock of colleges that might influence future provision. The RD estimate from this specification is shown in column 2 of Table 3.8. The RD estimate remains significant and is in fact slightly larger in magnitude than before.

Third, instead of using whether any state government college was opened in a given constituency in a given year as the dependent variable, I consider whether any state college was opened in the constituency during the ***electoral term*** as the outcome. The result for this specification is shown in column 3 of Table 3.8. The RD estimate indicates that constituencies with MLAs with serious charges are 18.7% less likely to receive a state college during a whole electoral term compared with constituencies where MLA is not charged with serious crimes.

Fourth, instead of a binary dependent variable indicating whether or not a constituency received a new state government college in a year, I use the total number of new state government colleges opened in a year as the dependent variables. Since

instances of more than one such college coming up in a constituency in a year are rare, these two definitions of dependent variable are roughly equivalent. Nevertheless, I show the results for the number of new state government colleges as the dependent variable in column 4 of Table 3.8. As expected, the RD estimate is significant and approximately the same in magnitude as before.

Finally, I implement the RD specification excluding the observations from election years. This is to mitigate any concerns about ambiguity in the attribution of new colleges that come up in election years to the incoming or the outgoing MLA. In my main specification, colleges opened during election years are attributed to the outgoing MLA with the reasoning that the new MLA cannot move things as quickly as to open a college in the same year. To assuage concerns that this attribution might be flawed and driving the results, I exclude election years and the resulting RD estimate is shown in column 5 of Table 3.8. We see that the RD estimate remains significant and of similar magnitude as before.

3.6 Heterogeneity

Studies looking at the impact of criminality of politicians on outcomes have focused on heterogeneity in effects along two dimensions– other characteristics of the candidate and characteristics of different states. For instance, Prakash et al. (2019) show that the negative effects of criminal politicians on economic growth are more pronounced where the MLA is accused of serious charges, financial charges, is not aligned with the ruling party, is less educated and has below median wealth. They also find worse outcomes for criminal politicians in states with low levels of development and in states with high corruption. Cheng and Urpelainen (2019) and Nanda and Pareek (2016) also report worse outcomes for criminal politicians not belonging to the ruling party when looking at local road construction and private investment respectively.

In the context of provision of higher education institutions, I look at three potential sources of heterogeneity based on candidate characteristics– the education level of candidate, whether the candidate belongs to the party of the Chief Minister and thirdly, whether the candidate was an incumbent who was re-elected. The reason for specifically looking at education of MLAs is the evidence in literature that educated politicians

might care more more education provision. Lahoti and Sahoo (2020a) find that MLAs with atleast a graduate degree are more likely to invest in education funding and infrastructure but only in states with high initial levels of development.

I split the sample by education level of the winning MLAs– into those who atleast have a graduate degree and those that do not have a graduate degree and implement the RDD separately for these samples. The results are shown in the first panel of Table 3.9. I find that the RD estimate is insignificant for candidates with a graduate degree however the estimates are not much different in magnitude in sub sample of graduate MLAs vs less than graduate MLAs, not indicating any conclusive heterogeneity in effects of criminal politicians based on their having college education.

Next, I look at the alignment of the winning MLA with the Chief Minister's party. The impact of belonging to a ruling party can go in either direction. While alignment might act as a disciplining mechanism and put a check on criminal candidates rent seeking behaviour and encourage effort (Prakash et al., 2019), on the other hand alignment with the ruling party is what gives MLAs the opportunity to actually make a difference to their constituencies since a non-aligned MLA is unlikely to be successful in lobbying a state government that favours it's own (aligned) MLAs in distribution (Asher and Novosad, 2017[19]; Dey and Sen, 2016[20]). Aligned MLAs are also likely to have higher control over the bureaucracy, another factor contributing to their ability to affect outcomes.

In the case of public colleges, it is intuitive to think that non-aligned MLAs whether criminal or non-criminal would be less able to impact provision of higher education, so the distinctive impact of criminal MLAs may only manifest when looking at the subset of aligned winners. I check this by splitting the sample into two- one where winners are aligned with the Chief Minister's party and the other where winners are not from the Chief Minister's party and implement the RDD specification separately for these samples. The second panel from Table 3.9 shows results from these specifications. The negative impact of a criminal politician is almost 7 times as large and significant for only aligned MLAs. This finding is noteworthy as it runs contrary to what Prakash et al. (2019) find in the case impact of criminal politicians on economic growth and highlights the distinctive nature and mechanism behind the impact of politician criminality on state educational

[19]This paper shows evidence in support of better firm growth, better employment growth in India when the MLA is aligned with ruling coalition in state

[20]This paper presents evidence in favour of presence of partisan alignment in allocation of funds for NREGS in local governments in the state of West Bengal

institutions.

Next, I split the sample based on whether the winner is an incumbent who was elected again, to see if consecutively re-elected criminal politicians perform any differently. The results are shown in the bottom panel of Table 3.9. I find that the RD estimate is almost 4 times as small and insignificant for the subset of re-elected incumbents. This might suggest that re-election might mitigate the undesirable behaviour on part of criminal politicians or alternatively it might mean that the subset of criminal politicians who are elected again are a better performing type.

I also check if the results vary based on the development and corruption level of states. First, I check if the negative effects on state colleges are amplified in the case of BIMAROU states- an acronym for the states of Bihar, Jharkhand, Odisha, Uttar Pradesh, and Uttarakhand which are considered to be at the bottom of the development ladder in India. Secondly, I use the same definition of least developed states as Prakash et al. (2019) who base it on such classification by the Ministry of Finance[21] and check if these states report any worse outcomes. Thirdly, I look at the subset of high corruption states[22].

These results from these regressions are reported in Table 3.10. Interestingly, I do not find any worse outcomes for constituencies with serious criminal politicians in BIMAROU states, least developed states or high corruption states– in a departure from previous research (Prakash et al., 2019). This, in fact, lends credence to the theory of low effort driving lower provision of colleges by criminal politicians. A lack of lobbying effort to direct state government's resources to their own constituencies is likely to produce worse outcomes, irrespective of the strength of the institutional structure of the state.

3.7 Mechanism

3.7.1 Increase in private colleges?

One of the commonly cited explanations for the detrimental effects of criminal politicians on economic and other outcomes is their tendency to engage in rent seeking behaviour at

[21]These include Arunachal Pradesh, Assam, Bihar, Jharkhand, Odisha and Uttar Pradesh

[22]I use the definition of high corruption states as in Prakash et al. (2019) who use ranking by Transparency International India on an index of corruption. These include Tamil Nadu, Haryana, Jharkhand, Assam, and Bihar

the expense of public welfare. To the extent that rent extraction opportunities are limited in the provision of government colleges, it might deter criminal politicians from pursuing efforts in this direction. However, it might be possible that the lower provision of state government colleges is compensated by increasing openings of private colleges. Private colleges generally charge higher fees than government institutes and are often criticized for illegal profiteering through charging "capitation" fees. Higher education being one of the most highly regulated sectors in post liberalisation India, the requirements of licensing and obtaining approvals for opening private colleges offers substantial room for influence by politicians (Krishnan, 2014; Kapur and Mehta, 2007).

To check if the decreased provision of state government colleges is accompanied by an increase in provision of private college[23], I use whether any private college was opened and alternatively the number of new private colleges opening up in a constituency in a year as the main outcome variable in my original regression discontinuity specification. The results from this specification are shown in Table 3.11. The regression discontinuity estimates for both definitions of the outcome variable are negative but insignificant, indicating that there is no evidence that private colleges are more likely to be provided in constituencies with leaders accused of criminal charges.[24]

3.8 Implication for human capital

In this section, I try to draw out the potential consequences of lower provision of state government funded colleges on human capital outcomes. To do so, I focus on the relationship between the number of state government funded colleges in a district and the probability of completing graduation for an individual residing in a district. I use data from a nationally representative survey– the 71st round of the National Sample Survey (NSS). This data set includes a rich set of household and individual level characteristics, including their age, sex, education, social and religious group. The analysis is limited to individuals who report having completed atleast higher secondary education and to individuals who would have started college education between the years 1991 and 2010[25].

[23] I include both private aided and private unaided colleges in my definition of private colleges.

[24] Section 3.A.1 of the appendix discusses heterogeneity in the impact of criminal politicians on private college openings based on other candidate and state characteristics.

[25] I use this period to roughly correspond with the post-liberalization era and secondly, individuals starting college after 2010 would have not completed their education by the time of the survey

The NSS 2014 data is combined with a district[26] level panel of the existing number of state government funded colleges. I run a regression of a dummy that takes value 1 if individual reports having completed graduation and 0 otherwise on the number of state government colleges in a district, after controlling for household size, religion, social group as well as individual's gender and relation to household head. I also allow for district fixed effects and fixed effects for the year of starting college. The result for this regression is shown in column 1 of Table 3.12. The coefficient of number of state government colleges is insignificant for the full sample.

However, since state government colleges offer higher education at highly subsidized and affordable costs, it is likely that their presence would matter most for lower income groups that possibly cannot afford education from private institutions. To check for this, I divide the sample into 5 quintiles based on the monthly reported household consumption expenditure from the NSS and run the regression separately for each of these quintiles.

Columns 2 to 6 of Table 3.12 display results from these regressions. As we might expect, the coefficient of number of state government colleges for the lowest consumption quintile is positive and significant indicating that state government colleges matter the most for the lowest income groups. The estimate shows that the presence of one additional state government funded college in the district is associated with an increase in probability of completing graduation by 2.5 percentage points on an average, which is roughly 6% of the probability of graduation in this quintile (average probability of graduating for this consumption quintile is about 42% conditional on having completed higher secondary education).

3.9 Discussion

While the main results conform to the previous evidence that establishes a negative impact of politician criminality on growth and welfare, the finer picture emerging from the heterogeneity exercises emphasizes the need to distinguish between different types of outcomes. For aggregate economic growth, it is plausible that the detrimental impact of criminal politicians works by reducing private investment and by encouragement of criminal activity that further dampens private consumption and investment.

[26]The NSS only reports the district of the household which does not allow for identification of location at a more granular (constituency/village) level.

This negative effect is mitigated when the criminal politician belongs to the ruling party in case of both economic growth (Prakash et al., 2019) and private investment (Nanda and Pareek, 2016). It is suggested in both cases that oversight by the ruling party might have the effect of restraining abuse of power and rent seeking activities by criminal politicians. For an outcome like state run higher education institutes, however, belonging to the ruling party might be necessary to be able to have an impact. It is plausible that state governments favour their own aligned MLAs in allocation of education, a fact supported by evidence in Asher and Novosad (2017). So a distinction in the impact of criminal and non-criminal MLAs is only visible within the subset of MLAs who are aligned with the Chief Minister's party and thus have the power to *successfully* lobby for colleges or propose for setting up new colleges in their constituencies.

The negative impact on state government colleges fits with both the hypotheses of rent seeking behaviour and lack of political effort advanced as explanations for the effects of criminal politicians. State government education institutions might not offer as much opportunities for rent extraction such as say mining regulation, where it has been shown that booms in mineral prices increase the likelihood of election of politicians with serious charges (Asher and Novosad, 2018). This makes it likely that criminal politicians ignore the area of higher education provision instead concentrating on more lucrative projects.

Similarly, the lack of political effort might also be a possible contributing factor. Gehring et al. (2019) show that Members of Parliament (MPs) with criminal charges exhibit lower rates of attendance in Parliament and lower rates of utilization of local development funds. If the same holds true for state level politicians with criminal charges, then it might translate into lower efforts in the direction of provision of higher education institutions. Since evidence from media reports suggests that bringing new government colleges involves considerable lobbying effort by MLAs, lower effort displayed by criminal politicians would mean lesser government colleges. This would also be true if politicians with serious charges value education less as compared to non-criminal politicians, however I do not find any variation in effect of politicians accused of serious crimes depending on whether they were themselves college educated.

The distinction between the results for serious criminal candidates who were re-elected against serious criminal candidates who were not re-elected reveals another interesting variation. Re-elected criminal politicians might make more of an effort to reward their

constituencies by offering education institutions in line with theories of political patronage. Alternatively, re-election in itself might be an indicator of better performance of politicians in other respects such as being able to attract resources (including for educational institutes) towards their own constituencies, even though they might be accused of serious crimes.

Finally, the positive and robust association between the number of state government funded colleges and the probability of attaining higher education for the lowest consumption quintiles suggest that such behaviour by criminal politicians might be the most detrimental to the most economically vulnerable sections of the society.

3.10 Conclusion

This paper looks at the impact of politician criminality on the hitherto unexplored area of a government provided good- public higher education institutions in India. Using a regression discontinuity design with close elections between candidates charged with serious crimes and candidates not charged with serious crimes in the context of elections to state legislative assemblies, I find that constituencies which elect criminally charged politicians are almost 4% less likely to witness the opening of a new state government college in a year. This result is robust to a variety of alternative specification and addition of fixed effects and covariates.

The reduced provision of state government colleges in constituencies with criminal MLAs is not offset by any accompanying increase in the provision of private colleges. With government colleges accounting for 33% of total enrollment in colleges and state government colleges constituting the majority of government colleges, this is likely to have a consequent impact on human capital outcomes. While it would been insightful to examine the direct impact of criminal politicians on actual higher educational attainment of the population, the lack of such data at the constituency level does not permit a mapping from MLA action to constituency level higher education status. In absence of such possibilities, this paper highlights one important route through which state representatives *could* impact higher education outcomes.

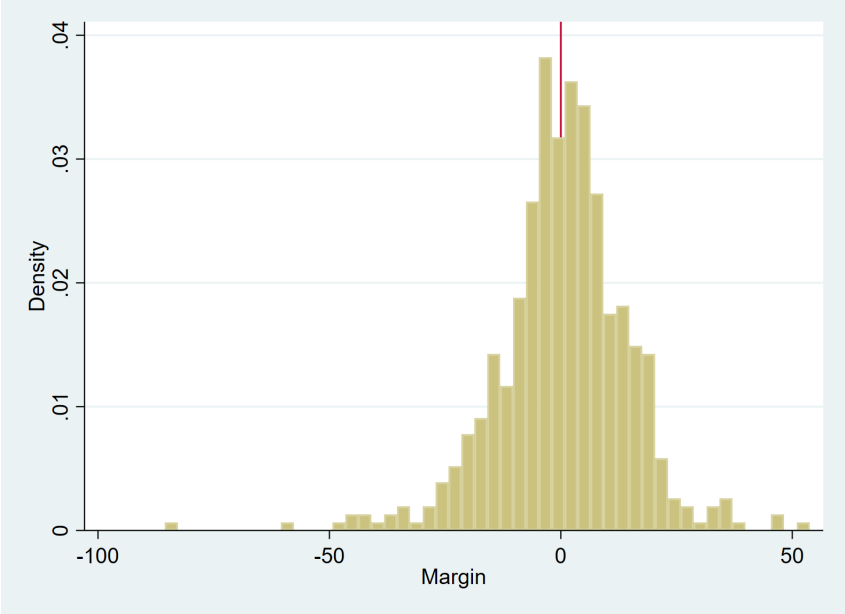
While the main result reaffirms that criminal politicians adversely impact outcomes, the heterogeneity in the effects based on other politician characteristics shines light on the

distinctive nature of higher education institutions as a state supplied good. Contrary to the evidence for aggregate economic growth and private investment, the effect of criminal MLAs on provision of state colleges is more pronounced in the case when such MLAs belong to the party of the Chief Minister. This highlights the importance of access to power to actually have the opportunity to affect outcomes, which only happens when MLAs belong to the ruling party. Also, the negative impact is more pronounced for serious criminal MLAs who were not re-elected which might indicate that such criminal politicians were re-elected as they had been performing well or could be due to better performance induced by re-election.

More broadly, this paper adds another dimension to the narrative of criminal politicians and especially criminal politicians with serious charges being detrimental to overall welfare by extending the scope of analysis to the sphere of higher education. It also shines light on the fact that such deleterious impact might disproportionately affect the poorest households which cannot afford education from private institutions. With higher education colleges constituting a capital good that continue to affect enrollment and education attainment in constituencies beyond the period of an MLA's electoral term, the results in this paper present another potential channel through which the effect of criminal politicians on economic outcomes might operate.

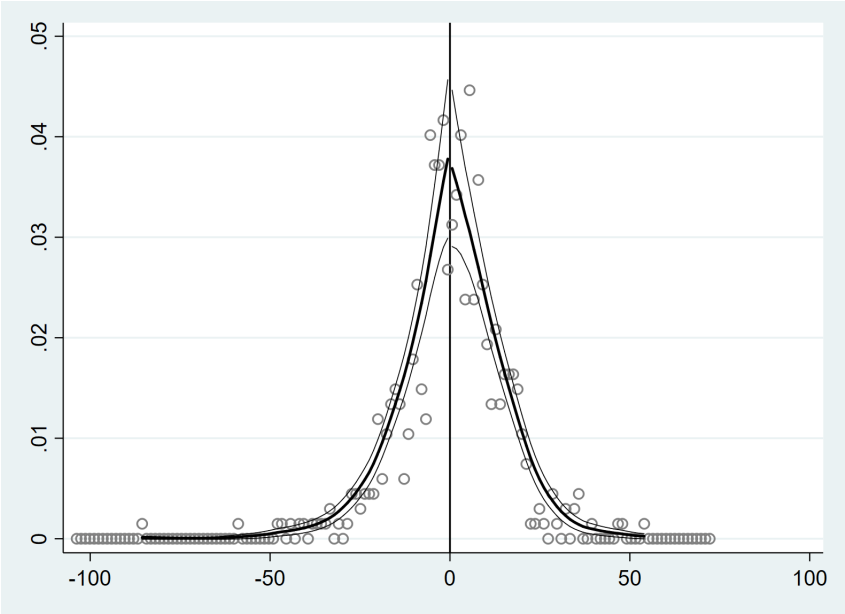
Figures and Tables for Chapter 3

Figure 3.1: Validity check- Continuity in density of forcing variable



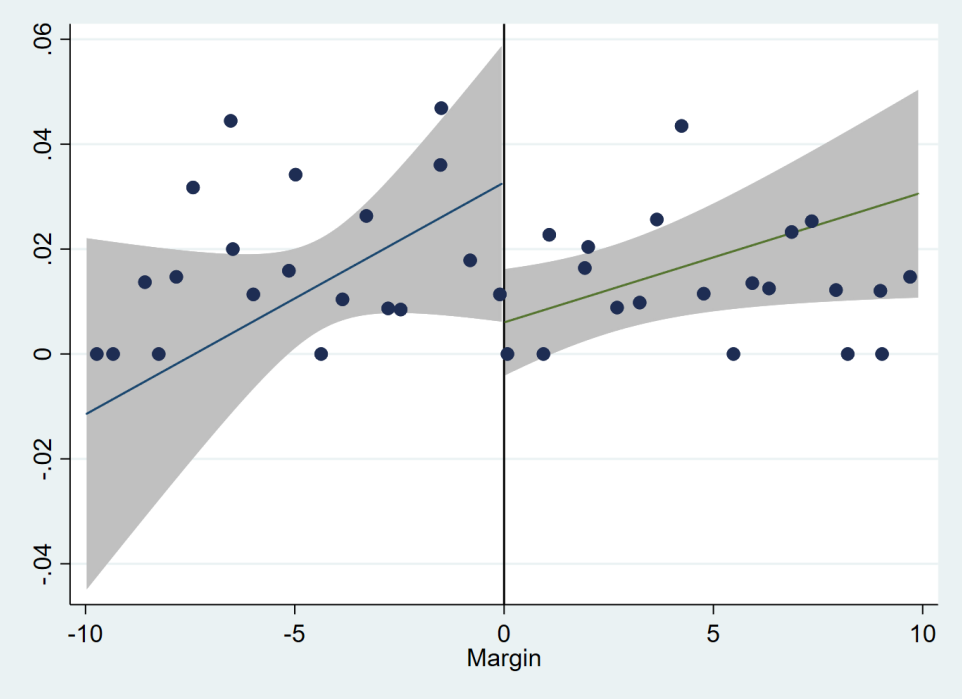
Histogram plotting the density of the vote margin of the candidate accused of serious crime

Figure 3.2: Validity check- McCrary Test



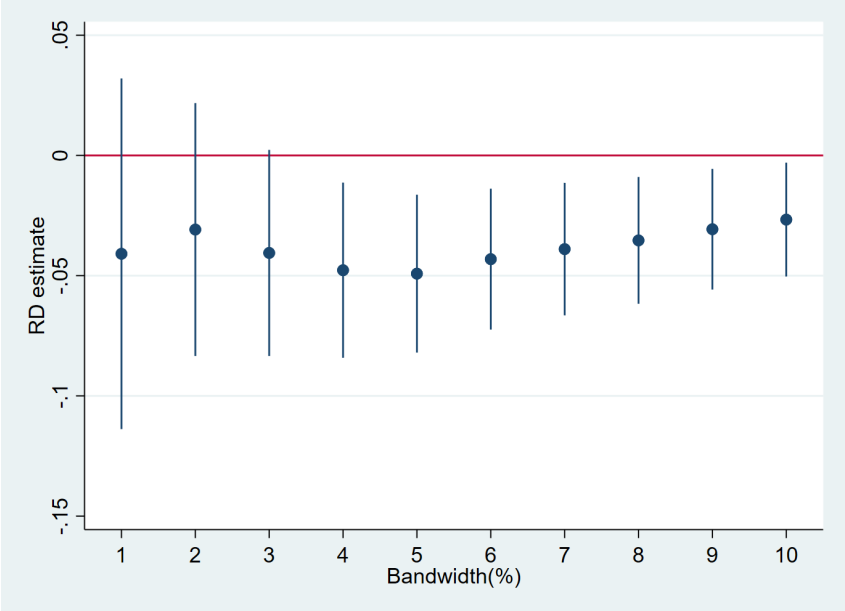
The estimated size of the discontinuity in the density of the margin at the threshold value of zero is 0.03979 with a standard error of 0.16152

Figure 3.3: Effect of electing serious criminal politicians on probability of receiving a new state govt college



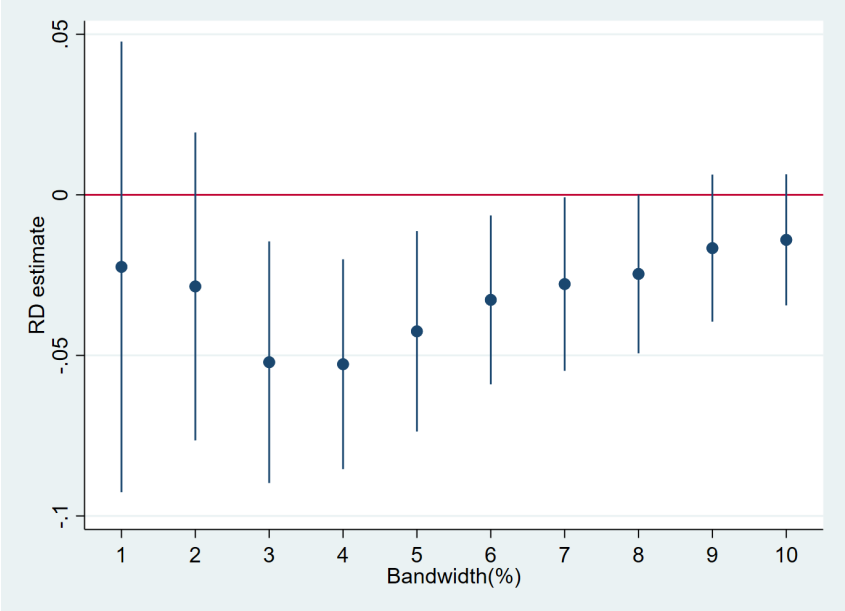
The margin of victory of the candidate accused of serious crime is the forcing variable. Positive values of margin denote where candidates accused of serious crime won against candidates not accused of any serious crime. Negative values of margin denote candidates accused of serious crimes losing against candidates not accused of any serious crime. The solid lines indicate the local linear regression fit separately on either side of the threshold value of margin which is zero with observations weighted using a triangular kernel. The dots in the scatter plot denote the average probability of receiving a state government college within margin bands of 0.5%.

Figure 3.4: RD estimates with varying bandwidths estimated using triangular kernel



This figure plots regression discontinuity estimates using triangular kernel with the bandwidth being varied from 1% to 10% on either side of the threshold of zero along with 90% confidence intervals.

Figure 3.5: RD estimates with varying bandwidths estimated using rectangular(uniform) kernel



This figure plots regression discontinuity estimates using uniform kernel with the bandwidth being varied from 1% to 10% on either side of the threshold of zero along with 90% confidence intervals.

Table 3.1: List of states and election years included in sample

State name	Election year
Sikkim	2004
Arunachal Pradesh	2004
Odisha	2004
Maharashtra	2004
Haryana	2005
Bihar	2005
Jharkhand	2005
Assam	2006
West Bengal	2006
Kerala	2006
Tamil Nadu	2006
Puducherry	2006
Himachal Pradesh	2007
Punjab	2007
Uttarakhand	2007
Uttar Pradesh	2007
Manipur	2007
Gujarat	2007
Goa	2007

Table 3.2: Descriptive statistics

	Full sample			Criminal vs non-criminal top 2		
	N	Mean	SD	N	Mean	SD
Whether any new state govt college	11649	0.016	0.127	2442	0.014	0.119
No. of new state govt colleges opened	11649	0.017	0.134	2442	0.015	0.124
No. of existing state govt collges	2362	0.712	1.373	497	0.531	0.877
No. of existing colleges (total)	2362	4.670	7.071	497	4.728	6.218
Log electors (previous election)	2290	11.947	0.552	475	12.058	0.451
Turnout percentage (previous election)	2290	64.369	10.998	475	63.536	11.097
Winner from ruling party (previous election)	2290	0.315	0.465	475	0.274	0.446
Winner incumbent (previous election)	2290	0.318	0.466	475	0.295	0.456
Winner male (previous election)	2290	0.930	0.255	475	0.943	0.232
SC reserved	2359	0.156	0.363	497	0.107	0.309
ST reserved	2359	0.070	0.254	497	0.040	0.197

The left panel gives summary statistics for the full sample irrespective of whether the winner or runner up was accused of serious crime. The right panel shows the summary statistics for the regression discontinuity sample where the winner is accused of serious crimes but the runner up is not and vice versa. This sample includes 2442 constituency year observations and 497 unique constituencies. The variables in the top two rows indicate the summary statistics for constituency year observations whereas the the rest of the variables pertain to the pre-determined characteristics and are therefore constituency specific. The state of Jharkhand had its first election in the year 2005 which is why there are no winner characteristics available for this state for the previous election and the number of observations drop from 497 to 475.

Table 3.3: Balance tests

	Serious crime	No serious crime	Difference	RD estimate	RD estimate (fixed bandwidth)
Whether any new state govt college (previous term)	0.010	0.006	0.003	-0.005	-0.009
	(0.098)	(0.079)	(0.004)	(0.007)	(0.007)
Observations	1331	1111	2442		
No. of existing state govt colleges	0.511	0.556	-0.045	-0.255	-0.286
	(0.876)	(0.880)	(0.079)	(0.205)	(0.232)
No. of existing colleges (total)	4.625	4.853	-0.228	-1.888	-1.965
	(6.800)	(5.444)	(0.561)	(1.446)	(1.614)
Log electors (previous election)	12.027	12.095	-0.069	0.135	0.133
	(0.506)	(0.373)	(0.042)	(0.084)	(0.084)
Turnout percentage (previous election)	63.133	64.020	-0.887	-3.785	-4.291
	(10.776)	(11.476)	(1.023)	(2.417)	(3.185)
Winner from ruling party (previous election)	0.236	0.319	-0.084*	0.002	-0.008
	(0.425)	(0.467)	(0.041)	(0.098)	(0.124)
Winner incumbent (previous election)	0.274	0.319	-0.045	0.043	0.057
	(0.447)	(0.467)	(0.042)	(0.106)	(0.116)
Winner male (previous election)	0.942	0.944	-0.002	0.060	0.054
	(0.234)	(0.230)	(0.021)	(0.058)	(0.066)
SC reserved	0.110	0.102	0.008	0.114	0.105
	(0.314)	(0.304)	(0.028)	(0.074)	(0.083)
ST reserved	0.048	0.031	0.017	0.043	0.043
	(0.214)	(0.174)	(0.018)	(0.038)	(0.035)
Observations	272	225	497		
Bandwidth type				MSE Optimal	7.248

The sample includes constituencies where the winner was accused of serious crime and the runner up was not and vice versa. The first two columns give the mean and standard deviation for the constituencies where the winner was accused of serious crime and where the winner was not accused of serious crime respectively. The third column shows the difference of means and the standard error of the difference in parentheses. The fourth column gives the discontinuity estimate for each of the variables used as the dependent variable in the regression discontinuity specification with the optimal bandwidth being determined separately for each estimate based on the procedure given by CCT. The last column gives the regression discontinuity estimate for each of these variables with the bandwidth fixed at 7.248 which is the optimal bandwidth for the main regression with new state govt college openings as outcome. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Impact of politicians accused of serious crime on state govt college openings

	Whether any new state govt college opened in year		
	(1)	(2)	(3)
RD Estimate	-0.0380**	-0.0454*	-0.0237*
	(0.0166)	(0.0236)	(0.0137)
Observations	1158	641	1567
Bandwidth size	7.2478	3.6240	10.8720
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)*0.5	MSE optimal (CCT)*1.5

All standard errors are clustered at the constituency level. RD estimate is the regression discontinuity estimate from using a local linear regression with a triangular kernel. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Varying bandwidth with triangular kernel

	Whether any new state govt college opened in year									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RD Estimate	-0.0409	-0.0308	-0.0405	-0.0477**	-0.0492**	-0.0431**	-0.0389**	-0.0353**	-0.0307**	-0.0267*
	(0.0443)	(0.0320)	(0.0261)	(0.0221)	(0.0200)	(0.0178)	(0.0167)	(0.0160)	(0.0153)	(0.0144)
Observations	133	296	507	699	842	1051	1145	1235	1382	1490
Bandwidth size	1	2	3	4	5	6	7	8	9	10
Bandwidth type	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual
Kernel	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Polynomial order	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

All standard errors are clustered at the constituency level. Each column denotes regression discontinuity estimates from a separate local linear regression with triangular kernel with bandwidths varying from 1% to 10% of vote margin on either side of the threshold. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Varying bandwidth with uniform kernel

	Whether any new state govt college opened in year									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RD Estimate	-0.0224	-0.0285	-0.0521**	-0.0527***	-0.0425**	-0.0327**	-0.0278*	-0.0246	-0.0166	-0.0140
	(0.0427)	(0.0291)	(0.0229)	(0.0199)	(0.0190)	(0.0160)	(0.0164)	(0.0150)	(0.0139)	(0.0124)
Observations	133	296	507	699	851	1051	1145	1240	1382	1490
Bandwidth size	1	2	3	4	5	6	7	8	9	10
Bandwidth type	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual	Manual
Kernel	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
Polynomial order	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear

All standard errors are clustered at the constituency level. Each column denotes regression discontinuity estimates from a separate local linear regression with uniform kernel with bandwidths varying from 1% to 10% of vote margin on either side of the threshold. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Higher order polynomials

	Whether any new state govt college opened in year			
	(1) Linear	(2) Quadratic	(3) Cubic	(4) Quartic
RD Estimate	-0.0380**	-0.0477**	-0.0490**	-0.0559**
	(0.0166)	(0.0195)	(0.0206)	(0.0236)
Observations	1158	1616	2132	2235
Bandwidth size	7.2478	11.8543	19.3362	21.8022
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)	MSE optimal (CCT)	MSE optimal (CCT)
Kernel	Triangular	Triangular	Triangular	Triangular
Polynomial order	Linear	Quadratic	Cubic	Quartic

All standard errors are clustered at the constituency level. Each column denotes regression discontinuity estimates from a separate regression with triangular kernel but with varying degree of the polynomial function of the forcing variable allowed on either side of the threshold. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Robustness checks

	(1) Adding fixed effects	(2) Adding existing no. of colleges as control	(3) Electoral term wise	(4) No. of new state colleges	(5) Excluding election years
RD Estimate	-0.0212*	-0.0273**	-0.1870**	-0.0333*	-0.0443**
	(0.0128)	(0.0136)	(0.0817)	(0.0170)	(0.0199)
Observations	1592	1499	238	1218	991
Year fixed effects	Yes	Yes	No	No	No
State fixed effects	Yes	Yes	No	No	No
Controls	No	Yes	No	No	No
Bandwidth size	11.2369	10.2293	7.3515	7.6887	7.6773
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)	MSE optimal (CCT)	MSE optimal (CCT)	MSE optimal (CCT)

All standard errors are clustered at the constituency level. Each column denotes results from a separate local linear regression. The dependent variable in columns 1, 2, 3 and 5 is a dummy that takes value 1 if a new state government funded college was opened in the year and 0 otherwise. The first column shows the RD estimate after adding year and state fixed effects to the main RD specification. Column 2 shows the RD estimate after adding the existing number of colleges as a control in the main specification. Column 3 collapses the data set into electoral term wise observations and shows the RD estimate from this specification. Column 4 uses the number of new state government colleges as the dependent variable. Column 6 shows the RD estimate from the original specification after excluding election year observations. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Impact of politicians accused of serious crime on state govt college openings by other candidate characteristics

	Whether any new state govt college opened in year	
	(1) Graduate	(2) Not graduate
RD Estimate	-0.0299	-0.0438**
	(0.0286)	(0.0205)
Observations	404	784
Bandwidth size	7.5851	8.3405
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new state govt college opened in year	
	(1) Belongs to CM party	(2) Does not belong to CM party
RD Estimate	-0.1122**	-0.0160
	(0.0473)	(0.0126)
Observations	345	803
Bandwidth size	5.7498	7.8207
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new state govt college opened in year	
	(1) Incumbent	(2) Not incumbent
RD Estimate	-0.0086	-0.0351**
	(0.0151)	(0.0157)
Observations	265	1163
Bandwidth size	6.0058	12.8026
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)

All standard errors are clustered at the constituency level. Each column denotes results from a separate local linear regression. The first panel estimates the regression discontinuity separately for winners who have a graduate degree and winners who do not have a graduate degree. The second panel gives RD estimates separately for samples where winners belong to the Chief Minister's party and where winners do not belong to the Chief Minister's party. The third panel gives the RD estimates where winners were incumbents and separately for where winners were not incumbents. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.10: Impact of politicians accused of serious crime on state govt college openings by state characteristics

	Whether any new state govt college opened in year	
	(1) BIMAROU states	(2) Non-BIMAROU states
RD Estimate	-0.0176	-0.0331
	(0.0184)	(0.0231)
Observations	555	838
Bandwidth size	7.8485	9.8430
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new state govt college opened in year	
	(1) Least developed states	(2) Other than least developed states
RD Estimate	-0.0165	-0.0361
	(0.0176)	(0.0244)
Observations	569	770
Bandwidth size	8.1778	8.8844
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new state govt college opened in year	
	(1) High corruption states	(2) Other than high corruption states
RD Estimate	-0.0337	-0.0430**
	(0.0385)	(0.0187)
Observations	365	805
Bandwidth size	7.9337	6.7255
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)

All standard errors are clustered at the constituency level. Each column denotes results from a separate local linear regression. The first panel estimates the regression discontinuity separately for BIMAROU states which include Bihar, Chhattisgarh, Jharkhand, Orissa, Uttar Pradesh and Uttarakhand and non-BIMAROU states separately. The second panel gives RD estimates separately for least developed states which include Arunachal Pradesh, Assam, Bihar, Jharkhand, Odisha and Uttar Pradesh. The third panel gives the RD estimates separtately for high corruption states which include Tamil Nadu, Haryana, Jharkhand, Assam, and Bihar. The definitions of least developed states and high corruption states are as in Prakash et al. (2019) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Impact of politicians accused of serious crime on private college openings

	Whether any new private college opened in year		
	(1)	(2)	(3)
RD Estimate	-0.0674	-0.0430	-0.0671
	(0.0772)	(0.1102)	(0.0633)
Observations	1402	768	1770
Bandwidth size	9.3102	4.6551	13.9653
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)*0.5	MSE optimal (CCT)*1.5
	Number of new private college opened in year		
	(1)	(2)	(3)
RD Estimate	-0.3467	-0.4091	-0.2956*
	(0.2239)	(0.3614)	(0.1742)
Observations	1392	754	1745
Bandwidth size	9.0547	4.5274	13.5821
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)*0.5	MSE optimal (CCT)*1.5

All standard errors are clustered at the constituency level. RD estimate is the regression discontinuity estimate from using a local linear regression with a triangular kernel. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.12: Probability of graduating

	(1) Full sample	(2) Quintile 1	(3) Quintile 2	(4) Quintile 3	(5) Quintile 4	(6) Quintile 5
No. of state govt colleges	0.000	0.025**	0.004	-0.003	-0.003	-0.001
	(0.003)	(0.012)	(0.011)	(0.008)	(0.006)	(0.004)
Mean	0.56	0.42	0.46	0.49	0.56	0.63
Observations	24295	1786	2634	3791	6184	9895
Household and individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Year of starting fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at the district level and given in parentheses. Each column denotes results from a separate regression. The dependent variable is whether individual completes graduation. Sample is limited to individuals who had finished at least higher secondary education. Data sourced from the 71st round of the NSS held in 2014. Quintiles refer to quintiles calculated on the basis of the monthly household consumption expenditure, where quintile 1 refers to the lowest consumption expenditure group. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.A Appendix to Chapter 3

3.A.1 Heterogeneity analysis for private college openings

In this section, I check for heterogeneity in the impact of politicians with serious criminal charges based on other candidate characteristics and based on state characteristics. In terms of candidate characteristics, just as in the case of state government colleges, I look at potential sources of heterogeneity based on whether the candidate was graduate or not, whether the candidate belongs to the party of the Chief Minister and based on whether the candidate was an incumbent who was re-elected. Table 3.A2 shows the results of the regression discontinuity estimation. Interestingly, contrary to the results for state government colleges, I find that the impact of candidate criminality is worse when the candidate is not a graduate. However, I find no distinction in the regression discontinuity estimate based on the alignment or incumbency status of the candidate.

Next, I check whether the impact of politician criminality varies based on the development and corruption level of the states. Table 3.A3 shows the results for this estimation exercise. Contrary to the results for state government college openings, but in line with the results for aggregate economic growth in Prakash et al. (2019), I find that the impact of criminal politicians on private colleges is amplified for BIMAROU and least developed states. On the other hand, the RD estimates for high corruption and other than high corruption states are similar.

3.A.2 Additional figures and tables

Figure 3.A1: Year wise average no. of new state government funded colleges opened

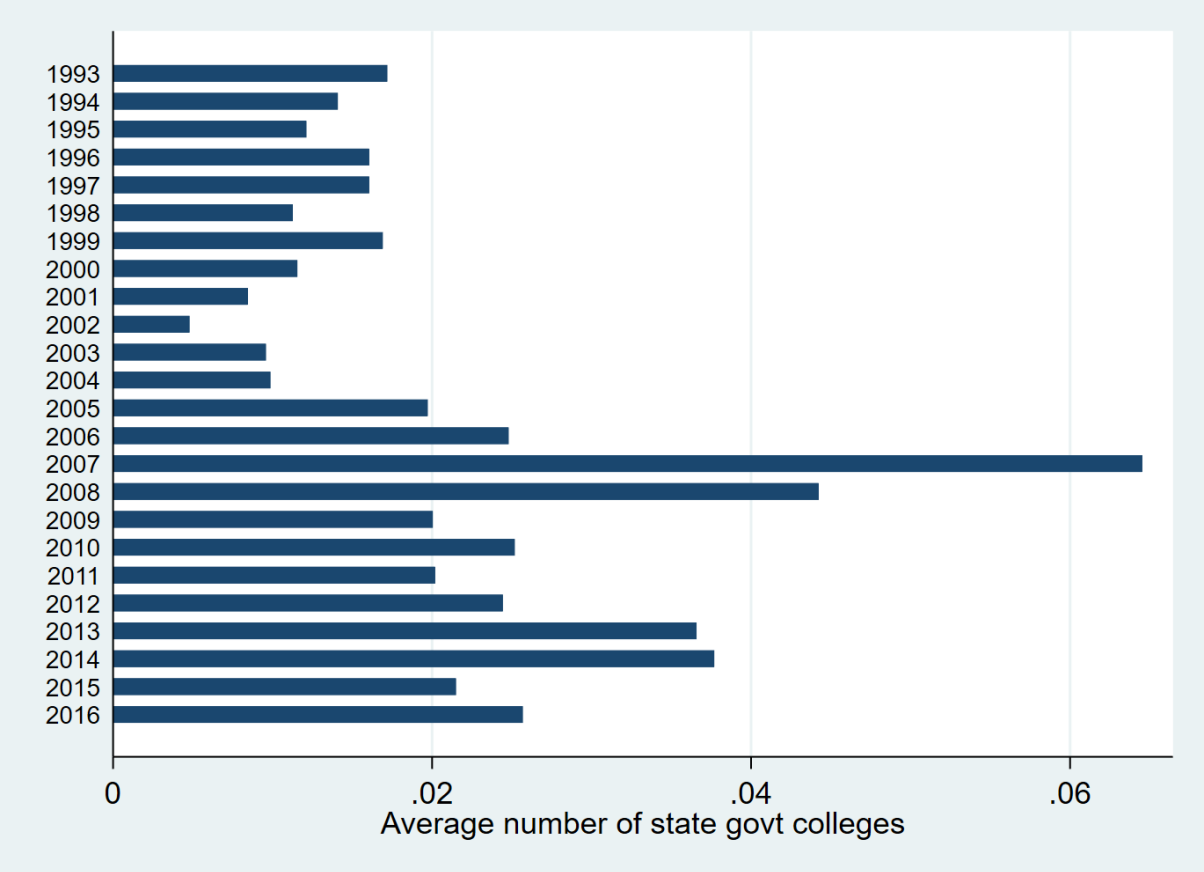


Table 3.A1: Balance test- Full sample

	Serious crime	No serious crime	Difference
Any state govt college opened (previous term)	0.008	0.012	-0.004
	(0.089)	(0.107)	(0.003)
Observations	1870	9779	11649
No. of existing state govt colleges	0.556	0.742	-0.186*
	(0.989)	(1.433)	(0.077)
No. of existing colleges (total)	4.772	4.650	0.121
	(6.717)	(7.138)	(0.396)
Log electors (previous election)	12.057	11.926	0.131***
	(0.457)	(0.566)	(0.031)
Turnout percentage (previous election)	62.850	64.656	-1.806**
	(10.647)	(11.042)	(0.628)
Winner from ruling party (previous election)	0.272	0.323	-0.051
	(0.446)	(0.468)	(0.027)
Winner incumbent (previous election)	0.319	0.318	0.000
	(0.467)	(0.466)	(0.027)
Winner male (previous election)	0.942	0.928	0.014
	(0.233)	(0.259)	(0.015)
SC reserved	0.092	0.168	-0.076***
	(0.289)	(0.374)	(0.020)
ST reserved	0.058	0.072	-0.014
	(0.234)	(0.258)	(0.014)
Observations	381	1981	2362

This table gives the means and difference in means for pre-determined characteristics of constituencies where the elected candidate was accused of serious crime as opposed to constituencies which did not elect a candidate accused of serious crime, irrespective of the criminal status of the runner up. Standard errors in parentheses.

Table 3.A2: Impact of politicians accused of serious crime on private college openings by other candidate characteristics

	Whether any new pvt college opened in year	
	(1)	(2)
	Graduate	Not graduate
RD Estimate	0.0647	-0.1652*
	(0.1420)	(0.0988)
Observations	497	722
Bandwidth size	10.5739	7.5002
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new pvt college opened in year	
	(1)	(2)
	Aligned with CM party	Not aligned with CM party
RD Estimate	-0.1023	-0.0478
	(0.1504)	(0.0837)
Observations	511	831
Bandwidth size	9.6500	8.1015
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new pvt college opened in year	
	(1)	(2)
	Incumbent	Not incumbent
RD Estimate	-0.1023	-0.0657
	(0.1249)	(0.1010)
Observations	384	895
Bandwidth size	10.3892	8.1889
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)

Table 3.A3: Impact of politicians accused of serious crime on private college openings by state characteristics

	Whether any new pvt college opened in year	
	(1)	(2)
	BIMAROU states	Non-BIMAROU states
RD Estimate	-0.2163*	0.0602
	(0.1213)	(0.0896)
Observations	642	750
Bandwidth size	10.1664	8.5508
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new pvt college opened in year	
	(1)	(2)
	Least developed states	Other than least developed states
RD Estimate	-0.2331*	0.0803
	(0.1224)	(0.0901)
Observations	652	745
Bandwidth size	10.0274	8.6016
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)
	Whether any new pvt college opened in year	
	(1)	(2)
	High corruption states	Other than high corruption states
RD Estimate	-0.0341	-0.0754
	(0.1251)	(0.0897)
Observations	444	1035
Bandwidth size	10.2371	9.8476
Bandwidth type	MSE optimal (CCT)	MSE optimal (CCT)

Chapter 4

Educated Leaders through Legislation but at what Cost?

4.1 Introduction

Recent literature has shown that the characteristics of leaders matter for development outcomes (Dreher et al., 2009; Alesina et al., 2015; Amaral et al., 2014). Within this literature, special emphasis has been accorded to the educational attainment of leaders. The education level of leaders is widely considered to be a valid proxy of their quality and is often used in empirical research as a measure of the same (Besley et al., 2005; Baltrunaite et al., 2014). For instance, it has been shown that educated leaders lead to higher economic growth (Besley et al., 2011) and can also improve the educational attainment of citizens (Diaz-Serrano and Pérez, 2013; Lahoti and Sahoo, 2020b).

Based on the premise that educated leaders are better at understanding and responding effectively to citizens' needs, communities have tried to introduce measures to promote the election of educated leaders and thereby improve the quality of their governance. One of the ways to achieve this is by incentivizing more educated candidates to contest elections (Kotakorpi and Poutvaara, 2011; Ferraz and Finan, 2011). Another way– which forms the backdrop of this paper, is by mandating minimum education requirements for candidates contesting elections.

We study such a restriction on candidature in the context of a large state in India– Rajasthan, where a legislation enacted in 2014 imposed minimum education requirements

for candidates contesting local elections in 2015.[1] While having leaders with education can lead to gains in aggregate efficiency if educated leaders perform better in terms of provision of public goods, there could be potential costs in terms of other desirable characteristics of leaders. For instance, such restrictions could adversely impact representation from historically disadvantaged communities or could disqualify more experienced yet uneducated candidates from elections.

Given the potential of such minimum education requirements to affect leader characteristics other than education, we investigate the impact of the legislation mandating minimum education for candidates on the caste category and gender of elected leaders in the context of local village council elections in Rajasthan. More specifically, using temporal data on elections held before the legislation (in 2010) and after the legislation (in 2015), we ask whether the probability of elected leaders belonging to disadvantaged groups changes in areas where the legislation hit the hardest, after accounting for reservation status. While the stated objective of introducing these education based restrictions on candidature is to improve the efficiency and accountability of local leaders and to curb corruption, these laws could be criticized for being undemocratic, discriminatory, elitist and disenfranchising for the reason that they would disproportionately affect candidates belonging to historically disadvantaged groups and women, both of which are likely to have lower levels of education (Pande, 2015; Jaffrelot, 2016; IndianExpress, 2014; BusinessLine, 2014; HindustanTimes, 2014).

To quantify the impact of legislation, we compare electoral districts (in this case, village councils or Gram Panchayats) where the existing incumbent elected in 2010 did not meet the education criteria instituted by the legislation to Gram Panchayats where the incumbent met the education criteria. The Gram Panchayats which did not have an educated incumbent as defined by the legislation criteria would presumably be the ones to witness the maximum impact of legislation[2].

It should be noted that, intuitively, the minimum education legislation should have no impact on say, the gender of the elected Sarpanch if the representative female candidate who did not meet the education criteria was supplanted by another female candidate who met the education criteria. Similarly, if the representative OBC candidate who was

[1] This was followed by the states of Haryana and Assam enacting similar legislation in 2015 and 2018 respectively.

[2] Discussed in detail in the section on empirical strategy

disqualified was replaced by another OBC candidate, we should not expect any changes in caste of the Sarpanch due to legislation. However, introduction of such a criteria in a context where disparities in education by gender and caste are dominant is likely to lead to exclusion of groups where educational attainment remains historically low. This is because not only are certain groups more vulnerable to disqualification, but it is also more likely that their replacements would belong to other "forward" groups.

Secondly, we examine if the introduction of this legislation actually led to an improvement in performance of the elected leader again by comparing Gram Panchayats on the basis of binding impact of the legislation by focusing on three important outcomes: first, the work generated under the NREGS (National Rural Employment Guarantee Scheme) which is one of the most ambitious social security and public works programmes in the world today (Zimmermann, 2020); second, economic activity as proxied by night time luminosity and finally, improvement in human capital as measured by school openings.

This paper employs a difference-in-differences strategy to measure the impact of the minimum education legislation on the selection and performance of the leader. Using data from elections held before introduction of legislation in 2010 and elections held post the introduction of legislation in 2015, we measure the impact of the minimum education legislation on the caste category and gender of the elected leader as well as performance of the elected leader with respect to the above mentioned outcomes. As stated before, we use the education qualification of the incumbent leader as a source of variation in the impact of the minimum education legislation. The legislation would conceivably hit the hardest in places where the existing leader did not meet the minimum education requirement.

This is based on the idea that the existing incumbent's education level is in a sense, indicative of the education of the most viable candidate in the Gram Panchayat. This is not an unreasonable assumption since the winner on an average takes 42% of the votes[3]. The minimum education restriction would hit the hardest where the most viable candidate was displaced on account of the legislation and these are the Gram Panchayats which would be most likely to witness changes in leader characteristics after accounting for reservation status. It is to be borne in mind that we do not rely on the notion that the same incumbent would or would not have contested again in the future election, but

[3]Alternatively, we also show the same results after considering the pool of uneducated candidates before the legislation as the potential source of variation in impact of legislation in the robustness checks.

simply hypothesize that Gram Panchayats with unqualified Sarpanches seen as indicative of the average Sarpanch candidate should see the maximum impact of legislation.

Next we focus on actual outcomes to evaluate whether educated leaders instated through this legislation performed better. To fully appreciate the import of this question, one must recognize that the impact of a legislation mandating minimum education, on actual outcomes would be distinct from the impact of simply having a leader with higher education. While educated leaders can potentially lead to less corruption and improved efficiency, such a legislation could also impact other leader characteristics such as the caste and gender of the elected leader– both of which have been shown to matter for distribution of public goods and targeting of programs (Bardhan et al., 2010; Besley et al., 2004). It could also induce entry of younger candidates and less experienced candidates who are more likely to meet the minimum education threshold which could in turn affect leader performance (Afridi et al., 2017),

At the same time, it is plausible that this restriction would also lead to changes in the extent of electoral competition– a fact that has been empirically established by studies looking at the impact of other kinds of barriers to candidacy[4]. [5] Both these aspects–one that the legislation could potentially lead to changes in caste and gender of the elected leader after accounting for reservation of seats by caste and gender and second, that it could affect the extent of political competition, which in turn has consequences for public goods and programs (Gupta and Mukhopadhyay, 2016; Afridi et al., 2021) imply that it would be naive to assume that such a legislation would automatically and necessarily improve leader performance.

Combining data on caste categories of elected leaders from elections held before and after the legislation with information on the education level the of incumbent leader at the time when legislation was introduced, estimates from the difference-in-differences

[4]Lack of availability of data about the age, experience, and vote shares and winning margin of elected leaders in the period after the introduction of the minimum education legislation prevents us from looking at these other outcomes of interest

[5]Afzal (2014) looked at the impact of minimum education requirements imposed for the 2002 federal legislature election in Pakistan and found that such ballot access restrictions led to a reduction in political competition where the disqualified incumbent belonged to a minor party and where literacy levels were lower and an increase in political competition where the disqualified incumbent had a higher winning margin. Ansolabehere and Gerber (1996) looked at the impact of legal barriers to competition in the form of filing fees and petition requirements in the context of the United States and found that higher ballot access requirements are associated with increased likelihood of uncontested seats. Drometer and Rincke (2009) found that decrease in petition requirements increased competition from third party and independent candidates whereas Stratmann (2005) found that stringent petition requirements reduce the number of major party candidates.

specification show that the imposition of minimum education requirement led to a decrease in the probability of election of a leader belonging to Scheduled Tribes (ST) by 1.6 percentage points in Gram Panchayats where the incumbent did not meet minimum education criteria. The corresponding decline for a leader belonging to Other Backward Castes (OBC) was 5 percentage points whereas the probability of election of a leader belonging to the general category[6] increases by 7 percentage points. As expected, we find that the probability of election of a woman falls by 8 percentage points. Looking at the intersection of caste and gender of elected leader, there is only a fall in the probability of election of a female ST and female OBC Sarpanch, with no such observed decline for their male counterparts.

We use data from 2005 and 2010 elections to check for pre-existing trends in the election of leaders belonging to different caste categories and gender and find that we cannot reject the assumption of parallel trends in the election of ST leaders in GPs where incumbent did not meet minimum education threshold vs. GPs where incumbent met the minimum education threshold. For election of leaders from OBCs and general categories as well as for election of women leaders, we find that the pre-existing trends run opposite to the direction of our estimated effect, suggesting that the existence of these prior trends cannot explain our results.

We find no improvement in the performance of leaders as measured by work generated under NREGS, night lights or new school openings in the aftermath of introduction of the minimum education restriction. The difference-in-differences estimates for all the three outcomes indicate that there is no change in performance of the leader in Gram Panchayats where the incumbent did not meet the minimum education compared with Gram Panchayats that already had an educated leader after introduction of the legislation. This finding underscores the need for a more nuanced approach to improving leader quality, while restricting candidature based on education brings educated leaders, it does not necessarily lead to better outcomes and at the same time leads to lower representation of disadvantaged groups.

Using a difference-in-differences strategy allows us to account for any time invariant unobservable Gram Panchayat (GP henceforth) characteristics that influence the selection of the leader through addition of GP fixed effects. Any time specific unobservables that

[6]General category essentially refers to the "upper" or "forward" castes

affect all GPs in the same way are accounted for by year fixed effects. Looking at changes in GPs with leaders not meeting education requirement relative to changes in GPs where leaders met the education requirement after the introduction of the legislation and controlling for time specific and GP specific unobservable characteristics allows us to quantify heterogeneity in the impact of the legislation. We control for the reservation status of the GP in all our regressions and additionally allow for trends by baseline GP demographic variables like the population, proportions of SC and ST population and overall and female literacy rates. These allow us to account for varying trends due to initial GP characteristics that might be correlated with having an uneducated[7] leader.

Additionally, a host of robustness checks are employed to establish the strength of our results. Firstly, we introduce time trends by additional baseline measures of GP education and socio-economic status and show that the main results stay the same. Second, we use propensity score matching and implement the difference-in-differences strategy on a sample of GPs with uneducated leaders (as per the legislation criteria) matched to GPs with educated leaders based on a host of initial GP characteristics and still find that the main results of decrease in probability of election of leaders from disadvantaged groups remain. Third, we use alternative definitions of the dependent variable as well as control for trends by initial reservation status and show that the results remain stable. Fourth, we employ a placebo test where we randomly assign the education status of the incumbent Sarpanch to GPs and construct a distribution of placebo beta coefficients through replications of this exercise. The distributions show that the probability of obtaining our results through chance alone are very low. Finally, we also show some evidence for heterogeneity in effects of the legislation based on the literacy rates of GPs.

This paper adds to the literature examining the impact of various barriers to candidacy. It is most closely related to Afzal (2014) which investigates the impact of minimum education requirements for contesting elections to the federal legislature introduced in Pakistan in 2002. To our knowledge, this is the only other instance where candidacy was qualified based on education attainment other than the premise of our paper based in India. While Afzal (2014) looks at the impact of this legislation on measures of political competition, we focus on the caste and gender identity of the elected leader– an issue which is salient in context of elections and their distributive consequences.

[7]The terms "educated" and "uneducated" are used interchangeably with being Class 8^{th}(Class 5^{th} in scheduled areas) pass and not being Class 8^{th}(Class 5^{th} in scheduled areas) pass respectively.

Papers looking at effects of erecting barriers to candidacy have mostly focused on their impact on political competition. Drometer and Rincke (2009) study the impact of an exogenous decrease in petition requirements in the state of Ohio in the United States and find that it led to a significant increase in the number of minor party and independent candidates. Stratmann (2005) looks at the impact of barriers to candidacy enforced through filing fees and signature requirements in the context of Lower House elections in the United States and finds that filing fees were more effective in deterring minor party candidates from contesting elections whereas more stringent signature requirements reduce the number of major party candidates. This paper expands the narrative surrounding ballot access restrictions beyond their potential impact on electoral competition to their possible effects on the characteristics of elected leaders– an area which remains relatively unexplored.

The importance of the findings of this paper rests on the evidence highlighting distributional effects of caste and gender identity of the leader. This theme has produced an expansive body of literature in recent years that has focused on using caste and gender identities instituted through reservations (or quotas) for disadvantaged groups to examine distributional outcomes. Chattopadhyay and Duflo (2004) in their paper use random reservation of seats for women based in the context of rural local elections in West Bengal and Rajasthan and find that women leaders are more likely to invest in public goods preferred by citizens of their own gender. Other papers look at the impact of reservation for members of disadvantaged groups on distribution of public goods towards these groups (Dunning and Nilekani, 2013; Besley et al., 2004; Chin and Prakash, 2011, Bardhan et al., 2010). Das et al. (2017) look at the impact of OBC reservations in the context of rural Rajasthan and show that the reservation increased the overall provision of public goods where OBCs constituted a greater share of the total population but that these goods were not specifically targeted to OBCs.

While acknowledging that the impact of representing minorities by affirmative action through a system of caste/gender based reservations is not the same as the impact of having a leader of a particular caste/gender, the aforementioned studies do help in understanding the potential effects of change in caste of elected leader as a result of the minimum education legislation. With respect to women especially, literature has established that there is a continuing positive impact of quotas for women even after they

are withdrawn on subsequent electoral participation by women (Beaman et al., 2009; Bhavnani, 2009; Gangadharan et al., 2016). Evidence also exists to show that election of women leaders even in unreserved seats encourages future political participation by women (Bhalotra et al., 2018). Given that we find that this legislation led to a decline in probability of electing women as leaders, this has possible ramifications for future electoral participation by women as well.

This paper also adds to the discourse surrounding the role of educated leaders. Previous studies have shown that educated leaders improve economic growth (Besley et al., 2011) and increase educational attainment of citizens (Diaz-Serrano and Pérez, 2013). In the context of India, Lahoti and Sahoo (2020b) show that leaders with a graduate degree positively impact enrollment, education funding and infrastructure but only in states where the level of development is high to begin with. While these existing studies tell us why educated leaders may be beneficial, we augment this literature by showing that forcing a minimum education requirement for candidates can have unintended consequences that may prove detrimental to the cause of minority representation especially in a society with unequal access to education.

The rest of the paper is organized as follows: Section 2 describes the electoral context and the minimum education legislation in detail. Section 3 presents the data sources and the descriptive statistics. Section 4 outlines the empirical strategy and Section 5 presents results from the empirical strategy. Section 6 presents the robustness checks. Section 7 discusses heterogeneity in results. Section 8 presents results for the impact of the legislation on night lights, NREGS and school openings and Section 9 concludes the paper.

4.2 Background

4.2.1 Village Council Elections

The 73rd Amendment (1993) to the Indian Constitution instituted a three-tier rural governance structure in all states referred to as the Panchayati Raj System. The Village Council or Gram Panchayat (GP) essentially forms the lowest level of this governance system with the Panchayat Samiti being the block level institution and the Zila Parishad

the district level institution[8]. The Gram Panchayat consists of a ward of representatives and a council head called the Sarpanch. While the ward members are directly elected every five years, the Sarpanch may or may not be directly elected depending on the state[9] but with a fixed term of five years. In the case of Rajasthan, the Sarpanch is directly elected along with the other ward members every five years.

The 73rd Amendment also mandated that atleast one-third of all positions of Sarpanch and other ward members would be compulsorily reserved for women[10]. It also provides for the reservation of positions of both ward members and Sarpanch for marginalized groups– the Scheduled Castes (SCs) and Scheduled Tribes (STs) in proportion to their population. Some states including Rajasthan additionally provide for reservation of positions for Other Backward Castes (OBCs) subject to the restriction that the combined reservation for SCs, STs and OBCs be less than 50 percent. In Rajasthan, the proportion of positions reserved for women was increased from one-third to half in 2008. While allocation of a GP for reservation for SCs and STs is decided on the basis of its share of SC/ST population, the reservation for women and OBCs is decided by way of a lottery. The reservation status based on both caste and gender is rotated amongst the GPs from election to election.

The primary responsibilities of the Gram Panchayat include provision of local public goods and infrastructure such as village roads, drinking water facilities, sanitation as well as the administration and targeting of public programs such as the National Rural Guarantee Scheme (NREGS). Although the Panchayati Raj Act calls for collective decision making by the Village Council members, in practice the Sarpanch wields a great deal of influence on the final allocation of public goods and selection of beneficiaries of welfare programs (Besley et al., 2012; Chattopadhyay and Duflo, 2004; Jeong et al., 2021). This provides the rationale for focusing on the changing characteristics of the Sarpanch as an important consequence of a policy restricting candidature in local village council elections.

[8]Villages are the lowest level in the hierarchy of administrative divisions in India. Each state is divided into districts, which are further divided into blocks which in turn consist of many villages.

[9]In states like Rajasthan, Madhya Pradesh, Uttar Pradesh, Bihar. Sarpanches and ward members are elected directly by the people. In other states like West Bengal, Karnataka and Kerala, the people elect ward members who then elect a leader from amongst themselves

[10]As per information available with the Ministry of Panchayati Raj, Government of India, 20 states namely Andhra Pradesh, Assam, Bihar, Chhattisgarh, Gujarat, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Sikkim, Tamil Nadu, Telangana, Tripura, Uttarakhand and West Bengal have made provisions for 50% reservation for women in Panchayati Raj Institutions in their respective State Panchayati Raj Acts.

4.2.2 Minimum Education Legislation

In December 2014, the BJP (Bharatiya Janata Party) Government in Rajasthan introduced an ordinance effecting an amendment to the Rajasthan Panchayati Raj Act 1994. This ordinance made it mandatory for those contesting Sarpanch elections to be Class 8th pass. In scheduled areas[11] where all Sarpanch positions are reserved for STs, the eligibility was fixed at Class 5th pass. Candidates contesting Zila Parishad and Panchayat Samiti polls were required to be Class 10th pass. No education requirements were instituted for contesting the posts of ward members. This announcement came a mere month before Panchayat elections were scheduled to be held in January 2015 and disqualified more than 60 percent of the existing Sarpanches from contesting elections.

The election in January 2015 was subsequently held with this minimum education restriction in place amidst widespread criticism from other political parties, social commentators and activist groups. This ordinance was passed as a bill in March 2015 with which Rajasthan became the first state in India to fix minimum education criteria for contesting Panchayat elections. The stated objective of enacting such legislation was to improve financial management by reducing embezzlement of funds at the hands of illiterate Sarpanches as well as to enhance accountability of Sarpanches given that they are responsible for managing funds allocated by public programs (IndianExpress, 2014).

However the legislation was criticized for being undemocratic, disenfranchising and discriminatory as it was most likely to disbar candidates from disadvantaged minorities as well as women from contesting elections because these groups were less likely to meet the minimum education criterion (Pande, 2015; Jaffrelot, 2016; BusinessLine, 2014). This was perceived as punishing the people for the failure of the state to provide education. Moreover, it is noteworthy that no such minimum education requirements exist for Members of Parliament(MPs) or Members of Legislative Assemblies in states (MLAs), both of which call for a much higher degree of financial responsibility and accountability.

That the minimum education legislation for candidates would be exclusionary towards the most vulnerable groups is not an unreasonable concern. The rural male literacy rate stands at 76% while the rural female literacy rate is 46% in Rajasthan as per the 2011

[11]Scheduled areas as identified in the Fifth Schedule of the Indian Constitution are regions which are predominantly inhabited by Scheduled Tribes or tribal communities. In Rajasthan, the districts of Dungarpur and Banswara were identified as fully tribal districts and the districts of Pratapgarh, Udaipur and Sirohi were partly tribal as of 2015. In 2018, this list was expanded to include additional villages in the districts of Pali, Sirohi, Rajsamand and Chittorgarh.

census. While the census does not give literacy rates by caste categories, if the pool of incumbent leaders before the legislation is taken as an indicator, 35% of leaders belonging to Scheduled Tribes (ST), 27% of leaders belonging to Scheduled Castes (SC) and 37% leaders belonging to Other Backward Castes (OBC) were atleast Class 8th pass[12]. The corresponding figure for leaders belonging to the General category on the other hand was 51% .

Minimum education requirements were subsequently also introduced for contesting Panchayat elections in the states of Haryana in 2015 and Assam in 2018. In December 2018, the new INC (Indian National Congress) government in Rajasthan scrapped the education criterion for contesting panchayat elections. This was formally revoked by a bill passed in the Rajasthan state legislative assembly in February 2019 and the next cycle of Panchayat elections were held in January 2020 without any minimum education restrictions on candidature.

4.3 Data

4.3.1 Data sources

The study uses election data from the 2010 and 2015 Panchayat General Elections in Rajasthan. The data for 2010 Sarpanch elections was manually obtained from the records of the Rajasthan State Election Commission and includes names, level of education, age, gender, caste category, vote share for each of the **top 5 candidates** in the GP. This information on education level is used to identify GPs where the incumbent Sarpanch did not meet the education criteria laid down by the minimum education legislation. This data set also carries the total number of candidates and reservation status of each GP and covers about 7600 of the 9166 GPs existing in Rajasthan at that time.

A delimitation exercise held before the elections in 2014 expanded the number of GPs to 9894 in 2015[13]. The 2015 election data comes from the report on Panchayat Election 2015 published by the Rajasthan Election Commission and made available on its website[14]. This data includes only the name, gender and caste category of the **elected**

[12]We are focusing on whether the leaders were at least Class 8th pass as that was the minimum education qualification mandated by the legislation in Rajasthan

[13]This number subsequently expanded to 11316 in 2019 on account of delimitation before the 2020 elections.

[14]http://sec.rajasthan.gov.in/StatisticsArchiveNew.aspx

sarpanch along with the reservation status of the GP and covers all of 9875 GPs where elections were held in 2015. [15]

Since delimitation entails reallocation of villages[16] to new and even existing GPs, we limit our sample to GPs which did not undergo any changes in composition on account of the 2014 delimitation exercise. This was done by manually tracking changes in constituting villages of GPs using delimitation documents published by the State Election Commission. This leaves us with about 6024 GPs in our final sample for which data is available in both 2010 and 2015 and which did not change due to delimitation. We also use Sarpanch caste, gender and reservation status data from the report on Panchayat Election 2005 published by the State Election Commission to investigate pre-legislation trends in GPs.

The data for GP level population and other demographic characteristics including the proportions of SC and ST population, proportion of female population, overall and female literacy rates as well as the number of schools is obtained from the 2011 census. Since the census reports village level data, a mapping from census village to GP was obtained from the Local Government Directory[17] to aggregate village level statistics to the GP level. Additionally the Socio-Economic Caste census (SECC) 2011 is employed to obtain data for the average educational attainment, average age and economic status of the GP population which are used for robustness checks.

Data for the performance of the NREGS (National Rural Employment Guarantee Scheme) is sourced from the official portal for the scheme [18]. This reports a range of GP wise yearly statistics on the allotment of work, actual work generated and caste and gender composition of beneficiaries under this scheme. We use the person days of work generated per capita[19] from 2012 to 2017 as the relevant metric to gauge NREGS performance. We were able to map 5403 GPs to the NREGS dataset from our original sample.

We use the Visible Infrared Imaging Radiometer Suite (VIIRS) night lights data from

[15]The paucity of data hinders analysis of other interesting dimensions like extent of political competition.

[16]A GP generally consists of as group of geographically contiguous villages subject to a broad combined population criteria

[17]https://lgdirectory.gov.in/

[18]https://www.nrega.nic.in/netnrega/mgnrega$_n ew/Nrega_h ome.aspx$

[19]Calculated by dividing the GP level total person days generated by the GP level population in 2011 census

2012-2019 sourced from the website of the NOAA (National Oceanic and Atmospheric Administration)[20] to calculate GP level proxies of economic activity. The VIIRS reports night time lights with a constant spatial resolution of 742 metres. We use GIS software and village level shapefiles to aggregate night lights by adding all reported grid light values in a GP in a year. Limiting to GPs for which we could track all constituent villages, we are left with 5600 of 6024 in our original sample. We use the per capita sum of night lights(and its logarithm) aggregated by year within the GP as the relevant measure of economic performance.

Finally, we use raw data from the DISE (District Information System for Education) from 2017 to look at year wise school openings. DISE is a database for all recognized schools in India. The raw data is a school level data set that includes details about the location, year of establishment, category (primary, middle or secondary), management (private or public) and enrollment in each school. The village names from the location field of the DISE data were matched to village names from the census. This DISE to census mapping was used to generate a DISE to GP mapping by using the village to GP mapping from the Local Government Directory. About 65 percent of all the schools could be mapped to their constituting GPs. This GP wise school location data was combined with the information about the year of establishment of schools to arrive at an aggregate of the new school openings within a GP in a given year. The new school openings serve as a measure of improvement in human capital.

4.3.2 Descriptive Statistics

Figure 4.1 shows the changes over time in our outcomes of interest– the caste category of the elected Sarpanch and the gender of the elected Sarpanch. The proportion of Sarpanches belonging to the ST, SC, OBC and General categories are more or less stable over time. However, for the gender of the elected Sarpanch, we see that the proportion of male sarpanches increased from 47 percent in 2010 to 49 percent in 2015. Because these graphs represent overall changes, they do not distinguish between incidence of Sarpanches of particular castes or gender in reserved vs unreserved seats.

To illustrate changes in Sarpanch identity over time after accounting for reservation status of GP, the fourth column in Table 4.1 shows the coefficient of the year dummy

[20] https://ngdc.noaa.gov/eog/download.html

from a simple pooled linear regression of Sarpanch caste and gender on a time dummy that equals 1 for observations from year 2015, controlling for whether GP was reserved for SC, ST, OBC or for women. Mirroring the statistics from the graph, we find that there is no change in the probability of Sarpanch being ST, SC or OBC from 2010 to 2015 whereas there is a small decrease in the probability of Sarpanch belonging to the general category. Interestingly, here too we find an increase in the probability of the Sarpanch being male by about 2.4 percentage points.

The top panel of Table 4.1 shows summary statistics for the incumbent Sarpanches before legislation. About 61 percent of GPs in our sample had Sarpanches who did not meet the education requirement stipulated by the 2014 legislation. This includes Sarpanches who were not Class 8^{th} pass in non scheduled areas and Sarpanches who were not Class 5^{th} pass in scheduled areas. Excluding the Sarpanch, about 64 percent of the top 5 candidates in a GP did not meet the minimum education threshold introduced by the legislation, on an average. The average combined vote share of the contestants in 2010 who failed to meet the education threshold (including the Sarpanch) from among the top 5 was 62.3 percent.

The second panel of Table 4.1 shows the means for the time varying control variables which in this case are the variables representing whether the GP was reserved for either SC,ST or OBC and separately for women. Since the total Sarpanch seats reserved for SCs or STs is supposed to be proportional to their share of population in the state, it is not surprising to note that the proportion of seats reserved for these categories remains stable across time. Similarly, for OBC seats which are basically determined by randomly allotting GPs unreserved for SCs or STs subject to the combined 50 percent reservation criteria, the table shows that about 15 percent Sarpanch seats are reserved for OBCs. About 48 percent of the GPs in our sample are reserved for women in each of the years.

Table 4.2 shows the summary statistics in the base year 2010 by the status of education of the incumbent Sarpanch before legislation. Columns 1 and 2 give the means and column 3 gives the difference of means. GPs where Sarpanch was less than 8^{th} pass (5^{th} pass in scheduled areas) were more likely to be reserved for women and SCs. Since a plain difference of means in Sarpanch caste/gender is also likely to reflect difference in reservation status, Column 4 shows coefficient from a simple linear regression of the variables on a dummy representing GPs with less than 8^{th} pass (5^{th} pass) Sarpanch after

controlling for reservation status.

20.7% of GPs where Sarpanch did not meet education requirements had Sarpanch belonging to STs and this proportion is similar to their percentage in GPs where the incumbent met the education requirement. 21.6% of uneducated Sarpanches were SCs as opposed to 13 percent SC Sarpanches in GPs with educated Sarpanches. However this difference is not statistically significant after accounting for reservation status. Percentage of OBC Sarpanches is higher in GPs where incumbent was less than 8^{th} pass (5^{th} pass) and that of general category Sarpanches was lower. Sarpanches not having the minimum education qualification were also more likely to be female. While the difference in means suggests a large 44 percentage point difference, Column 4 shows that after controlling for reservation status, uneducated Sarpanches are more likely to be female by about 9 percentage points.

The bottom two panels of Table 4.2 show summary statistics for baseline GP characteristics. GPs with less than 8^{th} pass (5^{th} pass) Sarpanches had a lower total population as well as lower share of ST population compared to GPs where Sarpanches were above 8^{th} pass (5^{th} pass). The overall literacy rate as well as female literacy rate was about 2 percent lower in GPs with uneducated Sarpanches as opposed to other GPs. The total number of primary, middle, secondary and senior secondary schools was lower in GPs where Sarpanch was less than 8^{th} pass (5^{th} pass) with the difference being progressively lesser as the level of schools increased. Given these baseline differences in demographics and other socio-economic indicators, we allow for differential time trends by all these characteristics as well as construct a sample matched on the basis of these characteristics to establish robustness of results.

4.4 Empirical Strategy

The ideal experiment to measure the impact of the minimum education legislation would be to randomly allocate some GPs to undergo elections with minimum education criteria in place with others having no such restriction. However since this legislation was enforced universally in Rajasthan, quantifying the impact of this policy necessitates identifying some source of variation in the impact of the legislation across GPs. A natural candidate for this variation is the educational qualification of the incumbent Sarpanch at the time

this legislation was put in place.

The GPs where the incumbent Sarpanch did not meet the minimum education requirement were presumably the GPs to be hardest hit by the legislation. The idea here is that the education qualification of the incumbent Sarpanch, in some sense, is indicative of the education qualification of the most viable candidate in the GP. This does not rely on the conjecture that the incumbent would or would not recontest the future elections. It simply hypothesizes that a GP with an incumbent who did not meet the education criteria is more likely to witness maximum changes in the characteristics of the elected leader compared with a GP where the incumbent met the minimum education qualifications. It is this particular dimension that we exploit to unravel heterogeneity in the impact of a legislation that was universally enforced.

Using data from the elections prior to and after the legislation, we essentially employ a difference-in-differences approach with those GPs whose Sarpanch did not meet the education criteria introduced by the legislation as the ***treated*** group. In particular we use the following specification:

$$Y_{gpt} = \alpha_g + \gamma Post_t + \beta Post_t * Treat_g + \theta \mathbf{X_{gpt}} + \tau \mathbf{D_g} t + \rho_p t + \epsilon_{gpt} \tag{4.1}$$

This is the usual difference-in-differences specification. Y_{gpt} is the main outcome variable of interest which takes value 1 if the Sarpanch elected in GP g in panchayat samiti p at time t belongs to a particular caste category or gender. So, while looking at whether the elected Sarpanch belongs to OBC category, Y_{gpt} takes value 1 if the elected Sarpanch is OBC and 0 otherwise. $Post_t$ takes value 1 for the post legislation election year 2015. This term captures temporal changes in Sarpanch caste/gender that do not vary across treatment and control groups. $Treat_g$ is equal to 1 for all GPs whose Sarpanch did not meet minimum education criteria put in place by the legislation and 0 otherwise[21]. The main coefficient of interest is β which captures the change in difference between the treated and control group after the legislation is introduced. All standard errors are clustered at the GP level.

α_g refers to GP fixed effects which are included to make sure that the results are not

[21]This is equal to 1 if incumbent Sarpanch was not Class 8 pass in non scheduled areas and if incumbent Sarpanch was not Class 5 pass in scheduled areas

affected by any time invariant GP characteristics. X_{gpt} is a vector of time varying GP specific controls which in our case includes dummies for whether the GP was reserved for ST, reserved for SC, reserved for OBC and reserved for women. Since GP level demographics are only available for the year 2011, we allow for differential trends by these baseline characteristics captured by $D_g t$. These include trends by GP population, proportion of SC and ST population, overall literacy rate in the GP, female literacy rate in the GP as per the 2011 census. We additionally allow for differential trends by Panchayat Samitis (which roughly correspond to census blocks in Rajasthan) given by $\rho_p t$ to make sure that results are not driven by treated GPs lying in specific blocks that have different trends.

The year fixed effects and GP fixed effects make sure that our estimates are not vitiated by any time specific or GP specific time invariant characteristics that might be correlated with the GP having a less educated Sarpanch. While one might argue that GPs with less educated Sarpanches could have different trends, the addition of time trends by baseline GP characteristics including overall and female specific literacy as well as number of schools partially account for these differential trends.

The main threat to the validity of the difference-in-differences estimate is posed by the possibility that the treated and control groups have different pre-existing trends. As standard practice suggests, we use election data from 2005 Panchayat elections in Rajasthan combined with the 2010 election data to check for pre-existing trends in sarpanch caste/gender by our treatment group. We check this by employing the same specification as in equation 1 with 2005 and 2010 election data with treated GPs being those whose elected Sarpanch in 2010 did not meet the legislation's minimum education criteria. We do so without introducing trends by baseline GP characteristics as these are only available for 2011.

4.5 Main Results

We show the results from our empirical strategy in a series of steps. First, we report results from the main specification with GP and election year fixed effects and controls for reservation status without adding time trends by GP demographic characteristics or trends by panchayat samiti (block). Next, we introduce trends by baseline GP demo-

graphics and finally we include block level time trends.

4.5.1 Sarpanch caste

Table 4.3 shows the result for elected Sarpanch belonging to ST and SC caste categories. Columns 1 & 4 show results without time trends by baseline GP demographics or panchayat samiti specific time trends. The coefficient of $Post*Treat$ in Column 1 is negative and significant suggesting that the probability of the elected Sarpanch belonging to STs decreased in treated GPs after legislation compared with untreated GPs. However, there is no change in probability of electing an SC Sarpanch. This is not surprising as the probability of electing an SC Sarpanch even before legislation was low in Rajasthan allowing for less room for change.[22]

Columns 2 & 5 show the same results after allowing for trends by baseline GP demographic characteristics. Columns 3 & 6 further add time trends by panchayat samiti. We find the coefficients of $Post * Treat$ do not change much for either election of ST or SC Sarpanch upon introducing these trends.

Table 4.4 shows the same results for the probability of election of Sarpanch belonging to OBC and general categories. Columns 1 & 4 show results without any time trends by GP characteristics or blocks. We find that the probability of the Sarpanch belonging to OBCs fell by 5.6 percentage points in treated GPs and correspondingly that of the Sarpanch belonging to the general category increased by 7 percentage points. Together, these results suggest a substitution of leaders belonging to disadvantaged castes (STs and OBCs) with leaders belonging to upper caste (general category) in the aftermath of the legislation.

Columns 2 & 5 show the same results after allowing for trends by baseline GP demographic characteristics and we find similar results with the probability of elected Sarpanch belonging to the general category increasing and that of belonging to OBCs declining in treated GP in 2015. The coefficients are similar in magnitude to those in Columns 1 & 4 suggesting that the results are not driven by treated GPs having different trends owing to demographic differences. Columns 3 & 6 show results with panchayat samiti (block) specific time trends and the coefficient of $Post*Treat$ stays similar in magnitude

[22]The probability of electing an SC Sarpanch in unreserved seats was only 2 percent in 2010 compared with a 6 percent probability of electing an ST Sarpanch. This is against a 52 percent probability of electing an OBC sarpanch in unreserved seats in 2010.

and direction. The stability of the coefficients across specifications indicates that the results are not driven by differential time trends due to reasons other than the impact of legislation.

4.5.2 Sarpanch gender

Table 4.5 reports coefficients from the main specification with the elected Sarpanch being male as the outcome. As before, the first column shows results excluding differential time trends by baseline GP characteristics and panchayat samiti (block) time trends. The coefficient of $Post * Treat$ is positive and significant indicating that the probability of a male Sarpanch being elected went up in treated GPs after legislation relative to untreated GPs by about 8.2 percentage points. This is a non trivial increase and suggests that the substitution of leaders belonging to "backward" castes by "upper" castes was accompanied by a substitution of females by males.

Addition of trends by baseline GP demographics in Column 2 and trends by panchayat samiti (block) in Column 3 does not change the direction of the coefficient and only reduces its magnitude by 1.2 percentage points lending credence to the robustness of the results.

An interesting experiment to examine the gender substitution is offered by looking at GPs in scheduled areas where all GPs remain reserved for STs in both election years. Even looking at the overall proportions, we find that the percentage of female Sarpanches in GPs unreserved for women falls by 2 percentage points from 14 percent in 2010 to 12 percent in 2015. Implementing the main specification with this limited sample of GPs in scheduled areas, we find that the likelihood of election of a female Sarpanch falls by a large 17 percentage points in treated areas relative to untreated after legislation (Appendix Table 4.A1).

4.5.3 Pre-existing trends

To check for the possibility of the results being driven by pre-existing differences in trend between the treated and untreated GPs, we implement the same difference-in-differences specification but with the 2005 and 2010 election data. Treated GPs are defined as fixed by our earlier exercise. $Post$ here takes value 1 for election year 2010 and 0 for 2005. The coefficient of $Post*Treat$ captures the change in difference between treated and untreated

GPs between 2005 and 2010. We focus on the simplest specification with GP fixed effects and without any time trends by baseline GP characteristics (as they correspond to 2011) or panchayat samiti (block) specific time trends.

The results for this specification are shown in Table 4.6. The coefficient of $Post * Treat$ is insignificant for elected Sarpanch belonging to ST and Sarpanch belonging to SC suggesting no violation of the parallel pre-existing trends assumption. The coefficient of $Post * Treat$ for elected Sarpanch being an OBC is however of the opposite sign that we find in our main results. This suggests that if the same trend had continued from 2010 to 2015, in the absence of legislation we would have found results opposite to what we actually see, suggesting that the presence of this pre-existing trend cannot explain our results. Similarly, results for the Sarpanch belonging to the general category and Sarpanch being a male, run opposite to the main results. Therefore, we can state that these pre-existing trends run in the opposite direction to what we find and cannot be the driving factors for our main results.

4.5.4 Intersection of caste and gender

While the previous results establish the changes in the caste and gender of the elected Sarpanch due to legislation separately, they do not illustrate how these two characteristics change together in the aftermath of legislation. For instance, we show that the probability of election of an ST Sarpanch falls in treated GPs and that of a male Sarpanch goes up but that does not provide any insight as to whether there is a change in likelihood of election of female ST Sarpanches or male ST Sarpanches or both. To look into this, we implement the original empirical specification with outcomes now being defined based on both caste and gender of the elected Sarpanch.

We look at the following outcomes- whether the elected Sarpanch is ST and female, whether elected Sarpanch is ST and male, whether the elected Sarpanch is SC and female, whether elected Sarpanch is SC and male, whether elected Sarpanch is OBC and female, whether elected Sarpanch is OBC and male, whether elected Sarpanch is of general category and female and whether elected Sarpanch is of general category and male.

Table 4.7 shows the results for these outcomes. Column 1 shows the coefficient of interest for probability of election of a male ST Sarpanch and Column 2 for election of a female ST Sarpanch. It is interesting to note that the probability of election of a female

ST Sarpanch goes down whereas that of the male ST Sarpanch is unchanged in treated GPs post legislation. Even for OBCs, the likelihood of election of a male OBC Sarpanch is unchanged but that of a female OBC Sarpanch decreases. The gender wise results for Sarpanch belonging to the general category corroborate the same story, the probability of election of a male Sarpanch goes up but that of female Sarpanch belonging to the general category remains the same.

With the exception of results for the category of SC Sarpanches, all other results point to a supplanting of female Sarpanches belonging to disadvantaged groups by male Sarpanches from the "forward" castes. This is consistent with hypotheses that the minimum education legislation would hit women belonging to disadvantaged groups the hardest on account of the prevalence of low education levels among them.

4.6 Robustness Checks

4.6.1 Additional time trends

To mitigate concerns that the education status of the Sarpanch of treated GPs is indicative of the education status of the treated GPs and therefore such GPs would follow different trends in election of Sarpanches of particular caste and gender, we already include trends by the GP literacy rate as well as the female literacy rate. To additionally strengthen the confidence that the main results are not driven by differential trends in education, we allow for trends by the total number of primary, middle, secondary and senior secondary schools in the GP as per the census 2011 as well as the average education of the GP population which is obtained from the Socio-Economic Caste Census (SECC) 2011 in the original specification. Secondly, we add trends by the proportion of deprived households in the GP obtained from SECC 2011 as an indicator of the socio-economic status of the GP population.

The results are shown in Table 4.8. Since the SECC data as well as data for number of schools is not available for all villages and therefore, all GPs, the number of observations is smaller compared with the original sample. We find that all the original results are still valid with the coefficients of $Post * Treat$ remaining stable in magnitude and direction.

4.6.2 Impact of 2010 reservation

The definition of treated GPs depends on whether the incumbent Sarpanch elected in 2010 met the minimum education criterion imposed by the legislation. Since the education of the elected Sarpanch in 2010 might be correlated with the reservation status of the GP in 2010, there is a concern that the results are driven by GPs with different reservation status in 2010 having differentially educated Sarpanches.

While controlling for the reservation status as we do in all our regressions partially accounts for this source of concern, we additionally introduce trends by the reservation status of the GP in 2010 to our main specification and check if the results remain the same. Table 4.9 shows the result from these regressions. The coefficients of $Post * Treat$ remain similar to our original results for all the dependent variables suggesting that reservation status in 2010 is not driving our results.

4.6.3 Uneducated sarpanch or candidate pool?

One can argue that a better candidate for identifying GPs which were the most affected by the legislation would be the pool of contestants in the GP from the 2010 election who did not meet the minimum education requirement. This is based on the reasoning that the education of the candidate pool in 2010 is a better indicator for the education of the potential future Sarpanch as it spells out the potential substitute candidate pool in case the incumbent does not meet the education requirement. The idea is that even if the Sarpanch was uneducated but there existed an educated substitute candidate pool presumably similar to the Sarpanch in other characteristics, the legislation would have less of a bite.

To check whether it is the Sarpanch not meeting the education requirement that matters or whether the extent of the candidate pool excluding the sarpanch not meeting the education requirement that matters for the result, we add separately the proportion of candidates in 2010 (from the top 5 that data is available for) other than the Sarpanch that were less than 8^{th} pass (5^{th} pass). The results are shown in Table 4.10. The coefficient of the trend by proportion of the candidate pool excluding Sarpanch that is less than 8^{th} pass (5^{th} pass) is insignificant in all the regressions.

This is hardly surprising as the average vote share of the winning Sarpanch is about 42

percent with similar averages for both the treated and untreated GPs in 2010, suggesting that the Sarpanch is in fact the most important candidate. We show in the next section that the original results do not change if we measure treatment by the proportion of the uneducated candidate pool including the Sarpanch or by the vote share of the candidate pool that was uneducated including Sarpanch. The fact that the Sarpanch is the most important candidate supports the rationale for using whether the incumbent Sarpanch met the minimum education as the appropriate measure of the impact of legislation.

4.6.4 Alternative independent variables

We move away from a dichotomous definition of treatment and look at two other definitions that use the entire top 5 candidate pool in 2010 that did not meet the education criteria. The first measure is the proportion of the top 5 candidates in 2010 who were not Class 8^{th} pass or alternatively Class 5^{th} pass in scheduled areas. The second measure uses the combined vote share of the less than 8^{th} pass (5^{th} pass) candidates among the top 5 candidates as a measure of effect of the legislation. Both these measures essentially define a varying *intensity* of treatment among GPs due to the minimum education legislation.

The results from the difference-in-differences specification for these intensity of treatment measures are shown in Tables 4.11 & 4.12. We find similar results for these alternative definitions of treatment in case of all the outcomes with a decline in probability of election of ST and OBC Sarpanches as well as female Sarpanches with a corresponding rise in likelihood of election of general category and male Sarpanches.

4.6.5 Matching on covariates

One concern that becomes evident from the baseline summary statistics in treated vs untreated GPs is that they differ in characteristics in terms of population, literacy rates as well as number of schools and are therefore likely to have different trends. While allowing for trends by all these characteristics alleviates some of these concerns, we additionally check whether we find the same results when we look at GPs with incumbent Sarpanch not meeting the minimum education that are similar to GPs with Sarpanch meeting the minimum education in terms of all these baseline characteristics.

To this end, we use propensity score matching to construct a sample from treated and untreated GPs matched on the basis of the 2011 GP population, share of ST and

SC population, literacy rate, female literacy rate, number of primary, middle, secondary and senior secondary schools from the 2011 census and the average education of the GP from SECC 2011. We allow for one-to-one matching without replacement with a specified caliper of 0.001. Our matching exercise results in a sample of 1895 treated GPs matched to 1895 unique untreated GPs. Table 4.13 shows the difference in demographic characteristics in the matched sample. The differences are statistically insignificant suggesting that treated GPs are ideally comparable to untreated GPs in this matched sample.

We then implement the initial difference-in-differences specification using only the matched sample. The results are shown in Table 4.14. As before we find a decline in the probability of an ST Sarpanch with the magnitude of decline being greater than our original results. There is no change in probability of election of SC Sarpanch, but a decline in probability of election of OBC Sarpanch. This is accompanied by increasing probability of election of general category and male Sarpanches, right in line with our original results. Even the magnitudes of the coefficients of $Post * Treat$ for Sarpanches belonging to OBCs, general category and being male are similar to the main results.

4.6.6 Placebo tests

We carry out placebo tests where we randomly reassign the education status (whether 8th pass) of the incumbent Sarpanch in 2010 to GPs. For each of these random assignments, we carry out the difference-in-differences estimation and arrive at a placebo coefficient of $Post * Treat$ for each of the outcomes. This exercise is replicated a 1000 times for each outcome, allowing us to create a distribution of these placebo coefficients.

The results from this exercise are shown in Figures 4.2, 4.3, 4.4, 4.5 and 4.6. We find that the p-values from these distributions of placebo coefficients are extremely low (almost zero) for probability of election of ST Sarpanch, OBC Sarpanch, General Sarpanch and male Sarpanch. This shows that the probability of obtaining coefficients of the magnitude that we find in our main results, through chance alone are very low, lending credibility to our estimates. Secondly, the p-value for the elected Sarpanch being SC is 0.745 which is in line with our main results which shows no change in the probability of the elected Sarpanch being an SC.

4.7 Heterogeneity

The hypothesis that the probability of election of Sarpanches belonging to STs and OBCs as well as female Sarpanches would decline where the incumbent did not meet the education criteria imposed by the legislation hinges on the unavailability of candidates from the these caste and gender groups who could enter the electoral fray. Following from this line of thought, one can expect that the effects of the legislation would be diminished in GPs where these disadvantaged groups are more educated, signifying a greater pool of viable candidates. While we would ideally want to see if these effects vary depending on the education level of caste groups in a GP, the unavailability of caste wise education or literacy data at the village or GP level prevents us from doing so. We check this using overall literacy rate and the female literacy rate of the GP instead.

The GPs in our sample were divided into five quintiles based on the overall literacy rate and separately based on the female literacy rate from the 2011 census, with the higher quintiles corresponding to higher levels of literacy. We then added interactions of these quintile groups with $Post * Treat$ to see if the impact of the legislation changed as we moved to higher literacy GPs. The first quintile corresponding to the GPs with lowest literacy is omitted from all regressions and taken as the base category.

Figure 4.7 plots the coefficients of $Post * Treat * Literacyquintile$ with 90 percent confidence intervals. The coefficients are not significantly different as we move up quintiles for election of ST, SC, OBC and general category Sarpanch indicating no difference in impact of legislation for higher educated GPs. However, we find that the impact for the probability of election of male Sarpanch is lesser in the highest educated quintile compared with the least, lending some support to our hypothesis. Table 4.15 shows the detailed results from these regressions.

Figure 4.8 plots coefficients for interactions with dummies for female literacy rate quintiles. Again we find no changing impact of legislation for all caste groups, however for election of male Sarpanches we find a countervailing positive coefficient for the highest educated GPs, showing that the likelihood of substitution of females by males is declining in the education of females. Table 4.16 shows detailed results for these regressions.

4.8 Impact on outcomes

The stated objective of introduction of the minimum education legislation was to improve the efficiency and financial management in working of Panchayats and to curb corruption in allocation of budgets for public programs and public goods (IndianExpress, 2014). The Sarpanch being the leader of the Village Council, is instrumental for the functioning and final decision making of the Gram Panchayat- a fact that has been shown in literature (Besley et al., 2012).

The rationale for restricting candidature based on education qualification hinges on the hypothesis that educated Sarpanches perform better than uneducated Sarpanches. There is some evidence that points that educated leaders lead to better outcomes in terms of economic growth, education of citizens and is therefore a credible proxy for the *quality* of a leader (Diaz-Serrano and Pérez, 2013; Besley et al., 2011; Lahoti and Sahoo, 2020b).

In the specific context of Gram Panchayats in Rajasthan, McManus (2014) used close elections between candidates who have completed high school and those who did not and showed that GPs with educated Sarpanches had a lower likelihood of delayed projects under NREGS than Sarpanches without high school education. However, he found that the difference in total expenditures under NREGS was not statistically significant between GPs with educated vs uneducated Sarpanches. He also found that GPs that elect educated Sarpanches showed greater increases in female literacy rates and that educated Sarpanches were less likely to locate NREGS projects near their own residence, an indication of lower corruption.

While recognizing that educated leaders can potentially lead to less corruption and improved efficiency, it must be emphasized that a legislation mandating minimum education also affects other characteristics of Sarpanches apart from education. We show that Sarpanches are less likely to belong to disadvantaged castes and are less likely to be women in GPs most affected by the legislation. Both these facets have been shown to matter for distribution of public goods to minorities and women (Chattopadhyay and Duflo, 2004; Bardhan et al., 2010; Besley et al., 2004). It is also worth noting that erecting barriers to candidacy is also likely to affect the extent of political competition (Afzal, 2014; Ansolabehere and Gerber, 1996; Drometer and Rincke, 2009; Stratmann, 2005) which can again impact Sarpanch performance (Afridi et al., 2021; Gupta and Mukhopadhyay, 2016). Additionally, it may well be that where GPs have uneducated

incumbents, there is an influx of more educated candidates who are also younger and perform worse on account of less experience (Afridi et al., 2017).

So while a naive point of view might suggest that imposing education requirements should automatically lead to better Sarpanch performance, this stance ignores the impact of such restrictions on other aspects of the electoral process and on other characteristics of the candidate, both of which have been shown to matter for actual outcomes. On one hand higher education of Sarpanch might improve public goods/programs delivery, on the other hand concomitant changes in Sarpanch caste and gender might have the opposite effect. Similarly, if such candidature restrictions lead to a decrease in electoral competition as has been shown in case of other ballot access restrictions, it might lead to worse performance based on the evidence that shows that higher political competition induces better leader performance (Besley et al., 2010).

In this section, we try to look at the impact of this legislation on three different types of outcomes. Firstly, we look at the impact of this legislation on performance of the NREGS by focusing on the important indicator of per capita person days of work generated in the program. Second, we examine whether the legislation led to better economic growth as captured by night time luminosity and finally, we look at whether having educated leaders due to legislation led to improvement in human capital outcomes in the form of new school openings.

We examine the impact of the legislation on outcomes using two different specifications. First, we look at the impact of legislation on the outcome of interest for each year using the following:

$$Y_{gpt} = \alpha_g + \gamma_t + \sum_{t=2013}^{2019} \beta_t Year_t * Treat_g + \theta \mathbf{X_{gpt}} + \tau \mathbf{D_g} t + \rho_p t + \epsilon_{gpt} \tag{4.2}$$

Y_{gpt} is the outcome of interest in GP g in Panchayat Samiti p in year t. The coefficients of interest are the coefficients of the interactions of the dummy for treated GPs $Treat_g$ and the year dummies $Year_t$. α_g are GP fixed effects and γ_t are year fixed effects. X_{gpt} is a vector of time varying GP specific controls which in our case includes dummies for whether the GP was reserved for ST, reserved for SC, reserved for OBC and reserved for women. Since GP level demographics are only available for the year 2011, we allow

for differential trends by these baseline characteristics captured by $D_g t$. These include trends by GP population, proportion of SC and ST population, overall literacy rate in the GP, female literacy rate in the GP as per the 2011 census. We additionally allow for differential trends by Panchayat Samitis given by $\rho_p t$.

We omit year dummy for 2012 and it's interactions as the base category. β_t is the change in the outcome in treated GPs relative to untreated GPs between 2012[23] and year t. The beta coefficients between years 2012 to 2014 help in understanding whether GPs with disqualified Sarpanches had differential trends in outcomes prior to the introduction of legislation. This is important to establish the validity of estimated beta coefficients after legislation came into effect. The beta coefficients for years 2015 to 2019 capture the changes in treated areas due to the impact of the legislation.

The second specification we implement to examine the impact of minimum education legislation is the earlier difference-in-differences specification:

$$Y_{gpt} = \alpha_g + \gamma Post_t + \beta Post_t * Treat_g + \theta \mathbf{X_{gpt}} + \tau \mathbf{D_g} t + \rho_p t + \epsilon_{gpt} \tag{4.3}$$

Y_{gpt} is the outcome of interest in GP g in Panchayat Samiti p in year t. Here, everything else is defined as before except that we have multiple years of observations than just two. $Post_t$ takes value 0 for years 2012 to 2014[24] and value 1 for years 2015 to 2019. β is the coefficient of interest which captures changes in outcome in GPs where existing Sarpanches did not meet education requirements set by legislation relative to GPs where the incumbent Sarpanch met the requirements of the legislation between the years before and after legislation.

4.8.1 NREGS

We use the total person days of work generated under the NREGS as the relevant indicator of the performance of the scheme. This data is available publicly for each financial year beginning from 2012-2013 for each GP. We use data from 2012-2013 until 2018-2019 for

[23]This is because data is available for NREGS and VIIRS night lights from 2012 onwards. For school openings where we have data from previous years, we consider the data series from year 2010 onwards.

[24]This is because data is available for NREGS and VIIRS night lights from 2012 onwards. For school openings where we have data from previous years, $Post_t$ takes value 0 for years 2010 to 2014

our analysis. We were able to match about 5403 GPs of the 6024 GPs in our original sample to the NREGS data set. Due to possibility of cyclicality in NREGS outcomes corresponding with electoral cycles, we limit our sample to financial years 2012-13, 2013-14, 2017-18 and 2018-19.[25]

We define our outcome of interest as the per capita total days of work generated by dividing the total person days of work reported by the total population of the GP obtained from census 2011. The results from the specification following equation (2) are shown in Table 4.17. Firstly, column 2 the coefficients of interactions $Year * Treat$ are not statistically different from the base year in the years before introduction of legislation, indicating that one can not reject the assumption of pre-treatment parallel trends between treated and untreated GPs.

Secondly, the coefficients of the interactions terms $Year * Treat$ remain statistically insignificant even for years after the introduction of the legislation indicating no change due to legislation. We obtain similarly insignificant results for the difference-in-differences specification from equation (3) shown in column 1 of Table 4.17. The coefficient of $Post * Treat$ is insignificant suggesting that the introduction of the legislation had no impact on per capita work generated in GPs where the incumbent Sarpanch did not meet the minimum education criteria and hence an increase in education of elected Sarpanch after 2015 election.

4.8.2 VIIRS Night time luminosity

The next outcome we consider is economic activity as proxied by night time luminosity. We use the VIIRS nighttime lights data from the NOAA (National Oceanic and Atmospheric Administration) from years 2012 to 2019. The VIIRS data reports night time lights with a constant spatial resolution of 742 m and does a good job of reporting night lights of low intensity (as in rural areas) over smaller regions (like villages) (Gibson et al., 2021).

Using GIS sofware, we aggregate the night lights by summing all reported grid values in a GP in a year. We deflate this sum of night lights by the GP population to arrive

[25]With GP elections taking place every 5 years, financial year 2017-18 is the post-legislation electoral cycle counterpart of the pre-election year 2012-13 and similarly 2018-19 is the comparable post legislation year to the year 2013-14. We omit years 2014-15, 2015-16 and 2016-17 as data for their pre(post) legislation counterpart is not available or it is an election year. However, our results stay the same even if we include all years from 2012 to 2018

at our outcome of interest- per capita sum of night lights. Again to mitigate concerns about presence of electoral cycles in economic activity, we limit our sample to years 2012, 2013, 2014 before legislation and 2017, 2018, 2019 after legislation[26]. To examine the the impact of legislation on economic activity, we use this per capita sum of night lights as dependent variable in specifications given by equation (2) and equation (3).

Column 2 of Table 4.18 shows result from the first specification that reports year by year change in night lights for treated GPs relative to untreated GPs. The coefficients of the interactions of $Year$ and $Treat$ are insignificant in the years preceding the legislation indicating no violation of parallel trends. The coefficients of the interactions remain insignificant in the years after the legislation suggesting that there was no impact of the legislation on economic activity in GPs that saw replacement of uneducated Sarpanches with educated ones. Column 1 of Table 4.18 shows the result from specification from equation (3). Again the coefficient of $Post * Treat$ is insignificant showing no change in treated GPs after legislation.

4.8.3 New school openings

This sections derives from the literature highlighting a positive association between educated leaders and educational outcomes of citizens. Diaz-Serrano and Pérez (2013) use changes in education of leader from one mandate to other in a cross country dataset from years 1970-2004 and find a positive relationship between education of leader and educational attainment of the population. Lahoti and Sahoo (2020b) look at the impact of having leaders with a graduate degree on learning levels, enrollment, school funding and infrastructure in context of state legislative assembly elections in India. They find that educated leaders positively impact enrollment, education funding and infrastructure only in the most developed states.

While it may be true that educated leaders focus more on improving education outcomes, another reason to expect that a legislation mandating minimum education for leaders will have an effect on education outcomes is the possibility of a *role model effect*. This idea posits that if elected leaders are also looked upon as role models by citizens, they might be encouraged to pursue, invest in and demand higher education in response to leaders being more educated. For instance, Anukriti and Chakravarty (2019) look at

[26]Results are unchanged even with including years 2015 and 2016 in the analysis

the impact of restrictions on candidature based on number of children in the context of Gram Panchayat elections in India and find that such restrictions led to reduction in fertility in rural couples but worsened the sex ratio in favor males in places with strong son preference.

While it would be interesting to look at changes in literacy rates or enrollment in schools, the lack of availability of data for both the pre and post legislation time periods prevents us from doing so. Instead we look at the impact of the legislation on investment in human capital measured by the opening of new schools. We use the DISE (District Information System for Education) raw data from the year 2017 for the state of Rajasthan which is a census of all schools containing details about the location, year of establishment, type (primary, middle, secondary), management of each school. By matching the school location data to the census village, we were able to map school locations to GPs. This location data was combined with information about the year of establishment of schools to create a GP level panel of new school openings from years 2010 to 2017.

We use the total number of new schools opened and separately the number of new primary, middle, secondary and senior secondary schools opened in a GP in a year as the main outcome of interest. We implement a year by year difference-in-differences specification given by equation (2) and the general difference-in-differences specification given by equation (3). All regressions additionally control for the number of existing schools at the beginning of the year. The results are shown in Table 4.19.

The coefficients of all year and treatment interactions are insignificant for the total number of schools in the pre-treatment years. The same is true for most of the coefficients of interactions for the other types of schools. Table 4.20 shows results from specification (3). We find no evidence of change in school openings after 2015 legislation using either of the specifications, suggesting no improved investment in education by the educated leader or increased demand through role model effect.

4.9 Conclusion

We show that the impact of legislation stipulating minimum education for candidates in elections can have unintended consequences for the identity of the elected leaders. More specifically, we show that a legislation disbarring candidates based on education in a

context where there is wide inequality in education attainment based on caste and gender is likely to lead to lower representation of candidates from historically disadvantaged groups as well as decline in representation by women. Even among the marginalized castes, the decline in representation is more likely to driven by women from marginalized castes in leadership positions.

Secondly, we find that the legislation did not lead to any noticeable improvements in outcomes such as work generated under NREGS, economic development as measured by night lights or human capital outcomes as measured by new school openings. This evidence shows that while education in itself might be a desirable quality in a leader, an educated leader instated through a system of candidature restrictions does not necessarily lead to better leader performance. This argues for a more nuanced approach to improve leader quality.

Given the enormous efforts that have been made in the past decades to increase the representation of these groups through affirmative action policies, imposing an education criteria qualifies seems like a step in the wrong direction. Not only did the legislation lead to decline in representation of disadvantaged groups, we show that there was no corresponding benefit achieved in terms of economic or human capital outcomes. Moreover, given the evidence that having women leaders encourages women in subsequent elections to participate, the effects of such legislation might perpetuate beyond the period of their enforcement.

Examining the impact of such restrictions on other candidate characteristics such as experience and measures of electoral competition such as winning margin could arguably yield a more comprehensive picture of changes in electoral landscape due to introduction of legislation. This might have proved useful in determining the reasons improvement in education of leaders fails to induce an improvement in their performance, however lack of data on other candidate characteristics and measures of electoral competition prevents us from exploring this channel in this paper.

Our evidence paints a note of caution for policy makers. While education in itself might be a desirable characteristic to have in political leaders, qualifications must be made about the kind of measures that should be taken to achieve this. Especially in a country like India where sections of population are denied equal access to education, steps to induce educated leaders by imposing minimum education requirements are indeed

undemocratic and contradict principles of fairness and equality.

Figures and Tables for Chapter 4

Figure 4.1: Proportions of Sarpanch by Caste and Gender over time

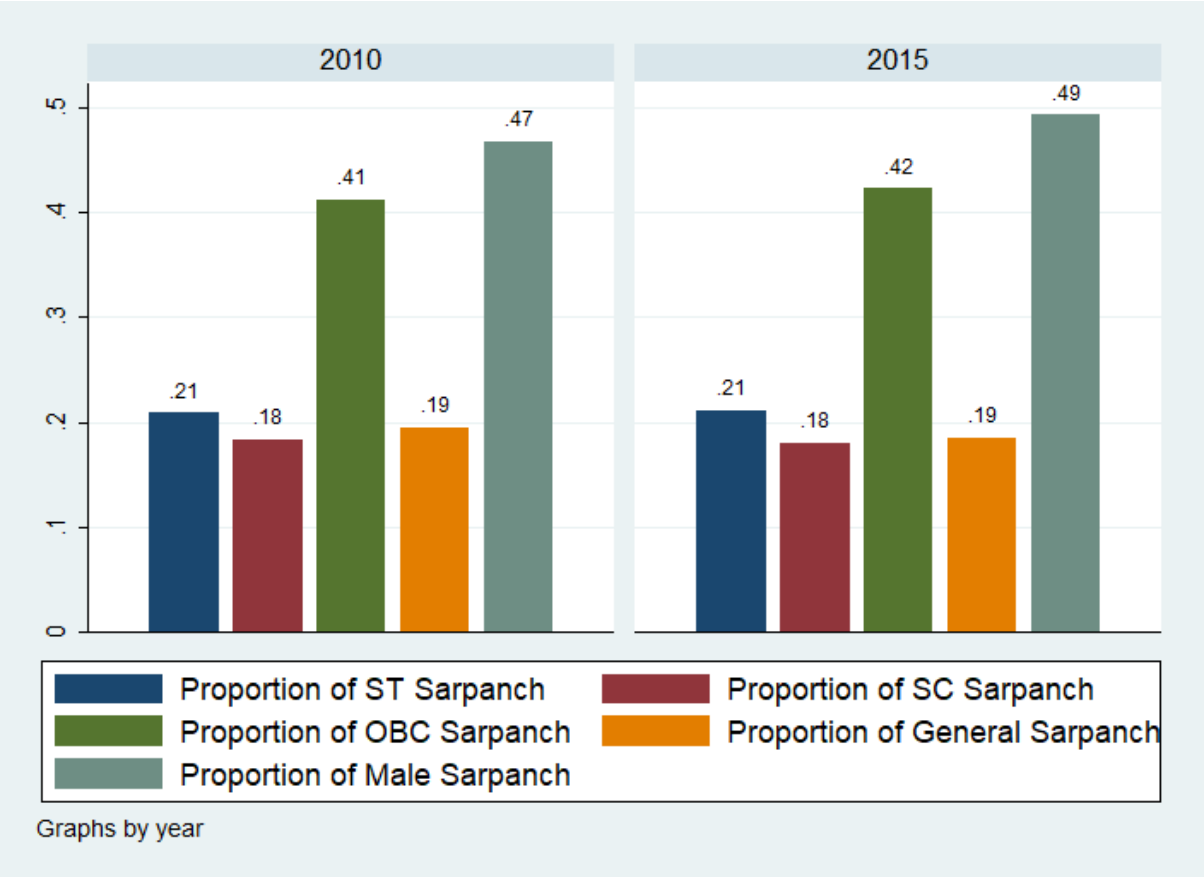
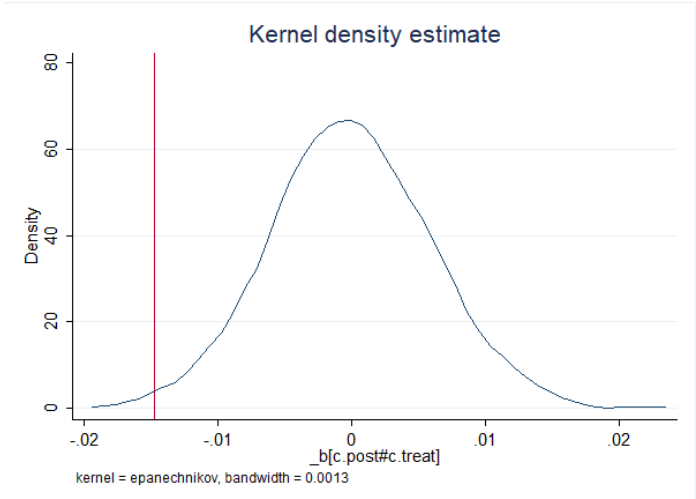
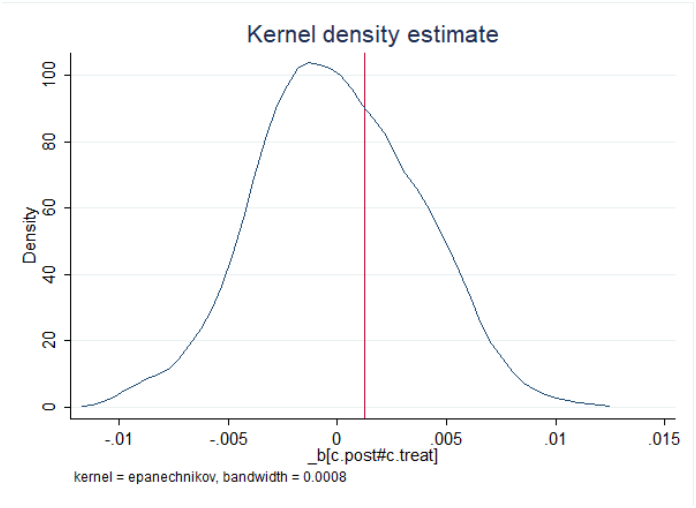


Figure 4.2: Placebo test-Sarpanch is ST



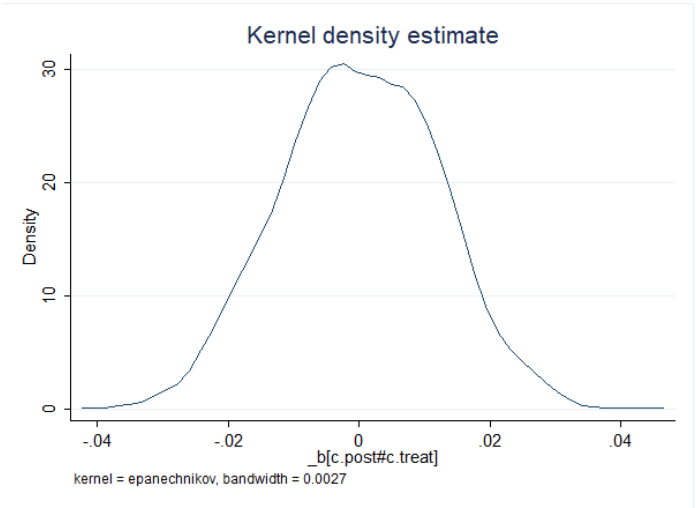
Distribution of simulated coefficients of $Post*Treat$ in specification with elected Sarpanch being ST as outcome. The red line indicates the value of the coefficient obtained from the main specification. The calculated p-value which is probability of obtaining a coefficient as large we do is 0.013.

Figure 4.3: Placebo test-Sarpanch is SC



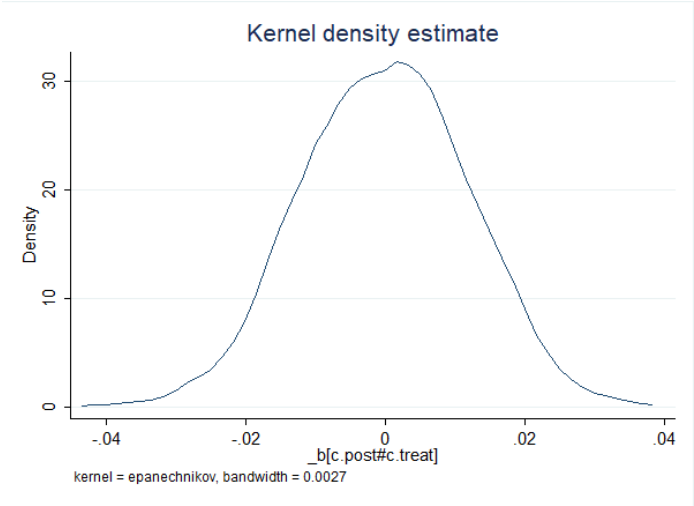
Distribution of simulated coefficients of $Post{*}Treat$ in specification with elected Sarpanch being SC as outcome. The red line indicates the value of the coefficient obtained from the main specification. The calculated p-value which is probability of obtaining a coefficient as large we do is 0.745.

Figure 4.4: Placebo test-Sarpanch is OBC



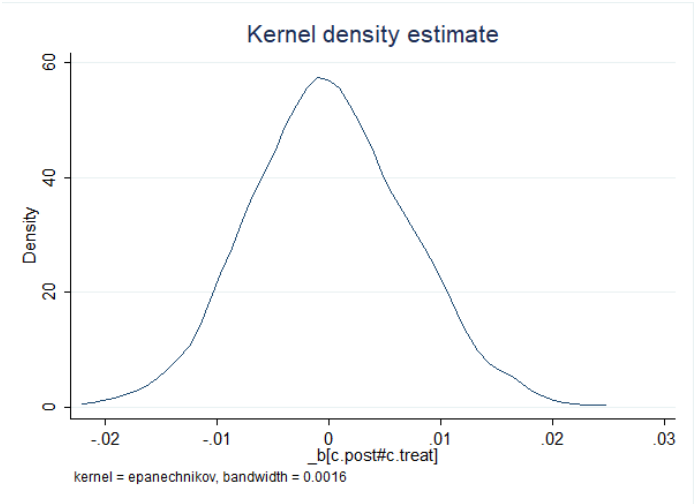
Distribution of simulated coefficients of $Post{*}Treat$ in specification with elected Sarpanch being OBC as outcome. The calculated p-value which is probability of obtaining a coefficient as large we do is 0.00. The red line indicating the value of the coefficient obtained from the main specification moves to the extreme left tail of the distribution.

Figure 4.5: Placebo test-Sarpanch is from General category



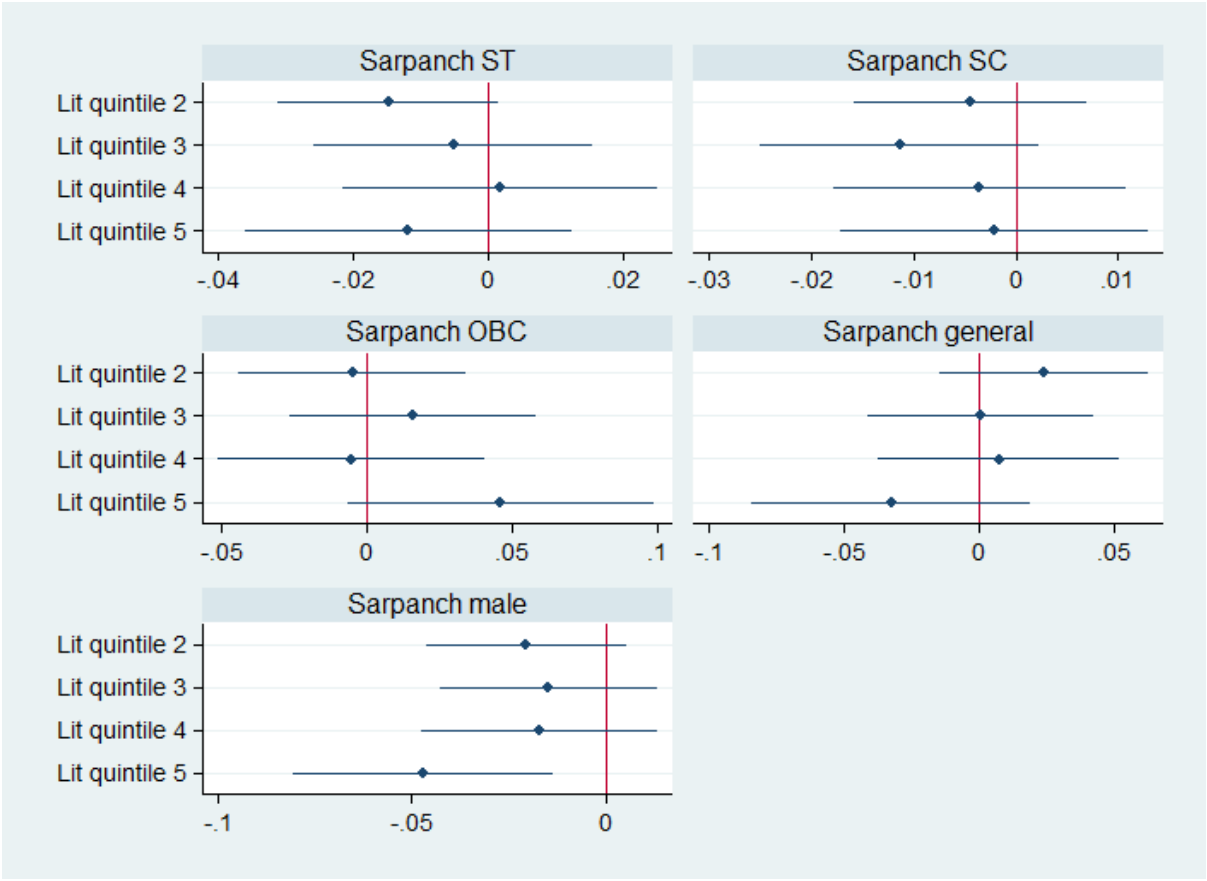
Distribution of simulated coefficients of $Post{*}Treat$ in specification with elected Sarpanch being of general category as outcome. The calculated p-value which is probability of obtaining a coefficient as large we do is 0.00. The red line indicating the value of the coefficient obtained from the main specification moves to the extreme right tail of the distribution.

Figure 4.6: Placebo test-Sarpanch is male



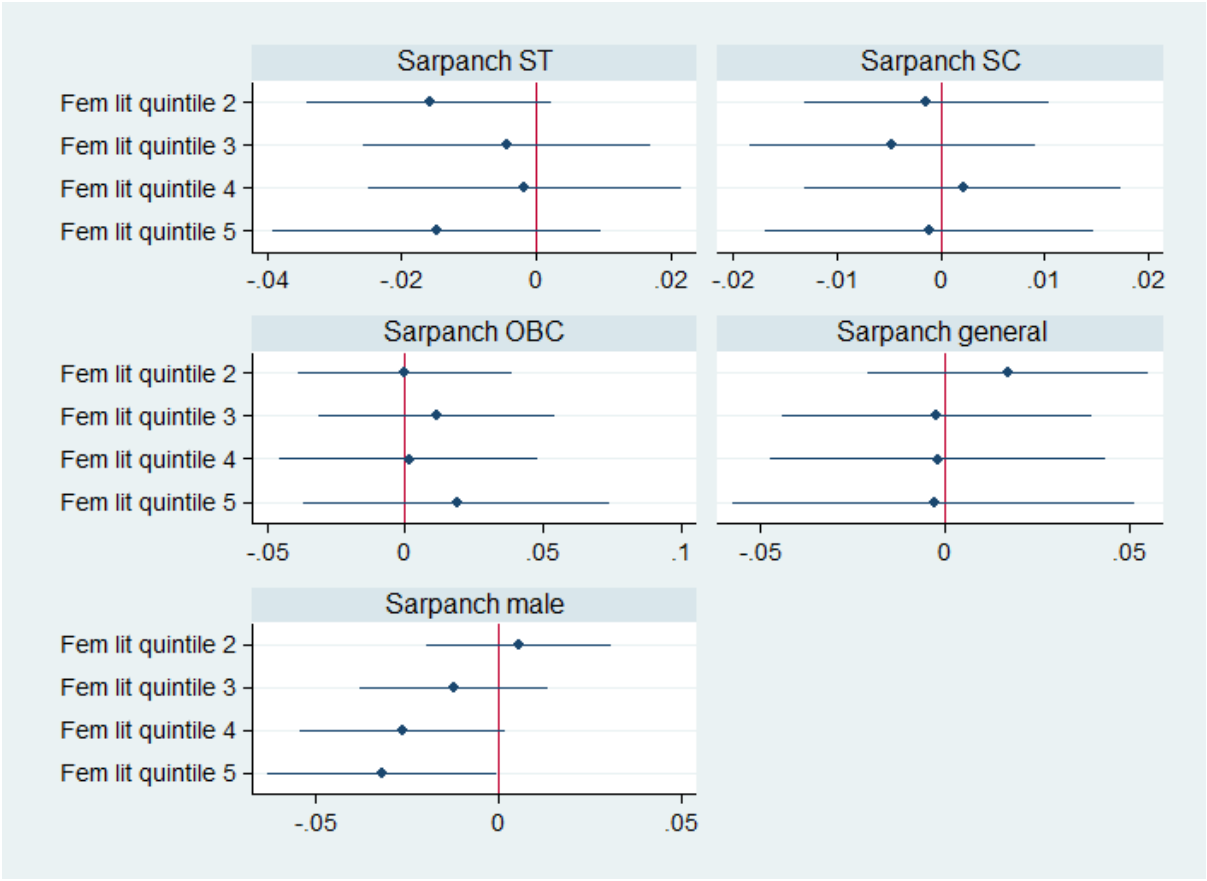
Distribution of simulated coefficients of $Post{*}Treat$ in specification with elected Sarpanch being male as outcome. The calculated p-value which is probability of obtaining a coefficient as large we do is 0.00. The red line indicating the value of the coefficient obtained from the main specification moves to the extreme right tail of the distribution.

Figure 4.7: Interactions of Post*Treat*Literacy rate quintile



Coefficients of interactions of $Post * Treat * Literacyquintile$ with 90 percent confidence intervals. Lit quintile 1 which corresponds to the population with lowest literacy rate forms the reference category.

Figure 4.8: Interactions of Post*Treat*Female literacy rate quintile



Coefficients of interactions of $Post * Treat * Femaleliteracyquintile$ with 90 percent confidence intervals. Fem quintile 1 which corresponds to the population with the lowest female literacy rate forms the reference category.

Table 4.1: Summary statistics by year

	2010	2015	Difference	Difference regression	Obs.
Independent variables					
Sarpanch less than Class 8th(Class 5th) pass	0.612				6024
	(0.487)				
Proportion of candidates less than Class 8th (Class 5th) pass	0.633				6011
	(0.344)				
Proportion of candidates less than Class 8th (Class 5th) pass excluding Sarpanch	0.639				6011
	(0.368)				
Vote share of less than Class 8th(Class 5th) pass candidates	0.623				6011
	(0.366)				
Control variables					
Reserved for ST	0.181	0.178	-0.002		12048
	(0.385)	(0.383)	(0.007)		
Reserved for SC	0.174	0.170	-0.003		12048
	(0.379)	(0.376)	(0.007)		
Reserved for OBC	0.152	0.150	-0.002		12048
	(0.359)	(0.357)	(0.007)		
Reserved for Women	0.482	0.480	-0.002		12048
	(0.500)	(0.500)	(0.009)		
Dependent variables					
Sarpanch ST	0.209	0.212	0.003	0.004	12048
	(0.407)	(0.409)	(0.007)	(0.003)	
Sarpanch SC	0.184	0.180	-0.004	-0.000	12048
	(0.387)	(0.384)	(0.007)	(0.002)	
Sarpanch OBC	0.413	0.423	0.010	0.008	12048
	(0.492)	(0.494)	(0.009)	(0.006)	
Sarpanch General	0.195	0.186	-0.009	-0.012**	12048
	(0.396)	(0.389)	(0.007)	(0.006)	
Sarpanch Male	0.467	0.493	0.026**	0.024***	12048
	(0.499)	(0.500)	(0.009)	(0.003)	
Observations	6024	6024	12048	12048	

This table gives the year wise means and difference in means for all the independent variables, control variables and dependent variables. Column 4 shows the coefficient of a dummy for year 2015 from a regression of sarpanch caste/gender on the year dummy after controlling for reservation status. Standard errors in parentheses.

Table 4.2: Baseline summary statistics by treatment status

	Sarpanch less than Class 8th (5th) pass	Sarpanch was Class 8th (5th) pass	Difference	Difference regression	Obs.
Control variables					
Reserved for ST	0.178	0.186	-0.008		6024
	(0.382)	(0.389)	(0.010)		
Reserved for SC	0.206	0.123	0.083***		6024
	(0.404)	(0.328)	(0.010)		
Reserved for OBC	0.151	0.153	-0.002		6024
	(0.359)	(0.360)	(0.009)		
Reserved for Women	0.641	0.232	0.409***		6024
	(0.480)	(0.422)	(0.012)		
Dependent variables					
Sarpanch ST	0.207	0.213	-0.006	0.005	6024
	(0.405)	(0.409)	(0.011)	(0.005)	
Sarpanch SC	0.216	0.133	0.083***	-0.000	6024
	(0.411)	(0.340)	(0.010)	(0.003)	
Sarpanch OBC	0.422	0.398	0.024	0.077***	6024
	(0.494)	(0.490)	(0.013)	(0.010)	
Sarpanch General	0.156	0.257	-0.101***	-0.082***	6024
	(0.363)	(0.437)	(0.010)	(0.010)	
Sarpanch Male	0.294	0.740	-0.445***	-0.090***	6024
	(0.456)	(0.439)	(0.012)	(0.006)	
Demographic variables					
Total population	5405.818	5511.806	-105.989*	-132.481**	6024
	(1724.568)	(2015.388)	(48.721)	(53.482)	
Proportion SC population	0.190	0.186	0.005	0.006*	6024
	(0.109)	(0.109)	(0.003)	(0.003)	
Proportion ST population	0.158	0.182	-0.024***	-0.032***	6024
	(0.237)	(0.266)	(0.007)	(0.006)	
Literacy rate	0.611	0.627	-0.016***	-0.021***	6024
	(0.091)	(0.092)	(0.002)	(0.003)	
Female literacy rate	0.451	0.472	-0.021***	-0.027***	6024
	(0.098)	(0.100)	(0.003)	(0.003)	
Other indicators					
No. of primary schools	1.864	2.051	-0.187***	-0.231***	5897
	(1.839)	(2.173)	(0.053)	(0.058)	
No. of middle schools	1.238	1.419	-0.181***	-0.235***	5897
	(1.592)	(1.888)	(0.046)	(0.050)	
No. of secondary schools	0.683	0.819	-0.136***	-0.176***	5897
	(1.120)	(1.352)	(0.032)	(0.036)	
No. of senior secondary schools	0.271	0.370	-0.099***	-0.122***	5897
	(0.715)	(0.922)	(0.021)	(0.024)	
Average years of education of population	3.615	3.840	- 0.225***	-0.279***	5161
	(1.023)	(1.062)	(0.030)	(0.032)	
Proportion of deprived households	0.474	0.464	0.010*	0.008*	5161
	(0.159)	(0.166)	(0.005)	(0.005)	
N	3685.000	2339.000	6024.000		

This table gives the baseline means by treatment status for all variables. Column 4 shows the coefficient of a dummy indicating treatment status from a regression of sarpanch caste/gender before legislation on the treatment dummy after controlling for reservation status. Standard errors in parentheses.

Table 4.3: Main results-Sarpanch caste I

	Sarpanch ST			Sarpanch SC		
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Treat	-0.0160***	-0.0147**	-0.0177***	0.0010	0.0013	-0.0003
	(0.0060)	(0.0062)	(0.0064)	(0.0042)	(0.0041)	(0.0044)
Post	0.0140***	-0.1484***	-0.0722	-0.0011	0.0080	-0.0655
	(0.0046)	(0.0409)	(0.0522)	(0.0030)	(0.0234)	(0.0646)
GP reserved for ST	0.7982***	0.7987***	0.7971***	-0.0188***	-0.0189***	-0.0200***
	(0.0121)	(0.0120)	(0.0117)	(0.0045)	(0.0044)	(0.0045)
GP reserved for SC	-0.0721***	-0.0718***	-0.0722***	0.9713***	0.9713***	0.9715***
	(0.0050)	(0.0050)	(0.0051)	(0.0037)	(0.0037)	(0.0037)
GP reserved for OBC	-0.0454***	-0.0451***	-0.0460***	-0.0170***	-0.0171***	-0.0166***
	(0.0035)	(0.0035)	(0.0036)	(0.0026)	(0.0026)	(0.0026)
GP reserved for women	-0.0029	-0.0026	-0.0032	0.0008	0.0008	0.0003
	(0.0033)	(0.0033)	(0.0033)	(0.0024)	(0.0024)	(0.0024)
Total population*Post		0.0000	0.0000		-0.0000	-0.0000
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Proportion of SC pop.*Post		0.0109*	-0.0117		-0.0007	0.0007
		(0.0061)	(0.0075)		(0.0074)	(0.0102)
Proportion of ST pop.*Post		0.0128***	0.0178**		-0.0019	0.0033
		(0.0038)	(0.0077)		(0.0016)	(0.0032)
Literacy rate*Post		0.1064***	0.0570		-0.0121	-0.0056
		(0.0294)	(0.0394)		(0.0165)	(0.0229)
Female literacy rate*Post		-0.0843***	-0.0377		0.0136	0.0146
		(0.0259)	(0.0352)		(0.0154)	(0.0208)
Observations	12048	12048	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Trends by 2011 GP characteristics	No	Yes	Yes	No	Yes	Yes
Block specific time trends	No	No	Yes	No	No	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4: Main results-Sarpanch caste II

	Sarpanch OBC			Sarpanch General		
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Treat	-0.0557***	-0.0562***	-0.0534***	0.0707***	0.0697***	0.0713***
	(0.0126)	(0.0128)	(0.0131)	(0.0123)	(0.0125)	(0.0128)
Post	0.0427***	0.0681	0.1708	-0.0556***	0.0723	-0.0332
	(0.0097)	(0.0597)	(0.1625)	(0.0096)	(0.0570)	(0.1509)
GP reserved for ST	-0.3999***	-0.4003***	-0.3983***	-0.3794***	-0.3795***	-0.3788***
	(0.0126)	(0.0126)	(0.0127)	(0.0128)	(0.0128)	(0.0129)
GP reserved for SC	-0.5475***	-0.5485***	-0.5479***	-0.3517***	-0.3509***	-0.3513***
	(0.0104)	(0.0104)	(0.0104)	(0.0099)	(0.0099)	(0.0099)
GP reserved for OBC	0.4498***	0.4496***	0.4516***	-0.3874***	-0.3874***	-0.3890***
	(0.0109)	(0.0109)	(0.0109)	(0.0108)	(0.0108)	(0.0107)
GP reserved for women	-0.0111	-0.0113	-0.0099	0.0131*	0.0131*	0.0127*
	(0.0072)	(0.0072)	(0.0072)	(0.0070)	(0.0070)	(0.0070)
Total population*Post		-0.0000	-0.0000		0.0000	0.0000
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Proportion of SC pop.*Post		0.0097	0.0308*		-0.0199*	-0.0198
		(0.0124)	(0.0164)		(0.0118)	(0.0161)
Proportion of ST pop.*Post		-0.0067	-0.0179**		-0.0041	-0.0032
		(0.0043)	(0.0082)		(0.0041)	(0.0080)
Literacy rate*Post		-0.0266	-0.0749		-0.0677*	0.0235
		(0.0435)	(0.0704)		(0.0411)	(0.0687)
Female literacy rate*Post		0.0316	0.0492		0.0390	-0.0261
		(0.0414)	(0.0658)		(0.0394)	(0.0647)
Observations	12048	12048	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Trends by 2011 GP characteristics	No	Yes	Yes	No	Yes	Yes
Block specific time trends	No	No	Yes	No	No	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.5: Main results-Sarpanch gender

	Sarpanch male		
	(1)	(2)	(3)
Post*Treat	0.0821***	0.0836***	0.0829***
	(0.0082)	(0.0083)	(0.0082)
Post	-0.0262***	-0.0695*	0.0870
	(0.0052)	(0.0387)	(0.0905)
GP reserved for ST	-0.0036	-0.0033	-0.0015
	(0.0082)	(0.0082)	(0.0081)
GP reserved for SC	0.0044	0.0049	0.0061
	(0.0060)	(0.0061)	(0.0059)
GP reserved for OBC	0.0104	0.0105	0.0106
	(0.0067)	(0.0067)	(0.0065)
GP reserved for women	-0.9116***	-0.9113***	-0.9141***
	(0.0048)	(0.0048)	(0.0047)
Total population*Post		-0.0000	-0.0000
		(0.0000)	(0.0000)
Proportion of SC pop.*Post		-0.0058	-0.0055
		(0.0069)	(0.0092)
Proportion of ST pop.*Post		0.0043	0.0051
		(0.0036)	(0.0056)
Literacy rate*Post		0.0297	-0.0290
		(0.0284)	(0.0433)
Female literacy rate*Post		-0.0138	0.0405
		(0.0255)	(0.0391)
Observations	12048	12048	12048
GP fixed effects	Yes	Yes	Yes
Trends by 2011 GP characteristics	No	Yes	Yes
Block specific time trends	No	No	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.6: Pre-existing trends in Sarpanch caste and gender- 2005-2010

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch general	(5) Sarpanch male
Post(=1 if 2010)*Treat	-0.0028	0.0010	0.0599***	-0.0582***	-0.0663***
	(0.0058)	(0.0038)	(0.0123)	(0.0123)	(0.0075)
Post(=1 if 2010)	-0.0064	-0.0005	-0.0392***	0.0461***	0.0053
	(0.0043)	(0.0028)	(0.0093)	(0.0093)	(0.0046)
GP reserved for ST	0.8333***	-0.0231***	-0.4277***	-0.3826***	-0.0030
	(0.0110)	(0.0040)	(0.0128)	(0.0127)	(0.0089)
GP reserved for SC	-0.0762***	0.9777***	-0.5243***	-0.3773***	0.0076
	(0.0053)	(0.0031)	(0.0103)	(0.0100)	(0.0061)
GP reserved for OBC	-0.0505***	-0.0165***	0.4530***	-0.3860***	0.0041
	(0.0040)	(0.0026)	(0.0109)	(0.0106)	(0.0064)
GP reserved for women	0.0054	0.0005	-0.0094	0.0035	-0.9113***
	(0.0042)	(0.0027)	(0.0087)	(0.0086)	(0.0055)
Observations	12030	12030	12030	12030	12030
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Trends by 2011 GP characteristics	No	No	No	No	No
Block specific time trends	No	No	No	No	No

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. Post(=1 if 2010)*Treat is the coefficient of interest. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.7: Sarpanch caste with gender

	(1) Sarpanch ST male	(2) Sarpanch ST female	(3) Sarpanch SC male	(4) Sarpanch SC female
Post*Treat	0.0140	-0.0317***	-0.0197**	0.0194**
	(0.0100)	(0.0102)	(0.0084)	(0.0085)
Observations	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	(1) Sarpanch OBC male	(2) Sarpanch OBC female	(3) Sarpanch general male	(4) Sarpanch general female
Post*Treat	0.0209	-0.0743***	0.0676***	0.0037
	(0.0127)	(0.0130)	(0.0103)	(0.0112)
Observations	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.8: Robustness-Additional time trends

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0148**	0.0020	-0.0571***	0.0698***	0.0790***
	(0.0069)	(0.0046)	(0.0145)	(0.0141)	(0.0088)
No. of primary schools*Post	0.0009	0.0015**	0.0004	-0.0028	0.0006
	(0.0008)	(0.0006)	(0.0017)	(0.0017)	(0.0009)
No. of middle schools*Post	-0.0001	-0.0010	0.0024	-0.0013	-0.0010
	(0.0010)	(0.0007)	(0.0023)	(0.0023)	(0.0012)
No. of secondary schools*Post	-0.0024**	-0.0010	-0.0026	0.0060**	0.0003
	(0.0011)	(0.0008)	(0.0026)	(0.0026)	(0.0014)
No. of senior secondary schools*Post	0.0023*	0.0004	-0.0019	-0.0007	0.0019
	(0.0014)	(0.0011)	(0.0029)	(0.0028)	(0.0014)
Avg years of education pop.*Post	0.0018	-0.0019	-0.0000	0.0001	-0.0014
	(0.0018)	(0.0012)	(0.0038)	(0.0037)	(0.0022)
Proportion of deprived households*Post	-0.0025	-0.0080	0.0164	-0.0058	-0.0013
	(0.0097)	(0.0052)	(0.0164)	(0.0156)	(0.0094)
Observations	10108	10108	10108	10108	10108
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.9: Robustness- Trends by 2010 reservation status

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0155**	-0.0003	-0.0579***	0.0737***	0.0914***
	(0.0064)	(0.0045)	(0.0135)	(0.0132)	(0.0088)
Reserved for ST in 2010*Post	-0.0559**	0.0127	0.0125	0.0307	0.0086
	(0.0256)	(0.0119)	(0.0335)	(0.0322)	(0.0203)
Reserved for SC in 2010*Post	0.0020	-0.0092	-0.0236	0.0308	0.0273*
	(0.0121)	(0.0090)	(0.0255)	(0.0246)	(0.0152)
Reserved for OBC in 2010*Post	-0.0149*	-0.0079	0.0098	0.0130	0.0056
	(0.0087)	(0.0071)	(0.0257)	(0.0252)	(0.0154)
Reserved for Women in 2010*Post	-0.0097	0.0006	0.0355	-0.0264	-0.0657***
	(0.0114)	(0.0070)	(0.0221)	(0.0216)	(0.0134)
Observations	12048	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficient of $Post * Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.10: Robustness- Proportion of candidates less than Class $8^{th}(5^{th})$ pass excluding Sarpanch

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0159**	-0.0000	-0.0547***	0.0707***	0.0833***
	(0.0065)	(0.0047)	(0.0138)	(0.0136)	(0.0088)
Post*Proportion candidates less than Class 8^{th}					
(5^{th}) pass excluding Sarpanch(top 5)	-0.0116	-0.0013	0.0131	-0.0002	-0.0015
	(0.0087)	(0.0060)	(0.0189)	(0.0185)	(0.0126)
Observations	12022	12022	12022	12022	12022
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.11: Robustness- Alternative independent variable I

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Proportion of candidates less than Class 8^{th} (5^{th}) pass 2010 (top 5)	-0.0277***	-0.0004	-0.0411**	0.0692***	0.0809***
	(0.0098)	(0.0061)	(0.0197)	(0.0193)	(0.0126)
Observations	12022	12022	12022	12022	12022
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The interaction of post with proportion of candidates less than Class 8(Class 5) pass 2010 is the coefficient of interest. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.12: Robustness- Alternative independent variable II

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Vote share of candidates less than Class 8^{th} (5^{th}) pass 2010 (top 5)	-0.0297***	-0.0027	-0.0482***	0.0805***	0.0917***
	(0.0090)	(0.0058)	(0.0184)	(0.0181)	(0.0115)
Observations	12022	12022	12022	12022	12022
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The interaction of post with vote share of candidates less than Class 8(Class 5) pass 2010 is the coefficient of interest. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.13: Matched sample using PSM- Baseline summary statistics by treatment status

	Sarpanch disqualified	Sarpanch not disqualified	Difference
Total population	5490.858	5491.452	-0.594
	(1819.913)	(1927.111)	(60.890)
Proportion SC population	0.196	0.196	-0.000
	(0.108)	(0.107)	(0.003)
Proportion ST population	0.145	0.149	-0.005
	(0.226)	(0.225)	(0.007)
Literacy rate	0.632	0.631	0.001
	(0.088)	(0.087)	(0.003)
Female literacy rate	0.476	0.475	0.001
	(0.094)	(0.095)	(0.003)
No. of primary schools	1.992	1.987	0.005
	(1.984)	(2.035)	(0.065)
No. of middle schools	1.385	1.365	0.020
	(1.700)	(1.758)	(0.056)
No. of secondary schools	0.794	0.759	0.034
	(1.212)	(1.215)	(0.039)
No. of senior secondary schools	0.335	0.320	0.014
	(0.762)	(0.783)	(0.025)
Average years of education of population	3.795	3.787	0.008
	(1.041)	(1.034)	(0.034)
Proportion of deprived households	0.464	0.466	-0.003
	(0.165)	(0.167)	(0.005)
N	1895.000	1895.000	3790.000

Table shows balance in all GP level variables used for propensity score matching exercise by treatment status in the matched sample. The first and second columns show the means whereas the third column shows the difference in means. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.14: Matched sample analysis using PSM- Sarpanch caste and gender

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0224***	0.0055	-0.0506***	0.0675***	0.0728***
	(0.0079)	(0.0057)	(0.0164)	(0.0159)	(0.0103)
Post	-0.0642	-0.0283	0.0836	0.0089	0.1764
	(0.0715)	(0.0506)	(0.2290)	(0.2204)	(0.1435)
GP reserved for ST	0.7923***	-0.0196***	-0.3921***	-0.3807***	-0.0036
	(0.0157)	(0.0060)	(0.0167)	(0.0170)	(0.0095)
GP reserved for SC	-0.0702***	0.9729***	-0.5460***	-0.3567***	0.0086
	(0.0063)	(0.0046)	(0.0131)	(0.0125)	(0.0071)
GP reserved for OBC	-0.0439***	-0.0174***	0.4390***	-0.3777***	0.0037
	(0.0044)	(0.0036)	(0.0137)	(0.0134)	(0.0080)
GP reserved for women	-0.0044	-0.0024	-0.0119	0.0187**	-0.9193***
	(0.0042)	(0.0034)	(0.0095)	(0.0093)	(0.0061)
Observations	7580	7580	7580	7580	7580
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficient of $Post * Treat$ is the difference-in-differences estimate for the matched sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.15: Heterogeneity by overall literacy rate

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0089	0.0056	-0.0669***	0.0703***	0.1041***
	(0.0104)	(0.0064)	(0.0224)	(0.0222)	(0.0149)
Post*Treat*Lit quintile 2	-0.0147	-0.0045	-0.0050	0.0242	-0.0207
	(0.0100)	(0.0069)	(0.0239)	(0.0234)	(0.0157)
Post*Treat*Lit quintile 3	-0.0050	-0.0113	0.0156	0.0008	-0.0150
	(0.0125)	(0.0083)	(0.0258)	(0.0254)	(0.0169)
Post*Treat*Lit quintile 4	0.0017	-0.0036	-0.0055	0.0074	-0.0173
	(0.0141)	(0.0087)	(0.0280)	(0.0270)	(0.0183)
Post*Treat*Lit quintile 5	-0.0117	-0.0021	0.0462	-0.0324	-0.0472**
	(0.0147)	(0.0091)	(0.0322)	(0.0314)	(0.0203)
Observations	12048	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	No	No	No	No	No
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficients of interactions of $Post * Treat$ with literacy quintiles give the difference between the DID estimate for that quintile and the DID estimate for the lowest literacy quintile which forms the base category. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.16: Heterogeneity by female literacy rate

	(1) Sarpanch ST	(2) Sarpanch SC	(3) Sarpanch OBC	(4) Sarpanch General	(5) Sarpanch Male
Post*Treat	-0.0073	0.0022	-0.0627***	0.0678***	0.0974***
	(0.0115)	(0.0071)	(0.0230)	(0.0228)	(0.0138)
Post*Treat*Fem lit quintile 2	-0.0158	-0.0013	-0.0000	0.0172	0.0056
	(0.0110)	(0.0071)	(0.0235)	(0.0230)	(0.0154)
Post*Treat*Fem lit quintile 3	-0.0044	-0.0047	0.0113	-0.0022	-0.0121
	(0.0130)	(0.0084)	(0.0259)	(0.0255)	(0.0157)
Post*Treat*Fem lit quintile 4	-0.0017	0.0021	0.0013	-0.0018	-0.0261
	(0.0141)	(0.0093)	(0.0285)	(0.0277)	(0.0170)
Post*Treat*Fem lit quintile 5	-0.0147	-0.0011	0.0188	-0.0029	-0.0318*
	(0.0148)	(0.0096)	(0.0336)	(0.0330)	(0.0190)
Observations	12048	12048	12048	12048	12048
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Election year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	No	No	No	No	No
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression with a different dependent variable. The coefficients of interactions of $Post * Treat$ with female literacy quintiles give the difference between the DID estimate for that quintile and the DID estimate for the lowest female literacy quintile which forms the base category. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.17: Performance under NREGS

	Limited years		All years	
	(1) Person days work (p.c)	(2) Person days work (p.c)	(3) Person days work (p.c)	(4) Person days work (p.c)
Post*Treat	-0.0051		-0.0119	
	(0.0792)		(0.0613)	
Year 2013*Treat		-0.0844		-0.0865
		(0.0678)		(0.0678)
Year 2014*Treat				-0.0376
				(0.0722)
Year 2015*Treat				-0.0588
				(0.0829)
Year 2016*Treat				-0.0253
				(0.0874)
Year 2017*Treat		-0.0375		-0.0547
		(0.0896)		(0.0877)
Year 2018*Treat		-0.0591		-0.0784
		(0.0953)		(0.0944)
Observations	21612	21612	37821	37821
GP fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. The first two columns show results from specifications (3) and (2) respectively, with sample limited to years 2012, 2013, 2017 and 2018 to compare years before and after legislation that correspond to the same period in the electoral cycle. Columns 3 & 4 show results from specifications (3) and (2) respectively, for data covering all years from 2012 to 2018. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.18: Night time luminosity (VIIRS)

	Limited years		All years	
	(1) Sum of night lights(p.c.)	(2) Sum of night lights(p.c.)	(3) Sum of night lights(p.c.)	(4) Sum of night lights(p.c.)
Post*Treat	0.0001		-0.0000	
	(0.0002)		(0.0001)	
Year 2013*Treat		-0.0002**		-0.0002**
		(0.0001)		(0.0001)
Year 2014*Treat		-0.0003		-0.0003
		(0.0002)		(0.0002)
Year 2015*Treat				-0.0004
				(0.0003)
Year 2016*Treat				-0.0003
				(0.0002)
Year 2017*Treat		-0.0002		-0.0002
		(0.0002)		(0.0002)
Year 2018*Treat		0.0001		0.0001
		(0.0002)		(0.0002)
Year 2019*Treat		-0.0002		-0.0002
		(0.0003)		(0.0003)
Observations	33600	33600	44800	44800
GP fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. The first two columns show results from specifications (3) and (2) respectively, with sample limited to years 2012, 2013, 2014 2017, 2018 and 2019 to compare years before and after legislation that correspond to the same period in the electoral cycle. Columns 3 & 4 show results from specifications (3) and (2) respectively, for data covering all years from 2012 to 2019. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.19: New school openings (DISE) I

	(1) Total new school openings	(2) New primary school openings	(3) New middle school openings	(4) New secondary school openings	(5) New sen. secondary school openings
Year 2011*Treat	-0.0205	-0.0144**	-0.0039	-0.0047	0.0020
	(0.0138)	(0.0072)	(0.0097)	(0.0038)	(0.0033)
Year 2012*Treat	-0.0043	0.0045	-0.0039	-0.0071*	0.0015
	(0.0132)	(0.0077)	(0.0087)	(0.0037)	(0.0033)
Year 2013*Treat	0.0210	0.0124	0.0034	0.0000	0.0028
	(0.0158)	(0.0118)	(0.0086)	(0.0034)	(0.0032)
Year 2014*Treat	-0.0064	-0.0006	-0.0033	-0.0056	0.0013
	(0.0139)	(0.0081)	(0.0091)	(0.0036)	(0.0039)
Year 2015*Treat	-0.0122	-0.0080	-0.0070	-0.0020	0.0018
	(0.0155)	(0.0098)	(0.0097)	(0.0035)	(0.0040)
Year 2016*Treat	0.0013	0.0016	0.0027	-0.0042	-0.0005
	(0.0147)	(0.0086)	(0.0095)	(0.0036)	(0.0040)
Year 2017*Treat	0.0006	0.0004	0.0005	-0.0038	0.0021
	(0.0152)	(0.0085)	(0.0105)	(0.0036)	(0.0036)
Observations	43504	43504	43504	43504	43504
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression based on specification (2) with a different dependent variable. Data includes all years from 2010 to 2017. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.20: New school openings (DISE) II

	(1) Total new school openings	(2) New primary school openings	(3) New middle school openings	(4) New secondary school openings	(5) New sen. secondary school openings
Post*Treat	-0.0018	-0.0027	0.0002	0.0003	-0.0004
	(0.0087)	(0.0060)	(0.0055)	(0.0015)	(0.0022)
Observations	43504	43504	43504	43504	43504
GP fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Block specific time trends	Yes	Yes	Yes	Yes	Yes
Trends by baseline GP demographics	Yes	Yes	Yes	Yes	Yes
Controls for reservation status	Yes	Yes	Yes	Yes	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression based on specification (3) with a different dependent variable. Data includes all years from 2010 to 2017. Post =1 for years 2015 to 2017 and 0 otherwise. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.A Appendix to Chapter 4

Table 4.A1: Main results-Sarpanch gender in Scheduled Areas

	Sarpanch male		
	(1)	(2)	(3)
Post*Treat	0.1732***	0.1795***	0.1828***
	(0.0388)	(0.0388)	(0.0390)
Post	-0.0746***	-0.2338	0.0747
	(0.0235)	(0.1738)	(0.2383)
GP reserved for women	-0.8558***	-0.8540***	-0.8510***
	(0.0192)	(0.0192)	(0.0196)
Total population*Post		-0.0000	-0.0000
		(0.0000)	(0.0000)
Proportion of SC pop.*Post		-0.0516	-0.0699
		(0.0951)	(0.1020)
Proportion of ST pop.*Post		0.0272	0.0056
		(0.0184)	(0.0232)
Literacy rate*Post		-0.0322	-0.1299
		(0.1238)	(0.1381)
Female literacy rate*Post		0.0854	0.1223
		(0.1241)	(0.1358)
Observations	984	984	984
GP fixed effects	Yes	Yes	Yes
Trends by 2011 GP characteristics	No	Yes	Yes
Block specific time trends	No	No	Yes

Standard errors are clustered at GP level and given in parentheses. Each column denotes result from a separate regression. Sample is limited to scheduled areas where all seats are reserved for STs. The coefficient of $Post{*}Treat$ is the difference-in-differences estimate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

Afridi, F., Bhattacharya, S., Dhillon, A., and Solan, E. (2021). Electoral competition, accountability and corruption: Theory and evidence from india.

Afridi, F., Iversen, V., and Sharan, M. (2017). Women political leaders, corruption, and learning: Evidence from a large public program in india. *Economic Development and Cultural Change*, 66(1):1–30.

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.

Aidt, T. S., Golden, M. A., and Tiwari, D. (2011). Incumbents and criminals in the indian national legislature.

Alesina, A. F., Troiano, U., and Cassidy, T. (2015). Old and young politicians. Technical report, National Bureau of Economic Research.

Algan, Y. and Cahuc, P. (2010). Inherited trust and growth. *American Economic Review*, 100(5):2060–92.

Amaral, S., Bandyopadhyay, S., Bhattacharya, S., and Sensarma, R. (2014). Crime and social conflict in india. *The Economics of Peace and Security Journal*, 9(1).

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Ansolabehere, S. and Gerber, A. (1996). The effects of filing fees and petition requirements on u. s. house elections. *Legislative Studies Quarterly*, 21(2):249–264.

Ansolabehere, S., Gerber, A., and Snyder, J. (2002). Equal votes, equal money: Court-ordered redistricting and public expenditures in the american states. *American Political Science Review*, 96(4):767–777.

Anukriti, S. and Chakravarty, A. (2019). Democracy and demography societal effects of fertility limits on local leaders. *Journal of Human Resources*, 54(1):79–121.

Asher, S., Lunt, T., Matsuura, R., and Novosad, P. (2021). Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform. World Bank Economic Review.

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

Asher, S. and Novosad, P. (2018). Rent-seeking and criminal politicians: Evidence from mining booms. *Available at SSRN 2812315*.

Atlas, C. M., Gilligan, T. W., Hendershott, R. J., and Zupan, M. A. (1995). Slicing the federal government net spending pie: Who wins, who loses, and why. *The American Economic Review*, 85(3):624–629.

Atlas, C. M., Hendershott, R. J., et al. (1997). Optimal effort allocation by us senators: The role of constituency size. *Public Choice*, 92(3):221–229.

Baltrunaite, A., Bello, P., Casarico, A., and Profeta, P. (2014). Gender quotas and the quality of politicians. *Journal of Public Economics*, 118:62–74.

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

Bardhan, P. (2002). Decentralization of governance and development. *Journal of Economic perspectives*, 16(4):185–205.

Bardhan, P., Mookherjee, D., and Torrado, M. (2010). Impact of political reservations in west bengal local governments on anti-poverty targeting. *Journal of Globalization and Development*, 1:5–5.

Bardhan, P. K. and Mookherjee, D. (2000). Capture and governance at local and national levels. *American economic review*, 90(2):135–139.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias?*. *The Quarterly Journal of Economics*, 124(4):1497–1540.

Besley, T., Montalvo, J. G., and Reynal-Querol, M. (2011). Do educated leaders matter?*. *The Economic Journal*, 121(554):F205–227.

Besley, T., Pande, R., Rahman, L., and Rao, V. (2004). The Politics of Public Good Provision: Evidence from Indian Local Governments. *Journal of the European Economic Association*, 2(2-3):416–426.

Besley, T., Pande, R., and Rao, V. (2012). Just rewards? local politics and public resource allocation in south india. *The World Bank Economic Review*, 26(2):191–216.

Besley, T., Persson, T., and Sturm, D. M. (2010). Political competition, policy and growth: theory and evidence from the us. *The Review of Economic Studies*, 77(4):1329–1352.

Besley, T. J., Pande, R., and Rao, V. (2005). Political selection and the quality of government: Evidence from south india.

Bhalotra, S., Clots-Figueras, I., and Iyer, L. (2018). Pathbreakers? women's electoral success and future political participation. *The Economic Journal*, 128(613):1844–1878.

Bhavnani, R. (2016). The effects of malapportionment on economic development: Evidence from india's 2008 redistricting.

Bhavnani, R. R. (2009). Do electoral quotas work after they are withdrawn? evidence from a natural experiment in india. *The American Political Science Review*, 103(1):23–35.

Bhavnani, R. R. (2018). The effects of malapportionment on cabinet inclusion: Subnational evidence from india. *British Journal of Political Science*, 48(1):69–89.

Bhogale, S., Hangal, S., Jensenius, F., Kumar, M., Narayan, C., U Nissa, B., and Verniers, G. (2019). Tcpd indian elections data v1.

BusinessLine (2014). Rajasthan panchayat ordinance on minimum education for poll contestants draws flak. https://www.thehindubusinessline.com/news/national/rajasthan-panchayat-ordinance-on-minimum-education-for-poll-contestants-draws-flak/article23122330.ece.

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

Castelló-Climent, A., Chaudhary, L., and Mukhopadhyay, A. (2018). Higher education and prosperity: From catholic missionaries to luminosity in india. *The Economic Journal*, 128(616):3039–3075.

Castelló-Climent, A. and Mukhopadhyay, A. (2013). Mass education or a minority well educated elite in the process of growth: The case of india. *Journal of Development Economics*, 105:303–320.

Catterberg, G. and Moreno, A. (2006). The individual bases of political trust: Trends in new and established democracies. *International journal of public opinion research*, 18(1):31–48.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.

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

Cheng, C.-Y. and Urpelainen, J. (2019). Criminal politicians and socioeconomic development: Evidence from rural india. *Studies in Comparative International Development*, 54(4):501–527.

Chin, A. and Prakash, N. (2011). The redistributive effects of political reservation for minorities: Evidence from india. *Journal of Development Economics*, 96(2):265–277.

Citrin, J. (1974). Comment: The political relevance of trust in government. *American Political Science Review*, 68(3):973–988.

Clausen, B., Kraay, A., and Nyiri, Z. (2011). Corruption and confidence in public institutions: Evidence from a global survey. *The World Bank Economic Review*, 25(2):212–249.

Coate, S. and Knight, B. (2007). Socially optimal districting: a theoretical and empirical exploration. *The Quarterly Journal of Economics*, 122(4):1409–1471.

Das, S., Mukhopadhyay, A., and Saroy, R. (2017). Efficiency consequences of affirmative action in politics: Evidence from india. IZA Discussion Paper No. 11093, Available at SSRN: https://ssrn.com/abstract=3056631.

Daxecker, U. (2019). Unequal votes, unequal violence: Malapportionment and election violence in india. *Journal of Peace Research*, page 0022343319884985.

DeccanHerald (2021). Mla seeks rs 2,000 crore for higher education department.

Denters, B. (2002). Size and political trust: evidence from denmark, the netherlands, norway, and the united kingdom. *Environment and Planning C: Government and Policy*, 20(6):793–812.

Desai, S. and Vanneman, R. (2018). India human development survey-ii (ihds-ii), 2011-12.

Desai, S., Vanneman, R., and National Council of Applied Economic Research, N. D. (2018). India human development survey (ihds), 2005.

Dey, S. and Sen, K. (2016). Is partisan alignment electorally rewarding? evidence from village council elections in india. *Evidence from Village Council Elections in India (August 1, 2016). Effective States and Inclusive Development (ESID) Working Paper*, (63).

Diaz-Serrano, L. and Pérez, J. (2013). Do more educated leaders raise citizens' education?

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):169–193.

Drometer, M. and Rincke, J. (2009). The impact of ballot access restrictions on electoral competition: evidence from a natural experiment. *Public Choice*, 138(3-4):461–474.

Dunning, T. and Nilekani, J. (2013). Ethnic quotas and political mobilization: caste, parties, and distribution in indian village councils. *American political Science review*, 107(1):35–56.

Ferraz, C. and Finan, F. (2011). Motivating politicians: The impacts of monetary incentives on quality and performance.

Gangadharan, L., Jain, T., Maitra, P., and Vecci, J. (2016). Social identity and governance: The behavioral response to female leaders. *European Economic Review*, 90:302–325.

Gehring, K., Kauffeldt, T. F., and Vadlamannati, K. C. (2019). Crime, incentives and political effort: Evidence from india. *European Journal of Political Economy*, 59:1–20.

George, S., Gupta, S., and Neggers, Y. (2018). Coordinating voters against criminal politicians: Evidence from a mobile experiment in india. Technical report, Working paper.

Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American political science review*, 94(3):653–663.

Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research. *Electoral studies*, 25(4):637–663.

Gibson, J., Olivia, S., Boe-Gibson, G., and Li, C. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics*, 149:102602.

Grönlund, K. and Setälä, M. (2007). Political trust, satisfaction and voter turnout. *Comparative European Politics*, 5(4):400–422.

Gul, F. and Pesendorfer, W. (2010). Strategic redistricting. *American Economic Review*, 100(4):1616–41.

Gulzar, S. and Pasquale, B. J. (2017). Politicians, bureaucrats, and development: Evidence from india. *American Political Science Review*, 111(1):162–183.

Gupta, A. and Spears, D. (2017). Health externalities of india's expansion of coal plants: Evidence from a national panel of 40,000 households. *Journal of environmental economics and management*, 86:262–276.

Gupta, B. and Mukhopadhyay, A. (2016). Local funds and political competition: evidence from the national rural employment guarantee scheme in india. *European Journal of Political Economy*, 41:14–30.

Hansen, S., Palfrey, T. R., and Rosenthal, H. (1987). The downsian model of electoral participation: Formal theory and empirical analysis of the constituency size effect. *Public Choice*, 52(1):15–33.

Hansen, S. W. (2013). Polity size and local political trust: A quasi-experiment using municipal mergers in denmark. *Scandinavian Political Studies*, 36(1):43–66.

Heß, S. et al. (2017). Randomization inference with stata: A guide and software. *Stata Journal*, 17(3):630–651.

Hetherington, M. J. (1998). The political relevance of political trust. *American Political Science Review*, 92(4):791–808.

Hetherington, M. J. (1999). The effect of political trust on the presidential vote, 1968–96. *American Political Science Review*, 93(2):311–326.

Hetherington, M. J., Hetherington, M. J., et al. (2005). *Why trust matters: Declining political trust and the demise of American liberalism*. Princeton University Press.

HindustanTimes (2014). Women's group opposes edu criteria for panchayat polls. https://www.hindustantimes.com/jaipur/women-s-group-opposes-edu-criteria-for-panchayat-polls/story-A7BfxJboDsGZHt4GXvS46N.html.

Hooghe, M. and Marien, S. (2013). A comparative analysis of the relation between political trust and forms of political participation in europe. *European Societies*, 15(1):131–152.

Horiuchi, Y. and Saito, J. (2003). Reapportionment and redistribution: Consequences of electoral reform in japan. *American Journal of Political Science*, 47(4):669–682.

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

IndianExpress (2014). Rajasthan governor fixes minimum education qualifications for panchayat polls. https://indianexpress.com/article/india/india-others/rajasthan-governor-fixes-minimum-education-qualifications-for-panchayat-polls/.

Iyer, L. and Mani, A. (2012). Traveling agents: political change and bureaucratic turnover in india. *Review of Economics and Statistics*, 94(3):723–739.

Iyer, L. and Reddy, M. (2013). *Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World's Largest Democracy?* Harvard Business School.

Jaffrelot, C. (2016). Laws prescribing educational qualifications for contesting elections are undemocratic. https://indianexpress.com/article/opinion/columns/narendra-modi-pm-degree-arvind-kejriwal-delhi-university-ba-degrees-of-exclusion-2792374/.

Jagnani, M. and Khanna, G. (2020). The effects of elite public colleges on primary and secondary schooling markets in india. *Journal of Development Economics*, 146:102512.

Jain, T., Gangadharan, L., and Maitra, P. (2015). Citizens' trust in local politicians and implications for good governance. *Ideas for India*.

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

Jensenius, F. R. and Verniers, G. (2017). Studying indian politics with large-scale data: Indian election data 1961–today. *Studies in Indian Politics*.

Jeong, D., Shenoy, A., and Zimmermann, L. V. (2021). Rethinking transparency: Corruption in local public office in india.

Kapur, D. and Mehta, P. B. (2007). Mortgaging the future. *India International Centre Quarterly*, 34(3/4):154–166.

Knight, B. (2008). Legislative representation, bargaining power and the distribution of federal funds: Evidence from the us congress. *The Economic Journal*, 118(532):1785–1803.

Kotakorpi, K. and Poutvaara, P. (2011). Pay for politicians and candidate selection: An empirical analysis. *Journal of Public Economics*, 95(7):877–885.

Krishnan, S. (2014). Political economy of india's tertiary education: Persistence and change. *Economic and Political Weekly*, pages 63–70.

Lahoti, R. and Sahoo, S. (2020a). Are educated leaders good for education? evidence from india. *Journal of Economic Behavior & Organization*, 176:42–62.

Lahoti, R. and Sahoo, S. (2020b). Are educated leaders good for education? evidence from india. *Journal of Economic Behavior & Organization*, 176:42–62.

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

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Levi, M. and Stoker, L. (2000). Political trust and trustworthiness. *Annual review of political science*, 3(1):475–507.

Levine, D. K. and Palfrey, T. R. (2007). The paradox of voter participation? a laboratory study. *American political science Review*, 101(1):143–158.

Listhaug, O. and Wiberg, M. (1995). Confidence in political and private institutions. *Citizens and the State*, 1:298–322.

Marien, S. and Hooghe, M. (2011). Does political trust matter? an empirical investigation into the relation between political trust and support for law compliance. *European Journal of Political Research*, 50(2):267–291.

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

McManus, J. (2014). Educational qualifications of village leaders in north india: Implications for public service delivery.

Mishler, W. and Rose, R. (2001). What are the origins of political trust? testing institutional and cultural theories in post-communist societies. *Comparative political studies*, 34(1):30–62.

Mori, Y. (2015). Voter turnout and the principle of 'one person, one vote': empirical evidence from the constituency freeze in india. *Applied Economics*, 47(37):3957–3970.

Nanda, V. K. and Pareek, A. (2016). Do criminal politicians affect firm investment and value? evidence from a regression discontinuity approach. *Evidence from a Regression Discontinuity Approach (May 11, 2016)*.

Nath, A. (2015). Bureaucrats and politicians: How does electoral competition affect bureaucratic performance? *Institute for Economic Development (IED) Working Paper*, 269(00016):12.

Newton, K. and Norris, P. (2000). Confidence in public institutions. *Disaffected democracies. What's troubling the trilateral countries*.

Nunn, N., Qian, N., and Wen, J. (2018). Distrust and political turnover. Technical report, National Bureau of Economic Research.

Oliver, J. E. (2000). City size and civic involvement in metropolitan america. *American Political Science Review*, 94(2):361–373.

Pande, R. (2015). Should the less educated be barred from village council elections? https://www.ideasforindia.in/topics/governance/should-the-less-educated-be-barred-from-village-council-elections.html.

Pharr, S. J. and Putnam, R. D. (2000). Officials' misconduct and public distrust: Japan and the liberal democracies.

Prakash, N., Rockmore, M., and Uppal, Y. (2019). Do criminally accused politicians affect economic outcomes? evidence from india. *Journal of Development Economics*, 141:102370.

Prakash, N., Sahoo, S., Saraswat, D., and Sindhi, R. (2021). When criminality begets crime: The role of elected politicians. *Ideas for India*.

Putnam, R. D. (2000). Bowling alone: America's declining social capital. In *Culture and politics*, pages 223–234. Springer.

Rahn, W. M. and Rudolph, T. J. (2005). A tale of political trust in american cities. *Public opinion quarterly*, 69(4):530–560.

Rudolph, T. J. and Evans, J. (2005). Political trust, ideology, and public support for government spending. *American Journal of Political Science*, 49(3):660–671.

Scholz, J. T. and Lubell, M. (1998). Trust and taxpaying: Testing the heuristic approach to collective action. *American Journal of Political Science*, pages 398–417.

Seligson, M. A. (2002). The impact of corruption on regime legitimacy: A comparative study of four latin american countries. *The journal of Politics*, 64(2):408–433.

Shachar, R. and Nalebuff, B. (1999). Follow the leader: Theory and evidence on political participation. *American Economic Review*, 89(3):525–547.

Stratmann, T. (2005). Ballot access restrictions and candidate entry in elections. *European Journal of Political Economy*, 21(1):59–71.

Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic association*, 8(4):677–716.

TheHindu (2019). 43% newly-elected lok sabha mps have criminal record: Adr. https://www.thehindu.com/elections/lok-sabha-2019/43-newly-elected-lok-sabha-mps-have-criminal-record-adr/article27253649.ece.

TheHindu (2021a). Construction of medical collegeat gurazala to begin soon. https://www.thehindu.com/news/national/andhra-pradesh/construction-of-medical-collegeat-gurazala-to-begin-soon/article34670658.ece.

TheHindu (2021b). TN government to consider pleas for arts and sciences colleges in madurai. https://www.thehindu.com/news/national/tamil-nadu/tn-government-to-consider-pleas-for-arts-and-sciences-colleges-in-madurai/article36271710.ece.

TheTimesofIndia (2013). DC asks official to identify land for medical college. https://www.babushahi.com/view-news.php?id=75588headline=MLA%27s-efforts-to-set-up-government-college-fructifies.

TheTimesofIndia (2018). UP government plans to open 44 new intermediate colleges named after atal bihari vajpayee. https://timesofindia.indiatimes.com/city/agra/up-govt-plans-to-open-44-new-inter-colleges-named-after-vajpayee/articleshow/65535258.cms.

TheTimesOfIndia (2020). Haryana cm announces plans to open 11 new colleges in the state.

TheTribune (2019). CM to lay stone of college at chabbewal soon, says mla. https://www.tribuneindia.com/news/archive/jalandhar/cm-to-lay-stone-of-college-at-chabbewal-soon-says-mla-732448.

TheTribune (2020). MLA visits construction site of college. https://www.tribuneindia.com/news/ludhiana/mla-visits-construction-site-of-college-115001.

TNM (2018). MLA's efforts to set up government college fructifies. https://timesofindia.indiatimes.com/city/ranchi/dc-asks-official-to-identify-land-for-medical-college/articleshow/27695324.cms?from=mdr.

Torgler, B. (2005). Tax morale and direct democracy. *European Journal of Political Economy*, 21(2):525–531.

TribuneNewsline (2018). Special secretary higher education visits jadla for setting up of government college. http://tribunenewsline.com/special-secretary-higher-education-visits-jadla-setting-government-college/.

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

Zak, P. J. and Knack, S. (2001). Trust and growth. *The economic journal*, 111(470):295–321.

Zimmermann, L. (2020). Why guarantee employment? evidence from a large indian public-works program. Technical report, GLO Discussion Paper.