



# Building connections: Political corruption and road construction in India

Jonathan Lehne, Jacob N. Shapiro, Oliver Vanden Eynde

► **To cite this version:**

Jonathan Lehne, Jacob N. Shapiro, Oliver Vanden Eynde. Building connections: Political corruption and road construction in India. PSE Working Papers n 2016-15. 2016. <halshs-01349350>

**HAL Id: halshs-01349350**

**<https://halshs.archives-ouvertes.fr/halshs-01349350>**

Submitted on 27 Jul 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2016 – 15**

**Building connections: Political corruption and  
road construction in India**

**Jonathan Lehne  
Jacob N. Shapiro  
Oliver Vanden Eynde**

**JEL Codes: D72, D73, L14, O18**

**Keywords:**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# BUILDING CONNECTIONS: POLITICAL CORRUPTION AND ROAD CONSTRUCTION IN INDIA \*

Jonathan Lehne<sup>†</sup> Jacob N. Shapiro<sup>‡</sup> Oliver Vanden Eynde<sup>§</sup>

05 July 2016

## Abstract

Politically-driven corruption is a pervasive challenge for development, but evidence of its welfare effects is scarce. Using data from a major rural road construction programme in India we document political influence in a setting where politicians have no official role in contracting decisions. Exploiting close elections to identify the causal effect of coming to power, we show that the share of contractors whose name matches that of the winning politician increases by 63% (from 4% to 6.4%). Regression discontinuity estimates at the road level show that political interference raises costs, lowers quality, and increases the likelihood that roads go missing.

JEL Codes: D72, D73, L14, O18

---

\* We would like to thank the International Growth Centre (IGC) and CEPREMAP for supporting the data collection of this project. We are also grateful to Nawal Aggarwal, Ashish Modi, Shrenik Sanghvi, Radha Sankar, and Paolo Santini for excellent research assistance. This paper has benefited from discussions with Sam Asher, Tarek Ghani, Guy Grossman, Dan Keniston, Ariane Lambert-Mogilianski, Karen Macours, Alexander Plekhanov, Akiko Suwa-Eisenmann, Liam Wren-Lewis, and Maiting Zhuang. We also thank seminar participants in PSE's IRG and India-China Workshops, the Namur Indian Political Economy Workshop, the Graduate Institute (Geneva), and the Sussex Development Workshop. All remaining errors are our own.

<sup>†</sup> Corresponding author, Paris School of Economics, jonathan.lehne@gmail.com.

<sup>‡</sup> Princeton University, Woodrow Wilson School of Public and International Affairs, jns@princeton.edu.

<sup>§</sup> Paris School of Economics, olivervandeneinde@gmail.com.

# 1. INTRODUCTION

A growing literature documents the private returns to holding public office in both the developed and developing world.<sup>1</sup> Fisman et al. (2014) find that the wealth of state-level politicians in India increases sharply when they narrowly win elections. And, politicians' personal characteristics and political ties clearly affect development outcomes (Prakash et. al. 2015, Asher and Novosad 2015). It is not clear from existing work whether political influence entails welfare-reducing corruption as opposed to politicians leveraging the influence of public office in beneficial ways.<sup>2</sup>

We provide evidence from a major public works program in India that there is politically driven reallocation among contractors with negative welfare consequences. Specifically, we use bidding data on more than 88,000 rural roads built under the Pradhan Mantri Gram Sadak Yojana (PMGSY) program to study how close-election victories shift spending. Using regression discontinuity (RD) estimates to identify the causal effect of coming to power we show in our preferred specification that the share of contractors whose name matches that of the winning politician increases from 4% to 6.4% (a 63% increase). The magnitude of the distortions are large relative to program size. Applying our RD estimate to the full sample (i.e. extrapolating from a LATE) would imply that state-level parliamentarians (MLAs) intervened in the allocation of roughly 1,600 of the 4,127 road contracts let to connected contractors, approximately \$470M of the \$1.2B spent on such roads and approximately 3% of the total spent on the program. These results are broadly representative of Indian polities. Our sample consists of 4,058 electoral terms from 2001 to 2013, of which 2,632 constituencies had PMGSY agreements signed both before and after the election, covering 24 of the 28 states which existed in our sample period.

The allocation of contracts to those with political connections does not conclusively prove that politicians' motives are corrupt. In an environment of imperfect information, MLAs could, in theory, be better informed about, and better able to monitor, contractors in their own network, which would imply improvements in road quality within connected contracts. This was not the case with PMGSY road construction. RD estimation at the road level, provides no evidence that

---

<sup>1</sup> See Eggers and Hainmuller (2009) for members of the UK House of Commons and Truex (2014) for Chinese deputies.

<sup>2</sup> The literature typically thinks of corruption as inefficient rent-seeking (e.g. Becker and Stigler, 1974; Krueger, 1974; Rose-Ackerman, 1975; Shleifer and Vishny, 1993), but under an alternative hypothesis it could actually help to overcome administrative burdens (Huntington 1968, Lui 1985).

political interventions promote efficiency or quality. On the contrary, roads allocated to connected contractors are both more expensive to construct and more likely to fail subsequent quality inspections. Note, the latter effect is the opposite of what we would expect if politicians were biasing the quality assurance process. These results suggest that political interference is costly and inefficient.

We also document direct negative welfare consequences for the people the program is supposed to serve. Census data at the village-level, collected after road construction was officially completed, reveal that a number of roads listed as having been completed in the PMGSY monitoring data, and for which payments were made, do not appear to exist. We define a road as “missing” if any village it was meant to reach subsequently lacked “all-weather road access” (PMGSY’s stated objective). The preferential allocation of roads appears to increase the likelihood of a missing all-weather road by 170%. The implied number of additional missing all-weather roads is 672, which would otherwise have served 1.16 million people.

These results are particularly striking since state-level legislators (MLAs) do not have any formal role in the allocation of contracts, but still manage to get more projects assigned to contractors who share their surname. Our paper’s first contribution is thus to demonstrate the existence of political influence in an environment where there is no formal institutional mechanism for it. In fact, this program’s bidding rules were designed in ways that should have forestalled political influence at the bidding stage (NRRDA 2015). The results of our paper are therefore of particular relevance for India and the functioning of its democratic institutions. A recent literature confirms the large influence of state legislators on economic outcomes. Asher and Novosad (2015) find higher employment in constituencies whose MLAs are aligned with the state-level government. Higher clearances of mining projects suggest that MLAs use their influence in the administration to push employment generating projects. An important source of influence for these politicians is their ability to reassign bureaucrats, as highlighted by Iyer and Mani (2012). Prakash et al. (2015) also confirm the economic importance of MLAs. These authors find that the election of criminal MLAs leads to lower economic growth in their constituencies. Finally, Fisman et al. (2014) show that the assets of marginally elected MLAs grow more than those of runner-ups, which confirms the idea that there are substantial private returns to holding public office in India. Compared to this recent literature, our paper sheds more light on the long causal chain that connects the characteristics of MLAs with aggregate economic outcomes. By showing how MLAs

use their power improperly to favour connected contractors, we provide micro-evidence on the channels of influence of these democratically elected politicians. Moreover, our paper suggests that not just the economic performance of the constituency is at stake when MLAs exert influence, through the preferential allocation of contracts they also affect the distribution of public goods.

Our second contribution is to demonstrate a new approach to quantify how much politicians influence contracting. The core challenges we confront in doing so are that: (1) there is no information on actual connections between politicians and the contractors active in their constituency; and (2), to the extent that politicians intervene in the allocation of roads on contractors' behalf, such improper interference would not be documented. We address the first problem by constructing a surname-based measure of proximity between candidates for state-level legislatures and contractors. This approach follows a number of papers that use Indian surnames as identifiers of caste or religion (e.g. Hoff and Pandey 2004, Field et al. 2008, Banerjee et al. 2014a). Dealing with the second issue – identifying improper intervention – requires isolating the variation in proximity to contractors that results from the MLA coming to power. We do so with a regression discontinuity approach that exploits the fact that in close elections, candidates who barely lost are likely to have similar characteristics to those who were barely elected. If MLAs are intervening in the assignment of contracts, one would expect a shift in the allocation towards contractors who share their name, and no equivalent shift for their unsuccessful opponents.

Our third contribution is to rule out one standard explanation for why politicians target patronage along in-group lines, which in India often means caste (Chandra 2004). Much of the literature posits that politicians allocate public goods to members of their own group as a form of vote-buying. We find no evidence that the preferential allocation of roads or the cost inflations increase immediately before or after election dates, so if vote-buying is going on it must be a long-run transaction. We also exploit India's 2008 re-drawing of electoral constituency boundaries to study the behaviour of MLAs in regions that have become "politically irrelevant" after the redistricting. We find no evidence of different behaviour in these regions. Our results are more consistent with either standard in-group favouritism or a subtler mechanism by which caste or kinship networks facilitate corrupt exchange. The latter of these possibilities fits our findings and the context of PMGSY. The involvement of the central government in the program guarantees a minimum level of monitoring. In line with the idea that contractors trade off rent-seeking and the cost of detection, we find no evidence that preferential allocation affects the performance markers

that are most easily observed in the administrative data collected at the central level: over-runs and delays.

The remainder of the paper proceeds as follows. Section 2 reviews the literature on political corruption in public goods provision and discusses our contribution to that literature. Section 3 provides context on PMGSY, the role of MLAs, and Indian surnames as identifiers of caste or religion. Section 4 describes the dataset used in the analysis. Section 5 outlines the empirical strategy. Section 6 presents the main results on re-allocation and robustness. Section 7 analyses the social costs of re-allocation. Section 8 rules out election cycles. Section 9 examines whether the introduction of e-procurement in PMGSY helped to reduce political influence. Section 10 concludes.

## **2. LITERATURE**

Our paper relates to a large literature on corruption, political connections, and ethnic favouritism. Theoretically, corruption is typically thought of as rent-seeking. Public officials use their control over the allocation of contracts or the provision of services to ask for bribes (e.g. Becker and Stigler, 1974; Krueger, 1974; Rose-Ackerman, 1975; Shleifer and Vishny, 1993). This behaviour is most likely to arise in contexts where enforcement is weak and officials are poorly remunerated.<sup>3</sup> The welfare effects of such behaviour are debated. The so-called “greasing the wheels” hypothesis argues that corruption can be optimal in a second-best world, by allowing agents to circumvent inefficient institutions and regulation (Huntington 1968, Lui 1985). In principle, both arguments could apply to the preferential assignment of PMGSY roads by Indian MLAs. However, the evidence we present on cost inflation, quality defects and missing roads under preferentially allocated contracts supports the rent-seeking hypothesis.

A growing number of papers seek to provide objective, quantitative estimates of corruption (Banerjee et al., 2012). Underlining the benefits of such an approach, Olken (2009) finds that villagers’ assessments of corruption correlate only weakly with an actual measure of missing expenditures in the context of rural road construction in Indonesia. In the case of PMGSY, there

---

<sup>3</sup> In the case of Indian MLAs, calculating efficiency wages (as suggested by Becker and Stigler, 1974) may be complicated by the fact that candidates frequently need to pay their parties significant sums for their place on the ticket. This could prompting them to engage in corrupt behaviour once elected (Jensenius,2013).

is no publicly available audit data that would provide a direct measure of corruption.<sup>4</sup> We therefore employ an approach that Banerjee et al. (2012) refer to as “cross-checking”: the comparison between (i) an actually observed outcome, and (ii) a counterfactual measure which should be equivalent to the former in the absence of corruption. In our setting, if politicians are not intervening in the allocation of road projects, they should be no ‘closer’ to contractors than their unsuccessful opponents. This empirical strategy is close to that of Do et al. (2013), who use a regression discontinuity design to compare the performance of firms connected to winning and losing candidates in close gubernatorial elections in the US. Other exponents of the “cross-checking” approach include Acemoglu et al. (2014), Golden and Picci (2005), Reinnika and Svensson (2004), Olken (2007), Fisman (2001), and Banerjee et al. (2014b). This literature offers more support for the rent-seeking than for the “greasing the wheels” hypothesis. Our findings on the characteristics of preferentially allocated roads point in the same direction.

Our paper complements existing work by providing direct evidence on the role of politicians. A small body of work considers the electoral incentives to engage in corrupt behaviour. Mironov and Zhuravskaya (2016) show that Russian firms who funnel money in the run-up to elections are significantly more likely to receive procurement contracts after the election. A different type of electoral incentives could arise in so-called “patronage-democracies”, where targeting patronage is easier within ethnic or caste groups (Chandra 2004, Horowitz 1985). Voters’ preference for patronage could motivate them to choose politicians of their own caste (Banerjee et al., 2014a).<sup>5</sup> While our paper documents the preferential allocation of road contracts, we find no evidence linking corrupt behaviour to electoral incentives. In that sense, our results might be most consistent with the hypothesis that caste and kinship networks facilitate corruption through trust and the ability to impose social sanctions (Lambsdorff 2002; Tonoyan, 2003). Corruption is illegal and therefore requires either trust among collaborators, or a predictable ability to sanction defections, both of which are more likely to exist between members of the same family, ethnic group, or network.

---

<sup>4</sup> Several countries conduct regular audits of local government expenditure and make the results publicly available. Examples of research based on these data include Ferraz and Finnan (2008 and 2011) and Melo et al. (2009) for Brazil, or Larreguy, Marshall and Snyder Jr (2014) for Mexico. Alternatively, studies can be designed to observe corruption independently (e.g. Bertrand et al., 2007; Olken and Barron, 2009).

<sup>5</sup> Ethnic favoritism is not specific to democracies. Burgess et al. (2014) show that the ethnic homelands of Kenyan presidents receive preferential coverage by road projects, but only under autocracy. Kramon and Posner (2016) show similar favouritism in schooling outcomes in Kenya, which survive in periods of democracy.



### 3. BACKGROUND

#### 3.1 PMGSY

In the year 2000, an estimated 330,000 Indian villages or habitations – out of a total of 825,000 – were not connected to a road that provided all-weather access (PMGSY 2004). Their inhabitants were at least partially cut-off from economic opportunities and public services (such as health care and education). To address this lack of connectivity, the Indian government launched the Pradhan Mantri Gram Sadak Yojana (PMGSY) in December 2000. Its goal was to ensure all-weather access to all habitations with populations over 1,000 by the year 2003, and to those with more than 500 inhabitants by 2007. In hill states, desert and tribal areas, as well as districts with Naxalite insurgent activity, habitations with a population over 250 were targeted (PMGSY 2004).

The programme has been described as “unprecedented in its scale and scope” (Aggarwal 2015), with roadwork for over 122,000 habitations completed and another 21,000 under construction as of June 2016.<sup>6</sup> A second phase of the scheme (PMGSY II), launched in 2013, targets all habitations with populations over 100. According to World Bank estimates, expenditures under PMGSY had reached 14.6 billion USD by the end of 2010, with a further 40 billion USD required for its completion by 2020 (World Bank, 2014).

Several studies have focused on the first-order research question that arises in relation to PMGSY: its impact on habitations and the lives of their inhabitants. Asher and Novosad (2016) analyse the employment effects of the programme in previously unconnected villages. They find that a new paved road raises participation in the wage labour market with a commensurate decrease in the share of workers employed in agriculture. Aggarwal (2015) also finds a positive effect on employment and reduced price dispersion among villages. While these studies analyse *what* PMGSY has achieved, this paper looks at *how* it has been implemented.

Compared to other public works programmes, the implementation of PMGSY stands out because of its reliance on private contractors combined with relatively strong monitoring and quality assurance provisions, designed to limit the scope for undue corruption. All tenders have to follow a competitive bidding procedure, of which the rules were prescribed by the National Rural Roads Development Agency (NRRDA) and set out in the so-called Standard Bidding Document

---

<sup>6</sup> OMMAS (Online Management, Monitoring and Accounting System), <http://omms.nic.in/>, accessed in July 2016.

(SBD). The SBD consists of a two envelope tendering process administered at the circle level. Each bid consist of both technical and financial volumes. The technical bids are opened first. Contractors have to fulfil eligibility criteria, taking into account factors such as their current workload and experience. Only the financial bids of contractors whose technical bids are found to meet the requirements are evaluated, and subject to meeting the technical standards the lowest bidder has to be selected. After the contract has been assigned, administrative data on the programme is gathered, while central and state-level inspectors can carry out quality inspections. In spite of these provisions, there remains clear scope for corruption, and the financial incentives are sizeable given the scale of the project.<sup>7</sup> A large number of newspaper reports document alleged corruption in PMGSY.<sup>8</sup> Corruption in PMGSY could take several forms, and the possible manipulation of road allocations is one of the challenges for impact evaluations of the programme (Asher and Novosad, 2016).<sup>9</sup> Our paper tests for a specific form of corruption: interventions by state-level parliamentarians (MLAs) in the allocation of road contracts (but not of the location of roads) within their constituencies.

An advantage of focussing on MLAs in this context is that under the programme guidelines, they should be in no way involved in the tendering process or the selection of contractors. In fact, they are granted practically no official role in the implementation of PMGSY whatsoever.<sup>10</sup> Funding for PMGSY comes primarily from the central government. The scheme is managed by local Programme Implementation Units (PIUs), which are under the control of State Rural Roads

---

<sup>7</sup> Existing work reports that the price bid of only one firm was evaluated in 95% of a random sample of 190 road contracts issued between 2001 and 2006 in Uttar Pradesh; i.e. only one bid submitted or all other bids were disqualified based on technical requirements (Lewis-Faupel et al., 2016). In conversations with bureaucrats involved managing PMGSY tenders, we were told that MLAs will often get technical requirements from bureaucrats and then share that information with favored contractors who then have a better chance of winning the tender (Author interviews, April 12, 2016).

<sup>8</sup> Examples include articles in “The Hindu” on April 11 2012, “The Economic Times” on March 8 2013, “The Arunachal Times” on March 6 2013, the online news-platform “oneindia” on July 31 2006, and “Zee News” on 30 August 2014. For example, the “oneindia” article reports that the former Chief Minister of Sikkim accused the current administration of “widescale corruption” in the implementation of PMGSY and “alleged that the works were awarded to relatives of Chief Minister, Ministers and MLAs of the state”.

<sup>9</sup> These authors find that the habitation population figures reported to PMGSY had been manipulated, particularly around the 1,000 and 500 population cut-offs used to target the program..

<sup>10</sup> MLAs are mentioned in the PMGSY guidelines, but only in reference to the initial planning stage. Intermediate panchayats and District panchayats were responsible for drawing up a planned “Core Network” which encompasses all future roadwork to be carried out under PMGSY. These plans were to be circulated to MPs and MLAs, whose suggestions were to be incorporated. MLAs could therefore have influenced which habitations were targeted ex-ante through official channels. However, this role is irrelevant for the timing of the construction work and assignment of road contracts, on which MLAs have no formal influence.

Development Agencies (SRRDA). These agencies are responsible for inviting tenders and awarding contracts. Given their lack of formal involvement, any systematic relationship between MLAs and the contractors working in their constituencies can therefore, in itself, be construed as evidence for an irregularity in the allocation of contracts.

### *3.2 The role of MLAs*

Is it plausible that MLAs would seek to intervene on behalf of specific contractors? While their official function is to represent their constituents in state legislative assemblies, surveyed MLAs overwhelmingly report this to be a minor part of their work (Chopra 1996). State assemblies meet rarely and according to Jensenius (2013), individual legislators have little impact on political decisions: “much more important to the MLAs are all their unofficial tasks of delivering pork, blessing occasions, and helping people out with their individual problems”. Qualitative accounts suggest that MLAs spend much of their time receiving requests from their constituents. Describing such meetings Chopra (1996) writes “constituents came to ask for favours that clearly contravened rules and laws”. MLAs often respond to requests by passing them on to ministers or high-ranking officials, but are also known to put pressure on bureaucrats by threatening them with reassignment (Iyer and Mani 2012, Bussell 2015). Existing work confirms the ability of MLAs to affect local economic outcomes: positively in the case of aligned politicians (Asher and Novosad, 2015), and negatively in the case of criminal politicians (Prakash et al., 2015). We will complement these studies by showing how MLAs can use their power to collect private rents for themselves and/or connected contractors. As stated earlier, MLAs have no formal role in the assignment of PMGSY contracts. However, their strong local power could enable them to exert undue influence in practice.

### *3.3 Surnames as a measure of interpersonal proximity in India*

To measure proximity between MLAs and contractors we construct a proxy based on politicians’ and contractors’ surnames.<sup>11</sup> Indian surnames can be an indicator of caste affiliation, religion, or geographic provenance. The strength of these associations varies regionally and across names within regions. Overall, the correlations are sufficiently strong for Indian surnames to have been

---

<sup>11</sup> Angelucci et al. (2010), and Mastrobuoni and Patacchini (2012) also uses name-based matching to study social networks.

used as identifiers of caste or religion in many empirical studies (Banerjee et al. 2014a, Hoff and Pandey 2004, Vissa 2011, Fisman et al. 2012, Field et al. 2008). This paper treats a match between the names of a politician and a contractor as a rough overall measure of proximity, without seeking to establish whether the individuals are of the same religion, caste, or (potentially) family. All of these types of connections are likely to increase the probability that a contractor would approach an MLA when bidding for a contract, and that the MLA would be receptive.

Name-based matching is an imperfect measure of proximity. Contractors may have connections to politicians without sharing a name, or equally, share a name but have no connection. Surnames that are not caste-identifiers, former honorific titles for example, are likely to dilute the accuracy of the measure. Hence, the estimates in this paper can be viewed as a lower bound for MLAs' true effect on contract allocation.

## **4. DATA**

The empirical strategy requires three kinds of data. Information on contractors and agreements is available in the administrative records of the PMGSY project, at the road level. Data on political candidates and elections are at the level of the assembly constituency. These two are linked using the population census, which allows for habitations to be matched to constituencies, as well as providing additional covariates used in the analysis.

### *4.1 PMGSY data*

The administrative records of projects sanctioned under PMGSY are publicly available in the Online Management and Monitoring System (OMMS). The dataset used for this paper contains the agreement details of 110,185 roads serving 188,394 habitations. This information includes: the date of contract signing, sanctioned cost, proposed length, proposed date of completion, name of the contracting company, and – crucially for this analysis – the name of the winning contractor. In addition to the agreement details, which precede road construction, the OMMS also contains later data on the physical progress of work, data on completed roads, and reports from subsequent quality inspections. These are used in section 7 to evaluate the effect of political interference on the efficiency and quality of road construction.

### *4.2. Assembly election data*

The Election Commission of India (ECI) publishes statistical reports on assembly elections that record each candidate's name, party, gender and vote share. Since 2003, candidates have moreover been required to submit sworn affidavits to the ECI with information on their assets, liabilities, educational attainment, and any pending criminal cases. Both the election reports and affidavits are publicly available from the ECI in pdf format. This paper draws on digitised versions of this information from four separate sources. Table A1 of the online appendix lists these sources – which cover different time periods and variables – and describes which variables from each source are used in the analysis (all these secondary sources are based on the ECI).<sup>12</sup>

Assembly elections operate on a plurality rule. While the median number of candidates per election is eight, typically only the top-2 candidates are competitive: the third placed candidates average 7% of the vote, the fourth placed candidates average 3%, the fifth 1.6% and the rest less than 1%. To estimate the RD we restrict attention to elections in which there are PMGSY contracts issued in the term before and after the election and focus on the winner and runner-up. This gives us a sample of 8,116 candidates in 4,058 elections from 2001 to 2013, covering 2,632 constituencies. In our preferred specification we estimate on the resulting sample of 8,116 candidate-terms. In a placebo test we show that the effects are not present for the contrast between the runner-up and third-placed candidate. Map 1 shows the constituencies included in the sample which cover 24 of the 28 states that existed during the timeframe under analysis.<sup>13</sup> Map 2 shows the constituencies which had at least one close election, the sub-sample for our local linear RD estimation.

[MAPS 1 and 2 here]

#### *4.3. Matching roads and electoral terms using census data*

The Population Census of India 2001 contains village-level data on demographic and socio-economic variables used as controls in the analysis. We use the Village Amenities part of the 2011 census, to identify the 'missing roads' evaluated in section 7.

---

<sup>12</sup>The matching process is complicated by discrepancies in the spelling of constituency and candidate names. These occur not only across datasets but also across time within datasets. Using different secondary sources helps us to construct a consistent data set. In a small number of cases, multiple constituencies within the same state have the same name. We drop all of these constituencies from our sample, to prevent false matches between election datasets and to avoid the risk of assigning roads to the wrong constituency.

<sup>13</sup> Goa, Meghalaya, Nagaland and Sikkim are not part of our sample.

The 2001 census is also the source for habitation-level data, which is collected by the PMGSY in order to determine the prioritisation of roads. This includes information on the size of the population (the project guidelines stipulate that habitations above certain population thresholds are to be prioritised), whether or not it was connected to a road in 2001, and if so, whether this road provided all-weather access. Moreover, it reports the MLA constituency in which each habitation was situated in 2001.

Using this information, it is possible to match PMGSY roads (at the habitation-level) to the assembly election data described in the previous sub-section. However, changes in the delimitation of MLA constituencies – which took effect in mid-2008 – led to changes in boundaries, the abolition of some constituencies, and the creation of new ones. For roads built in electoral terms after the new delimitation we use the coordinates of habitations and match these to GIS data on constituency boundaries.

While the census data allows for spatial matching of roads and constituencies, it is also necessary to match them in time. Road contracts are allocated to electoral terms based on the date of the agreement, as recorded in the PMGSY data. In order to precisely assign road contracts, it is necessary to set an exact date that marks the end of one term and the beginning of the next. We define this as the date on which the results of an election are announced.<sup>14</sup>

#### *4.4 Matching politicians and contractors using surnames*

In the electoral terms that preceded and followed the elections in the sample, 88,020 road agreements were signed. For each political candidate, we assess whether they share a surname with the contractors who received projects in their constituency in the term after the election. For every politician-contractor pair, we exclude all names except for each individual's final name and then look for matches among these surnames. The results are, however, robust to broader definitions of matches.<sup>15</sup> To account for different spellings of the same name, we implement a fuzzy matching algorithm optimised for Hindi names.<sup>16</sup>

---

<sup>14</sup> These dates were collected from the website [www.electionsinindia.com](http://www.electionsinindia.com) (accessed in 2015).

<sup>15</sup> Naming conventions differ across India; it is common for Indians to have multiple surnames and the same name can appear in different positions within the list of names. This is also true of caste identifiers. The results are robust to considering all matches among individuals' names (excluding their first name) or only matches based only on the last two names.

<sup>16</sup> All results are robust to considering exact matches only.

Matches are aggregated at the electoral term level as follows. The variable  $match_{nijt}$  takes the value of 1 if the contractor for a road agreement  $n$ , signed in constituency  $j$  in term  $t$ , shares a name with candidate  $i$ , and 0 otherwise. This variable is determined for the  $N$  road agreements signed in the constituency during an electoral term.  $share_{ijt}$  is defined as the share of contracts in term  $t$  allocated to contractors who share a candidate's name.  $share_{ijt-1}$  provides the equivalent share for contracts in the term prior to the election in which a candidate took part.

$$share_{ijt} = \frac{\sum_{n=1}^n match_{nijt}}{N_{jt}} \quad share_{ijt-1} = \frac{\sum_{n=1}^n match_{nijt-1}}{N_{jt-1}}$$

The dependent variable in the main regressions is the difference between these two:

$$\Delta share_{ijt} = share_{ijt} - share_{ijt-1},$$

which we calculate for all candidates  $i \in \{winner, runner-up\}$ .

A complication arises in elections where winning and losing candidates have the same surname. It is not possible to estimate the effect of winning an election in this situation, as candidates who lost will see their proximity to contractors evolve in parallel to that of the elected politicians. In the main regressions, we therefore exclude candidates from elections where this issue arises.

#### 4.5 Descriptive statistics

Table 1 reports descriptive statistics for the sample of candidates used in the main regressions. For the average term in the sample, the number of road contracts signed is 28. The average value of  $share_{ijt-1}$  – which can be construed as a baseline measure of the frequency of surname-matches – is 4%. There is however, significant geographic variation in the frequency of matches, ranging from a mean of 0% in Mizoram to a mean of 13% in Andhra Pradesh (Map 3 shows this variation at the constituency-level).<sup>17</sup> However, these means do not distinguish between winning and losing candidates – the variation exploited in the empirical strategy below.

[TABLE 1 ABOUT HERE]

---

<sup>17</sup> It is likely that these baseline frequencies lead to heterogeneous treatment effects. In states or constituencies, where the distribution of names is such that matches are relatively rare, a politician who is elected may not have many potential contractors of the same name to allocate roads to.

## 5. EMPIRICAL STRATEGY

A natural control group for elected politicians are those who aspire to the same office. If being an MLA is associated with the power to intervene in the allocation of roads in one's constituency, one would expect the share of contractors with the same name as a winning candidate to be higher than the corresponding share for losing candidates.

Taking the first difference of  $share_{ijt}$  should remove unobservable, time-invariant characteristics of an individual candidate that may be correlated with the number of matches with contractors. In our context, this is a way of controlling for specificities that individual names may have within certain constituencies. Some candidates' names will be more common than others. Some may be more prevalent among certain professions (e.g. contractors) for historical reasons. Under the assumption that winning and losing candidates had a common trend in their share of matches with contractors, a simple DiD approach would be sufficient for identification. However, given that winners are likely to be systematically different from losing candidates in many respects, it is possible that they may face divergent trends in  $share_{ijt}$  that are not determined by election outcomes. This suggests the use of a regression discontinuity (RD) design.

In order to identify whether there is a causal relationship between the election of politicians and the allocation of road contracts in their constituencies, we exploit the fact that in close elections, the assignment of victory can be considered conditionally independent of subsequent contracting patterns. The underlying assumption is that candidates who won an election by a very small margin are comparable to those who narrowly lost (Lee, 2008). We evaluate whether this assumption holds in our sample by running balance checks on observable characteristics (see below). In order to determine how close elections were, we define the variable  $margin_{ijt}$ :

$$margin_{ijt} = \begin{cases} vote\ share(winner)_{jt} - vote\ share(runnerup)_{jt} & \text{if } winner_{ijt} = 1 \\ vote\ share(runnerup)_{jt} - vote\ share(winner)_{jt} & \text{if } winner_{ijt} = 0 \end{cases}$$

We estimate equation (1) in a non-parametric RD for a range of bandwidths  $\mu$ , controlling for the assignment variable  $margin_{ijt}$  and its interaction with  $winner_{ijt}$  to allow for a different relationship between  $\Delta share_{ijt}$  and  $margin_{ijt}$  among winning and losing candidates:

$$\Delta share_{ijt} = \alpha + \beta winner_{ijt} + \delta margin_{ijt} + \rho winner * margin_{ijt} + \varepsilon_{ijt} \quad \forall i \text{ where } margin_{ijt} \in [-\mu, \mu] \text{ and } i \in \{winner, runner-up\} \quad (1)$$



In order to improve the efficiency of the estimates, we introduce constituency-level controls, individual-level controls, state fixed-effects, and year fixed-effects in most specifications although these are not required for identification.<sup>18</sup> Because we have the top-two candidates in each election we cluster standard errors at the election level.<sup>19</sup>

## 6. REALLOCATION RESULT

### 6.1 Randomisation test

Our identification strategy is based on the premise that restricting the sample to close elections ensures that the treatment and control groups are comparable. Table 2 presents the results of a randomization test for the optimal bandwidth of 6.2% (derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman 2011). None of the MLA characteristics display a discontinuity when the vote margin exceeds one. Panel B shows that the common trends assumption is likely to hold in that we see no differences in the share of matches in years leading up to close elections.

[TABLES 2 AND 3 ABOUT HERE]

### 6.2 Main results

The results of local linear regression RD estimation are presented in Figure 1 and Table 3. For each bandwidth there are two columns in the table. The first corresponds to the basic RD in equation (1). The second adds state fixed effects, year fixed effects and additional controls. These include whether or not a constituency is reserved for candidates from scheduled castes (SC) or scheduled tribes (ST), characteristics of the PMGSY roads built in the constituency prior to the election, and candidate-level controls. The latter set of variables includes a candidate's vote share, their age, gender, and whether they were an incumbent or a former runner-up.

For the 6.2% bandwidth, the effect of winning an election on the change in  $share_{ijt}$  is consistently positive and significant. The coefficient is around 0.024 in our preferred specification

---

<sup>18</sup> Legislative assembly terms are not synchronised across Indian states. In each year in our sample window, there were elections in multiple states.

<sup>19</sup> The main results are robust to clustering standard errors at the state-year level to account for within-state-political-season correlations in the errors.

including fixed effects and the full set of controls (column 4). Relative to the baseline, pre-election level of matches, the latter estimate implies that the effect of a candidate coming to power is a 63% increase in the share of roads allocated to contractors who share their surname.<sup>20</sup>

The results of non-parametric RD estimations can be sensitive to the choice of bandwidth., and there is trade-off between bias and efficiency inherent in this choice (Lee and Lemieux 2010). Reassuringly, the results are consistent across a wide range of bandwidths. Figure 2 plots the coefficient on  $winner_{ijt}$  for the main specification for different bandwidths. As the samples get smaller the estimates are less precise but the coefficient is relatively stable for all but very small bandwidths (less than 1%).

Relative to the total number of roads – most of which are allocated to contractors whose name does not match the MLA’s – the absolute value of the coefficient implies a small effect. Yet as explained in section 3.3, these estimates can be considered a lower bound on MLAs’ true intervention in PMGSY contract allocation. If the results are interpreted as evidence of improper political involvement in the assignment of roads, it raises the question whether this improper involvement *only* occurs on behalf of individuals with the same surname. In this sense the sign and significance of the coefficient might be seen as more important than the magnitude. Secondly, given the scale of PMGSY, even a relatively small fraction can translate into what can be considered a sizeable number of affected roads and substantial financial expenditure. This is illustrated by the following, back-of-the-envelope calculation. In our dataset (including the first electoral term), 4,127 road projects were allocated to contractors sharing a name with the MLA. The total sanctioned cost of these projects was 56 billion INR, or around 1.2 billion USD.<sup>21</sup> Applying our preferred RD estimate (6.2% bandwidth) to the full sample, would imply that MLAs had intervened in the allocation of roughly 1,600 road contracts worth around 470 million USD.<sup>22</sup> Of course, these estimates rely on an extrapolation from a LATE. Still, they serve to illustrate the economic significance of even proportionately small misallocations in PMGSY contracts.

---

<sup>20</sup> In appendix table A2 we report results using the level of the share of same name contractors (rather than the difference) as the main outcome. In these results, the key coefficient is slightly smaller and less consistent in magnitude across bandwidths, but it remains significant in all but one specification.

<sup>21</sup> Applying the average exchange rate over the period (December 2000 to December 2013): 1 INR=0.021 USD.

<sup>22</sup> The estimated impact in the RD with a full set of controls on a 6.2% bandwidth is a 63% increase. This implies that 38.6% of roads allocated to contractors with the same name as the politicians would otherwise have gone to another contractor.

The coefficients in Table 3 are based on a sample that includes practically all Indian states, but there are reasons to expect significant heterogeneity in the results across regions. Our name-based approach will provide a more accurate measure of proximity in areas where there is a strong association between castes and surnames. This is more likely to be the case in northern states than southern ones and in constituencies not reserved for members of Scheduled Castes or Scheduled Tribes. Figure 3 shows that restricting the estimation to these areas results in a higher coefficient, although the sample is underpowered to test heterogeneous effects. By contrast, when we focus only on Tamil Nadu, a state whose naming conventions imply that surnames will not provide an indicator of proximity, the coefficient is statistically insignificant and very close to zero. Heterogeneity could also result from differences in the underlying level of corruption across regions. Figure 3 shows that relative to the main sample, the effect is roughly 28% larger in the so-called BIMAROU states that are widely reputed to be more corrupt.<sup>23</sup> However, given that these are northern states with a relatively strong association between names and caste affiliation, we are unable to attribute this to systemic corruption.

A second possible source of heterogeneity is the party affiliation of the MLA. Asher and Novosad (2015) find that MLAs aligned with the party in power at the state-level appear to have greater control over the bureaucracy. In appendix table A3 we evaluate whether this applies to their involvement in PMGSY contract allocation, with inconclusive results. The differential effect for aligned MLAs is consistently positive but not statistically significant.<sup>24</sup>

The results of this section lend support to qualitative accounts on favouritism in the allocation of PMGSY contracts. Only recently, BJP leader Munna Singh Chauhan accused the Uttarkhand State Government of such misallocations:<sup>25</sup>

“There is a huge scam in tender allotment in Pradhan Mantri Gram Sadak Yojana (PMGSY) in Bahuguna government. Of a total of 113 mega road construction projects, 75 contracts were awarded to chosen ones close to the echelons of power on a single bid basis. [...]

---

<sup>23</sup> The definition of BIMAROU is loose. We use the broadest set which includes Bihar, Madhya Pradesh, Rajasthan, Orissa, and Uttar Pradesh, as well as new states created on their historical territory: Chhattisgarh, Jharkhand, and Uttarkhand.

<sup>24</sup> The effect for unaligned MLAs is also generally insignificant and the two coefficients are jointly significant throughout.

<sup>25</sup> See footnote 6 for references to similar newspaper articles.

Coincidentally, one of the contractors awarded the project is also the brother-in-law of state rural development minister Pritam Singh,” (Quoted in Zee News, 30 August 2013).

Our analysis suggests that episodes of suspected favouritism in particular states, like the one quoted above, match a wider pattern of corruption that shows up in our sample covering the whole of India.

## 6.2 Validity of the RD approach

The RD design requires that no variables other than the dependent variable exhibit discontinuities at the cut-off. The randomization test in Table 2 provided the first evidence that observable characteristics are comparable on either side of the cut-off.

Close elections can only be considered to provide quasi-random treatment assignment when the probability density function of candidates’ vote shares is continuous (Lee 2008). This will not be the case if candidates are able to strategically manipulate their vote share.<sup>26</sup> The standard test for strategic manipulation of the running variable in a RD design was formulated by McCrary (2008). Applying the McCrary test to the assignment variable in this analysis ( $margin_{ijt}$ ), would not make sense because the density is continuous by construction. For every winner with a positive  $margin_{ijt}$ , there is a runner-up with the equivalent negative value of  $margin_{ijt}$ . We therefore test for manipulation in the vote share based on an alternative variable: the margin of victory/defeat for the candidate in the constituency with the higher value of  $share_{ijt-1}$ . The McCrary test does not reject the continuity of this variable at the threshold. Figure A1 in the online appendix presents a graphical depiction of the test.

The online appendix provides further robustness checks, including results of a parametric RD regression estimated on the full sample (in Table A5), as well as the main results in levels instead of first differences (Table A2). The main result is robust to these alternative specifications. To evaluate whether our results can be interpreted as the causal impact of gaining public office, as opposed to the information revealed by performing well in an election, we conduct a placebo test comparing runners-up to third-placed candidates (Table A6). If shifts in allocation somehow

---

<sup>26</sup> Using data on close US house races, Caughey and Sekhon (2011) provide evidence of such strategic sorting. Eggers et al. (2015) examine over 40,000 close elections from a range of countries (including India) and find no other country that exhibits sorting.

reflect individuals' increased status following a strong electoral performance, rather than their official position, second-placed candidates might experience gains relative to third-placed candidates. We find no such effect; the coefficient on coming second is close to zero across all specifications.

## 7. SOCIAL COSTS OF MISALLOCATION

Theoretical work has contended that corruption could be socially beneficial (Leff 1964). In the case of political connections, proximity may be associated with better information ex-ante or greater sanctioning power ex-post, and is therefore desirable in contexts of adverse selection or moral hazard. Distinguishing between outright corruption and this efficiency motive is a challenge that is faced by many empirical studies on political connections. We analyse PMGSY projects at the road level, in order to evaluate how MLA's interventions affect the cost, timeliness, and quality of road construction, as well their impact on the likelihood that construction actually takes place.

Are roads built by contractors who are connected to politicians better or worse than other roads? We again employ an RD-approach that exploits close elections to identify the impact of political interference on the efficiency and quality of road construction.<sup>27</sup> We drop all roads from the sample that were not built either by a contractor who shares a name with the current MLA, or by a contractor who shares a name with the runner-up in the most recent election. Once the sample is restricted to close elections, the latter set of roads can be considered a more appropriate control group as it will be similar to the 'treated' roads.<sup>28</sup> Once again we control for the vote shares of winning and losing candidates. The equation for this non-parametric RD is given by:

$$\text{Road Characteristic}_{nsy} = \alpha + \beta * \text{MLAsamename}_{nsy} + \delta \text{margin}_{ijt} + \rho \text{winner} * \text{margin}_{ijt} + \gamma X_{nsy} + \theta_s + \vartheta_y + \varepsilon_{nsy}, \text{margin}_{ijt} \in [-\mu, \mu] \quad (3)$$

---

<sup>27</sup> One way to approach this question empirically would be to run regressions of road characteristics on a dummy variable that takes the value of one if the MLA and the contractor for road have the same name. However, this approach would fail to control for two important sources of unobserved variation. Firstly, contractors who have the same name as politicians may have systematically different characteristics from other contractors. Secondly, the locations where contractors of the same name as the MLA operate could be systematically different from other areas targeted by PMGSY.

<sup>28</sup> Assuming as above, that the names of politicians who just win elections are not systematically different from the names of candidates who just lost.

The first outcome we consider is the cost of road construction. If rent-seeking politicians are putting pressure on bureaucrats to reject the lowest bidder in favour of their preferred contractor, we would expect to see a rise in costs. Table 4 shows that roads built by contractors who share a name with an elected official are more expensive (per kilometre). This result is significant for the bandwidths we consider with the coefficient rising as the bandwidth declines.

[TABLE 4 ABOUT HERE]

It should be noted that, while the RD-design is likely to be an improvement on a naïve OLS approach, it may still be insufficient to identify a causal effect. To the extent that politicians only intervene on behalf of their network for some roads, and this selective intervention is not random, the ex-ante characteristics of the roads in the treatment group may differ from those in the control group. For example, politicians might try to ensure that more difficult projects are allocated to contractors from their network whom they trust. Given that the road-level outcomes we observe were predominantly determined ex-post – at the time of the contract or during construction – this possibility cannot easily be evaluated. We control for observable variables that might affect the cost of a road project and a politician’s desire to intervene in its allocation: characteristics of the terrain (altitude and ruggedness) and whether the project involved the construction of a bridge). The resulting estimates allow us to measure the bias in observable characteristics. Table A4 of the appendix, shows that the coefficient remains unchanged when these additional controls are added.<sup>29</sup> As a result, in order to fully explain the estimates of Table 4, the bias in unobservable characteristics would have to very large relative to the bias in observable characteristics (Altonji et al., 2005).

A rise in costs might not be socially detrimental if it were offset by improved quality. Table 5 therefore analyses five additional measures of quality using the same RD approach: (i) the number of days between the completion date specified in the contract and the actual date of completion; (ii) the ratio between the actual cost of the project and the cost sanctioned in the agreement; (iii) a dummy variable for whether a road was deemed “unsatisfactory” or “in need of improvement” in either the latest state quality inspection or the latest national quality inspection<sup>30</sup>; (iv) whether a

---

<sup>29</sup> In order to ensure comparability, we restrict the sample to roads for which we have information on altitude, ruggedness, and bridges in all the regressions of Table A4.

<sup>30</sup>The quality data available on the OMMAS has some shortcomings for the purpose of this analysis. Data is available on national and state quality inspections, and a single road may have multiple inspections in each category. However, only the grade assigned in the latest inspection is provided (for each category). The data therefore do not allow us to

road is “missing” by the most conservative definition: none of the villages located on the planned road had either a black-topped road, a water bound macadam road, or a gravel road, according to 2011 census data collected after that road was officially listed as completed<sup>31</sup>; and (v) a second measure of “missing roads” that applies PMGSY’s own criteria: whether all the villages located on the planned road had all-weather road access according to the 2011 census<sup>32</sup>.

For delays and cost discrepancies we find no significant difference between roads constructed by contractors whose name matches the MLA’s and those whose name matches the runner-up. However, roads allocated to connected contractors were more likely to fail subsequent quality inspections. At the optimal bandwidth, this result is only weakly significant (at the 90% level).

Finally, we find that preferential allocation increases the likelihood of missing roads. This result is robust to different measures of missing roads. For our most conservative definition, the coefficient for the optimal bandwidth implies that the probability of a road not being constructed increases by 145% when the politician and contractor share a name. Applying this estimate to our whole sample in a back-of-the-envelope calculation, suggests that preferential allocation accounts for 73 additional missing roads that would have served around 125,000 people<sup>33</sup>. When we identify missing roads based on targeted villages’ all-weather road access, we find that the probability increases by 170% for connected contractors. The same back-of-the-envelope calculation implies that preferential allocation resulted in an additional 672 missing all-weather roads, with around 1.16 million people affected<sup>34</sup>.

---

distinguish between roads that were satisfactory at the outset, and roads that initially did not pass inspection but were improved prior to subsequent inspections. Moreover, only a fraction of the roads in our sample appear in the quality data, and many of these only had one of the two inspection types (national or state). Pooling the two inspections is not ideal, but it provides the best available measure of initial road quality.

<sup>31</sup> By this measure, around 2.5% of roads listed as completed prior to the census are missing. This is likely to be an underestimate. Given that some existing roads (gravel roads in particular) may not have met PMGSY’s goal of all-weather access, villages that have such roads in the census may still never have received a PMGSY road.

<sup>32</sup> By this measure, around 26% of roads listed as completed prior to the census are missing. There are two reasons why a road could appear as missing, both of which are indicative of corruption. Firstly, roads may be listed as completed without ever being built. Secondly, roads could be built with sub-standard materials leading to complete or partial deterioration by the time of the 2011 census.

<sup>33</sup> 4,127 roads in our sample were built by connected contractors. Of these 3% are deemed to be missing. Our estimates imply that the share of these missing roads due to preferential allocation is 59% ( $1-1/(1+1.45)$ ). This yields an estimate of 73 roads. The average road in our sample serves villages with a total of 1726 inhabitants, giving an estimate of 126,343 people left unconnected.

<sup>34</sup> 4,127 roads in our sample were built by connected contractors. Of these 26% are missing all-weather road access. Our estimates imply that the share of these missing roads due to preferential allocation is 62%, or 672 roads. Multiplying this by the average number of inhabitants on a road, gives an estimate of 1,160,752 people affected.

To the extent that inferences can be drawn from this incomplete set of indicators, preferential allocation appears to reflect costly corruption with no mitigating improvements in the efficiency of road construction. Indeed, we find suggestive evidence that political intervention leads to roads that are not only more expensive, but also more likely to be either of poor quality or never to have been built at all.

[TABLE 5 ABOUT HERE]

## 8. ELECTORAL CYCLES IN CORRUPTION

We implicitly assume that kinship ties to politicians are relevant connections in the structure of local political corruption in India, and our results appear to validate this assumption. But why should patronage be targeted along caste or familial lines? The literature offers two main explanations: vote-buying and particularised trust. In this section we attempt to shed light on which is more applicable to corruption in PMGSY

[TABLE 6 ABOUT HERE]

If road contracts are awarded in exchange for political contributions or political support, one would expect the bias towards connected contractors to increase in election periods. To test for this we construct more disaggregated measures of proximity: the share of contractors with a candidate's name in the first 12 months after an election (*start of term* $_{ijt}$ ), the equivalent share for the last 12 months before the subsequent election (*end of term* $_{ijt}$ ), and finally the share for the intermediate, mid-term, period. The first three columns of Table 6 shows the results of applying our main estimation approach to this disaggregated sample and interacting dummies for *start of term* $_{ijt}$  and *end of term* $_{ijt}$  with *winner* $_{ijt}$ . The overall effect of winning the election is comparable to the term-level results, and we find no differential effects in election years.

Although the bias towards connected contractors does not increase in election periods, there could be different patterns for the within-term variation on the cost margin. Politicians might need to extract rents, buy support, or reward supporters were higher before or after elections. Including the interactions between *MLAsamename* $_{njt}$  and *start of term* $_{ijt}$  and *end of term* $_{ijt}$  in the road-level regressions, we again find no evidence of a political cycle in which election periods see



increased corruption (columns 4-6 of Table 6). The observed negative effect for both the start and end of term is more consistent with increased scrutiny in the run-up to elections acting as a deterrent to corruption.

Changes to the delimitation of parliamentary constituencies allow for an additional test of the vote-buying hypothesis. The changes proposed by the delimitation commission of 2002 were approved in February 2008. Subsequent assembly elections, starting with Karnataka in May 2008, were carried out under the new delimitation. After the reform had been announced and approved, the majority of MLAs elected under the old delimitation continued to hold office for several years until the next election. In constituencies where the boundaries were redrawn, this meant that only some areas would remain part of the constituency at the next election, while others would be of no consequence to the MLA's chances of re-election. We identify such areas with a dummy variable *politically irrelevant<sub>ijt</sub>* and also disaggregate temporally, splitting the applicable electoral term into the period before the announcement, and the period between February 2008 and the next election (the variable *post announcement<sub>ijt</sub>* denotes the latter). Given that the boundaries were defined by an independent commission following objective pre-set guidelines, the reform could provide plausibly exogenous variation in the incentive for vote-buying.<sup>35</sup> The final column of Table 6 presents the results of our main specification for the disaggregated sample and interaction terms. The coefficient of interest is the triple interaction term:

$$winner_{ijt} * politically\ irrelevant_{ijt} * post\ announcement_{ijt}$$

A negative and significant coefficient would suggest that political corruption is weaker in areas where politicians have no incentive to buy votes. The coefficient on the triple interaction is however, positive and statistically insignificant. Hence, we find no evidence of strategic vote-buying. This result is also consistent with recent work by Chhibber and Jensenius (2016), who use the same delimitation experiment and find that MLAs from “ethnic” or “well-organized” parties tend to target existing loyalists rather than the electorally valuable voters who joined their constituencies post-delimitation.

---

<sup>35</sup> According to the Electoral Commission of India's Guidelines and Methodology for Delimitation, “the delimitation of the constituencies in a district shall be done starting from North to North-West and then proceeding in a zig-zag manner to end at the Southern side.” Constituencies were to have equal populations, as far as possible, with maximum deviations of 10% from the State average, based on the 2001 Census.

In the absence of clear evidence for vote buying, it is possible that corruption arises within kinship networks because these provide the “particularised trust” needed to engage in risky collusive behaviour (Tonoyan, 2003). While we are unable to test this explanation explicitly, it fits the context of PMGSY in that the involvement of the central government provides a minimum level monitoring. We find no evidence that preferential allocation affects the performance markers that are most easily observed in the central administrative data: over-runs and delays, consistent with the idea that contractors face a trade-off between potential rents and the cost of detection.

PMGSY is subject to a central monitoring system but, by explicitly limiting political involvement, it foregoes a mechanism that could provide local accountability. If voters held their MLAs responsible for the services delivered under PMGSY, the latter would have an incentive to limit corruption. By contrast, a scheme in which local politicians have no formal role but over which they still retain influence through informal channels, can be seen as an ideal vehicle for rent-seeking. One interpretation of our results is that the design of PMGSY restricted political accountability rather than political control.

## **9. POLICY RESPONSES TO CORRUPTION**

The identification of corrupt interventions in PMGSY allows us to revisit the evidence of Lewis-Faupel et al. (2016) who find that e-procurement improved the quality of PMGSY road construction. They identify the entry of higher quality contractors from outside regions as a key mechanism behind the quality improvement, a development which could be consistent with a decline in preferential allocation. In fact, an NRRDA official described political influence over contract allocation in some regions, as one of the motivations for making e-procurement mandatory for all states (Author interview, January 18, 2016). Interacting our main effect with the state-specific introduction of e-procurement, we evaluate whether preferential allocation declined under the new policy (Table 7). The interaction term is never significant and the point estimate is close to zero. In short, we find no evidence that e-procurement has helped to prevent this specific form of corruption in PMGSY.

[TABLE 7 ABOUT HERE]

## 10. CONCLUSION

This paper provides direct empirical evidence that local politicians in India abuse their power to benefit members of their own network. We exploit the variation in political leadership due to the electoral cycle, to identify systematic distortions in the allocation of contracts for a major rural road construction programme (PMGSY). By matching contractors' and political candidates' surnames, we generate a measure of proximity which evolves as the pool of contractors changes. A regression discontinuity design based on close elections, suggests that the causal impact of a politician coming to power is a 63% increase in the share of roads allocated to contractors who share their surname. This result withstands a series of alternative specifications and robustness checks. Further regression discontinuity estimates at the road level, indicate that political interference in the allocation of roads raises the cost of construction, without providing any offsetting benefits in terms of efficiency or quality. Corruption is therefore welfare-reducing in this context.

A distinguishing feature of our analysis, is that we identify the effect of political connections to state-level legislators who have no official involvement in the road construction programme. Our results therefore not only indicate preferential treatment of the politically connected, they also provide indirect evidence that local politicians' power over purportedly neutral bureaucrats is sufficient to coerce them into corruption. From a policy perspective, these findings indicate that more could be done to insulate the officials implementing government programmes at the local level, including those involved in PMGSY.

While this paper is primarily about the measurement of corruption, its findings have significance beyond the potential number of misallocated roads or the amount of misdirected money. If corrupt arrangements were made based on random matching between individuals, the empirical strategy would have revealed nothing. Our results provide further evidence for the role of networks in facilitating corruption and point towards theories in which kinship networks facilitate corruption through trust or the ability to impose social sanctions. The irony is, that the setting for the analysis – PMGSY – is conceptually a profoundly inclusive programme, facilitating the integration of over 100 million people into the Indian economy (Aggarwal 2015). This paper

suggests that allowing them to compete equally for jobs, permits, licenses, or government procurement contracts, may require building more than roads.

## REFERENCES

- Acemoglu, Daron, Tarek Hassan, and Ahmed Tahoun. "The Power of the Streets: Evidence from Egypt's Arab Spring." Mimeo MIT (2014).
- Aggarwal, Shilpa. "Do rural roads create pathways out of poverty? Evidence from India." *University of California, Santa Cruz, unpublished* (2015).
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of political economy* 113, no. 1 (2005): 151-184.
- Angelucci, Manuela, Giacomo De Giorgi, Marcos A. Rangel, and Imran Rasul. "Family networks and school enrolment: Evidence from a randomized social experiment." *Journal of Public Economics* 94, no. 3 (2010): 197-221.
- Asher, Sam, and Paul Novosad. "Politics and Local Economic Growth: Evidence from India." *Cambridge, MA: Harvard University, unpublished* (2015).
- Asher, Sam, and Paul Novosad. "Market Access and Structural Transformation: Evidence from Rural Roads in India." Manuscript: Department of Economics, University of Oxford (2016).
- Banerjee, Abhijit, Sendhil Mullainathan, and Rema Hanna. *Corruption*. No. w17968. National Bureau of Economic Research, (2012).
- Banerjee, Abhijit, Donald P. Green, Jeffery McManus, and Rohini Pande. "Are poor voters indifferent to whether elected leaders are criminal or corrupt? A vignette experiment in rural India." *Political Communication* 31, no. 3 (2014a): 391-407.
- Banerjee, Abhijit, Esther Duflo, Clement Imbert, Santosh Matthew, and Rohini Pande. "Can e-governance reduce capture of public programs? Experimental evidence from a financial reform of India's employment guarantee." Mimeo. (2014b).
- Becker, Gary S., and George J. Stigler. "Law enforcement, malfeasance, and compensation of enforcers." *The Journal of Legal Studies* (1974): 1-18.
- Bertrand, Marianne, Simeon Djankov, Rema Hanna, and Sendhil Mullainathan. "Obtaining a driver's license in India: an experimental approach to studying corruption." *The Quarterly Journal of Economics* (2007): 1639-1676.
- Bussell, Jennifer. "Clients or Constituents? Distribution Between the Votes in India." mimeo, University of California Berkeley (2015)

- Burgess, Robin, Remi Jedwab, Edward Miguel, and Ameet Morjaria. "The value of democracy: evidence from road building in Kenya." *The American Economic Review* 105, no. 6 (2015): 1817-1851.
- Caughey, Devin, and Jasjeet S. Sekhon. "Elections and the regression discontinuity design: Lessons from close us house races, 1942–2008." *Political Analysis* 19, no. 4 (2011): 385-408.
- Chandra, Kanchan. "Why Ethnic Parties Succeed." *Patronage and ethnic head counts in India* (2004).
- Chhibber, Pradeep, and Francesca R. Jensenius. 2016. "Privileging one's own? Voting patterns and politicized spending in India." *Mimeo*.
- Chopra, Vir K. *Marginal players in marginal assemblies: The Indian MLA*. Orient Longman, 1996.
- Do, Quoc-Anh, Yen Teik Lee, and Bang Dang Nguyen. "Political connections and firm value: evidence from the regression discontinuity design of close gubernatorial elections." *Available at SSRN 2190372* (2013).
- Eggers, Andrew C. and Jens Hainmueller, "MPs For Sale: Returns to Office in Post-War British Politics", *American Political Science Review*, Vol. 103, No. 4, (2009): pp. 1-21.
- Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder. "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races." *American Journal of Political Science* 59, no. 1 (2015): 259-274.
- Ferraz, Claudio, and Federico Finan. "Exposing corrupt politicians: the effects of Brazilians publicly released audits on electoral outcomes." *Quarterly Journal of Economics* 123, no. 3 (2008): 703.
- Ferraz, Claudio, and Frederico Finan. "Electoral Accountability and Corruption: Evidence from the Audits of Local Government." *American Economic Review* 101 (2011): 1274.
- Field, Erica, Matthew Levinson, Rohini Pande, and Sujata Visaria. "Segregation, Rent Control, and Riots: The Economics of Religious Conflict in an Indian City." *The American Economic Review* (2008): 505-510.
- Fisman, Raymond. "Estimating the value of political connections." *American Economic Review* (2001): 1095-1102.
- Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. "Cultural proximity and loan outcomes". NBER Working Paper, No. 18096., 2012.
- Fisman, Raymond, Florian Schulz, and Vikrant Vig. "Private Returns to Public Office". *Journal of Political Economy* (2014).
- Golden, Miriam A., and Lucio Picci. "Proposal for a new measure of corruption, illustrated with Italian data." *Economics & Politics* 17, no. 1 (2005): 37-75.
- Hoff, Karla, and Priyanka Pandey. "Belief System and Durable Inequalities: An Experimental Investigation of Indian Caste." World Bank, Policy Research Paper: No. 3351, 2004.
- Horowitz, Donald L. *Ethnic groups in conflict*. Univ of California Press, 1985.

- Huntington, S. "Modernisation and Corruption 'in Political Order in Changing Societies, Yale University Press, New Haven." (1968).
- Imbens, Guido, and Karthik Kalyanaraman. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies* (2011).
- Iyer, Lakshmi, and Anandi Mani. "Traveling agents: political change and bureaucratic turnover in India." *Review of Economics and Statistics* 94, no. 3 (2012): 723-739.
- Jensenius, Francesca Refsum. "Power, performance and bias: Evaluating the electoral quotas for scheduled castes in India." (2013). PhD Thesis, University of California Berkeley
- Krueger, Anne O. "The political economy of the rent-seeking society." *The American economic review* (1974): 291-303.
- Lambsdorff, Johann Graf. "Corruption and rent-seeking." *Public choice* 113, no. 1-2 (2002): 97-125.
- Larreguy, Horacio A., John Marshall, and James M. Snyder Jr. *Revealing Malfeasance: How Local Media Facilitates Electoral Sanctioning of Mayors in Mexico*. No. w20697. National Bureau of Economic Research, 2014.
- Lee, David S. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142, no. 2 (2008): 675-697.
- Lee, David S., and Thomas Lemieux. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2010): 281-355.
- Leff, Nathaniel H. "Economic development through bureaucratic corruption." *American behavioral scientist* 8, no. 3 (1964): 8-14.
- Lewis-Faupel, Sean, Yusuf Neggers, Benjamin A. Olken, and Rohini Pande. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia," *American Economic Journal: Economic Policy*, (2016).
- Lui, Francis T. "An equilibrium queuing model of bribery." *The journal of political economy* (1985): 760-781.
- Mastrobuoni, Giovanni, and Eleonora Patacchini. "Organized crime networks: An application of network analysis techniques to the American mafia." *Review of Network Economics* 11, no. 3 (2012).
- McCrary, Justin. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142, no. 2 (2008): 698-714.
- Melo, Marcus André, Carlos Pereira, and Carlos Mauricio Figueiredo. "Political and institutional checks on corruption: Explaining the performance of Brazilian audit institutions." *Comparative Political Studies* (2009).
- Mironov, Maxim and Ekaterina Zhuravskaya. 2016. "Corruption in Procurement and the Political Cycle in Tunneling: Evidence from Financial Transactions Data." *American Economic Journal: Economic Policy*, 8(2): 287-321.
- NRRDA, "Standard Bidding Document for PMGSY." (2015).

- Olken, Benjamin A., and Patrick Barron. "The Simple Economics of Extortion: Evidence from Trucking in Aceh." *Journal of Political Economy* 117 (3) (2009).
- Olken, Benjamin A. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia" *Journal of Political Economy* 115 (2) (2007): 200-249
- Olken, Benjamin A. "Corruption perceptions vs. corruption reality." *Journal of Public economics* 93, no. 7 (2009): 950-964.
- Kramon, Eric and Daniel Posner. 2016. "Ethnic Favoritism in Education in Kenya.", *Quarterly Journal of Political Science*, 11(1).
- Prakash, Nishith, Marc Rockmore, and Yogesh Uppal. 2015. "Do criminally accused politicians affect economic outcomes? Evidence from India", *Mimeo*.
- Reinikka, Ritva, and Jakob Svensson. "Local capture: evidence from a central government transfer program in Uganda." *The Quarterly Journal of Economics* (2004): 679-705.
- Rose-Ackerman, Susan. "The economics of corruption." *Journal of public economics* 4, no. 2 (1975): 187-203.
- Shleifer, Andrei, and Robert W. Vishny. "Corruption." *The Quarterly Journal of Economics* 108, no. 3 (1993): 599-617
- Tonoyan, Vartuhí. "Corruption and Entrepreneurship: Impact of Trust and Civic Society; a Cross Cultural Comparison of 21 Countries; Paper Presented at the International Conference on " Trust and Entrepreneurial Behavior in East and West European Economies: Concepts, Developments and Comparative Aspects", University of Bremen 2003." ifm, 2003.
- World Bank, "PMGSY Rural Roads Project : P124639 - Implementation Status Results Report" (2014).
- Truex, R. "The returns to office in a 'rubber stamp' parliament", *American Political Science Review*, 108, no.2 (2014): 235-251.
- Vissa, Balagopal. "A matching theory of entrepreneurs' tie formation intentions and initiation of economic exchange." *Academy of Management Journal* 54, no. 1 (2011): 137-158.

## APPENDIX: MAIN TABLES

**Table 1A: Descriptive Statistics (Candidate/Constituency)**

Variable	Observations	Mean	Std. Dev.	Min	Max
<i>Roads allocated to contractors of the same name</i>					
Share <sub>t-1</sub>	8116	0.037	0.141	0.000	1.000
Share <sub>t</sub>	8116	0.035	0.134	0.000	1.000
ΔShare	8116	-0.002	0.151	-1.000	1.000
<i>Candidate characteristics</i>					
Vote share	8036	0.279	0.103	0.020	0.837
Margin	8116	0.000	0.102	-0.695	0.695
Incumbent	8116	0.282	0.450	0.000	1.000
Runner-up previous election	8116	0.147	0.354	0.000	1.000
Age	7357	49.164	10.207	23	87
Female candidate	8116	0.062	0.242	0.000	1.000
Candidate with criminal charges	4434	0.182	0.386	0.000	1.000
Total assets (1000000s of INR)	5010	106	4740	0.000	300000
Liabilities (1000000s of INR)	5286	1.841	17.800	0.000	644
University graduate	5286	0.596	0.491	0.000	1.000
Postgraduate degree	5286	0.192	0.394	0.000	1.000
Congress candidate	7239	0.295	0.456	0.000	1.000
BJP candidate	7239	0.203	0.403	0.000	1.000
Named Kumar	8116	0.058	0.234	0.000	1.000
Named Lal	8116	0.022	0.145	0.000	1.000
Named Patel	8116	0.009	0.094	0.000	1.000
Named Ram	8116	0.018	0.133	0.000	1.000
Named Reddy	8116	0.016	0.124	0.000	1.000
Named Singh	8116	0.112	0.316	0.000	1.000
Named Yadav	8116	0.014	0.117	0.000	1.000
<i>Constituency characteristics</i>					
Reserved seat	8116	0.335	1.349	0.000	84.092
Road count <sub>t</sub>	8116	27.691	30.822	1.000	479
Road count <sub>t-1</sub>	8116	22.086	25.744	1.000	388
Mean road length <sub>t</sub>	8116	5.833	3.999	0.350	42.654
Mean road length <sub>t-1</sub>	8116	4.963	3.838	0.410	53.985
Mean population	7822	961.697	633.986	30.000	7230
Mean SC/ST population	7822	244.078	193.401	0.000	2283
Mean connectivity	7822	0.561	0.308	0.000	1.000

Note: The number of observations varies due to missing values. Reserved seat refers to constituency reserved for MLAs from scheduled castes or tribes. Road count<sub>t</sub> is computed at the term-level by counting the number of road contracts signed in a constituency within a term. Mean road length is the average length of roads (in km) built in a constituency and term. Mean population and mean SC/ST population are averages of 2001 census data for all of a constituency's villages. Mean connectivity is the share of a constituency's villages that had all-weather road access at the time of the 2001 census.



**Table 1B: Descriptive Statistics (Roads built by same name contractors)**

Variable	Observations	Mean	Std. Dev.	Min	Max
Length of road	4924	3.995482	3.653428	0	41
Cost (1000000s of INR)	4924	133.5844	149.6136	0	2730.56
Days overrun	3204	587.9772	543.7035	-1750	3932
Actual cost/sanctioned cost	3868	0.953493	0.257386	0	9.349162
Failed Inspection	1513	0.364838	0.481544	0	1
Road missing	1817	0.024766	0.155454	0	1
All-weather road missing	1817	0.259219	0.438326	0	1
Months since election	4924	31.62145	16.78559	0	74

Note: Table 1b provides descriptive statistics for the sample used in Tables 4, 5 and columns 4-6 of Table 6. Cost, delays, cost overruns, and quality inspections are based on the PMGSY data. Quality is a dummy variable equal to one if the road is “unsatisfactory” or “in need of improvement” in latest inspection. “Missing roads” are defined on the basis of the 2011 census and the PMGSY data. For the first definition, we use roads officially completed before start of 2011 census in the PMGSY data, and set the missing dummy equal to one if all villages on the completed PMGSY road had no road of any type (tarmac, gravel, or wbm) according to the 2011 census. The dummy for the all-weather road missing variable takes the value of one if any village on the route of an officially completed road lacks all-weather road access.

**Table 2: Randomization for local linear regression at 6.2% bandwidth**

	Observations	Winner	Standard error
<i>Panel A: Candidate characteristics</i>			
Share of same name contractors <sub>t-1</sub>	4,396	-0.0072	(0.0084)
Incumbent	4,396	-0.0366	(0.0296)
Runner-up in previous election	4,396	-0.0119	(0.0214)
Age	4,036	0.2874	(0.6014)
Female candidate	4,396	-0.0004	(0.0129)
Candidate with criminal charge	2,536	-0.0421	(0.0293)
Total assets (1000000s of INR)	2,770	261.119	(718.93)
Liabilities (1000000s of INR)	3,049	0.1260	(0.5146)
Candidate with university degree	3,049	-0.0139	(0.0328)
Candidate with post-grad. degree	3,049	-0.0047	(0.0271)
BJP candidate	3,940	0.0104	(0.0274)
Congress candidate	3,940	-0.0194	(0.0312)
<i>Panel B: Share of roads built by contractors of same name in term prior to election</i>			
Share 5 years before election	2,502	0.0078	(0.0111)
Share 4 years before election	2,898	-0.0151	(0.0124)
Share 3 years before election	2,634	-0.0096	(0.0116)
Share 2 years before election	1,688	0.0037	(0.0133)
Share 1 year before election	1,866	-0.0131	(0.0152)
<i>Panel C: Prevalence of most common names</i>			
Named Kumar	4,396	0.0119	(0.0136)
Named Lal	4,396	-0.0019	(0.0084)
Named Patel	4,396	0.0026	(0.0061)
Named Ram	4,396	-0.0021	(0.0076)
Named Reddy	4,396	0.0074	(0.0054)
Named Singh	4,396	0.0195	(0.0171)
Named Yadav	4,396	0.0052	(0.0076)
Note: Coefficients are estimated by regressing the row variables on winner, the vote margin, and the vote margin interacted with winner in OLS regressions Standard errors are clustered at the election level. The bandwidth of 6.2% is derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011).			

**Table 3: Local linear regression RD**

$\Delta$ Share of same name contractors <sub>t</sub>	Whole Sample		Margin of Victory <6.2%		Margin of Victory <5%		Margin of Victory <2.5%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Winner	0.0094 (0.0052)	0.0099 (0.0054)	0.0252 (0.0092)	0.0242 (0.0097)	0.0202 (0.0102)	0.0205 (0.0109)	0.0267 (0.0141)	0.0261 (0.0152)
Margin	-0.0001 (0.0299)	0.0114 (0.0337)	-0.3153 (0.1856)	-0.2697 (0.1839)	-0.2949 (0.2389)	-0.2652 (0.2513)	-1.0043 (0.6660)	-1.1498 (0.7430)
Margin*winner	-0.0059 (0.0414)	-0.0353 (0.0445)	0.0835 (0.2724)	0.0577 (0.2671)	0.3363 (0.3507)	0.2797 (0.3663)	1.4266 (0.8778)	1.7690 (0.9724)
Incumbent		-0.0014 (0.0045)		0.0034 (0.0065)		0.0030 (0.0072)		-0.0119 (0.0092)
Runner-up in previous election		0.0068 (0.0055)		0.0074 (0.0079)		0.0069 (0.0085)		0.0024 (0.0125)
Female candidate		-0.0012 (0.0066)		-0.0130 (0.0107)		-0.0104 (0.0097)		-0.0048 (0.0138)
Age		0.0002 (0.0002)		-0.0000 (0.0003)		0.0001 (0.0003)		0.0004 (0.0004)
AC controls		X		X		X		X
State fixed effects		X		X		X		X
Election year fixed effects		X		X		X		X
N	8,116	7,290	4,396	4,012	3,760	3,432	2,104	1,915

Note: Local linear regression estimates. Standard errors are clustered at the election-level. Variables are defined either in the text or in the note for table 1. AC controls include: Reserved seat, Road count<sub>t-1</sub>, Mean population, Mean SC/ST population, Mean connectivity, Mean road length<sub>t-1</sub>. The bandwidth of 6.2% is derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011).

**Table 4: Road-level regression discontinuity - estimates for cost**

Ln(sanctioned cost/km)	Margin of victory <5%		Margin of victory <2.7%		Margin of victory <2.5%	
	(1)	(2)	(3)	(4)	(5)	(6)
MLAsamename	0.2328 (0.0895)	0.0683 (0.0376)	0.3122 (0.1166)	0.1238 (0.0522)	0.2735 (0.1170)	0.1320 (0.0561)
Margin	-3.5799 (2.3060)	-1.8487 (1.0848)	-9.0103 (5.4485)	-6.9359 (2.8075)	-10.3331 (5.8739)	-8.0188 (3.1738)
Margin* MLAsamename	-1.5417 (3.5541)	2.6023 (1.5201)	4.3953 (7.9768)	7.5140 (3.6619)	10.7941 (8.6207)	8.6252 (3.9456)
Months since election	0.0045 (0.0012)	-0.0008 (0.0006)	0.0054 (0.0013)	-0.0010 (0.0007)	0.0056 (0.0013)	-0.0008 (0.0008)
Ln(length)	0.0171 (0.0357)	-0.0554 (0.0257)	-0.0407 (0.0478)	-0.0718 (0.0349)	-0.0478 (0.0500)	-0.0695 (0.0366)
Reserved seat	-0.0155 (0.0936)	0.0491 (0.0402)	-0.1598 (0.1275)	0.0556 (0.0627)	-0.2165 (0.1353)	0.0562 (0.0719)
Mean population of habitations	-0.0299 (0.0302)	-0.0129 (0.0111)	-0.0401 (0.0355)	-0.0225 (0.0158)	-0.0502 (0.0369)	-0.0230 (0.0165)
SCST share of habitations	-0.1258 (0.0874)	0.0061 (0.0442)	-0.1210 (0.1144)	0.0352 (0.0640)	-0.0082 (0.1089)	0.0524 (0.0661)
State fixed effects		X		X		X
Agreement year fixed effects		X		X		X
N	2,418	2,418	1,542	1,542	1,435	1,435

Note: Standard errors clustered at the contractor level to account for intra-contractor correlation of the error term at the road level. We use Ln(length) to account for non-linear relationship between cost and distance. The bandwidth of 2.7% is derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011).

**Table 5: Road-level regression discontinuity - estimates for quality**

Dependent variable:	Days overrun	Ratio: actual cost to sanctioned cost	Failed inspection	Road missing	All-weather road missing
Margin of victory:	<3%	<3.5%	<4.3%	<4.4%	<2.5%
	(1)	(2)	(3)	(4)	(5)
MLAsamename	-33.5 (87.4)	-0.0492 (0.0402)	0.1415 (0.0771)	0.1144 (0.0542)	0.3604 (0.1056)
Margin	-508.6 (3,965.8)	2.0571 (1.7202)	-0.4117 (2.7463)	-1.4931 (1.2592)	-11.6416 (4.9986)
Margin*MLAsamename	-1,174.6 (5,195.0)	-1.6319 (2.1579)	-4.9992 (3.7064)	-1.0067 (1.9612)	1.3525 (7.2428)
Months since election	2.8 (1.3)	0.0006 (0.0005)	0.0014 (0.0015)	-0.0002 (0.0006)	0.0018 (0.0016)
Ln(length)	43.8 (20.2)	0.0180 (0.0100)	0.0137 (0.0239)	-0.0188 (0.0106)	0.0039 (0.0278)
Reserved seat	102.0 (92.6)	0.0575 (0.0808)	0.0589 (0.0607)	-0.0149 (0.0276)	0.0348 (0.0907)
Mean population of habitations	-28.5 (19.7)	0.0000 (0.0054)	0.0256 (0.0247)	-0.0000 (0.0044)	-0.0271 (0.0240)
SCST share of habitations	-55.3 (66.7)	-0.0136 (0.0304)	-0.1550 (0.0796)	-0.0017 (0.0193)	-0.1367 (0.0942)
State fixed effects	X	X	X	X	X
Agreement year FE	X	X	X	X	X
N	1,139	1,570	758	796	490

Note: Standard errors clustered at the contractor level to account for intra-contractor correlation of the error term at the road level. The optimal bandwidths of 3% (for days overrun), 3.5% (for cost overruns), 4.3% (for failed inspections), 4.4% (for missing roads) and 2.5% (for missing all-weather roads) are derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011). All regressions include a constant..

**Table 6: Three tests for electioneering**

$\Delta\text{Share}_t$	Electoral cycles: allocation			Electoral cycles: cost			Effect in “politically irrelevant” areas
	Start of term heterogeneity	End of term heterogeneity	Start and end of term heterogeneity	Start of term heterogeneity	End of term heterogeneity	Start and end of term heterogeneity	
	(1)	(2)	(3)	(4)	(5)	(6)	
Winner (MLAsamename in cols 4-6)	0.0288 (0.0089)	0.0246 (0.0108)	0.0236 (0.0113)	0.137 (0.053)	0.146 (0.054)	0.163 (0.056)	0.0279 (0.0109)
Margin	-0.1764 (0.2095)	0.0323 (0.2605)	0.0282 (0.2603)	-7.107 (2.797)	-7.053 (2.808)	-7.063 (2.776)	-0.0961 (0.1882)
Margin*winner (Margin*MLAsamename in cols 4-6)	-0.3650 (0.3590)	-0.4818 (0.4445)	-0.4740 (0.4442)	7.866 (3.654)	7.486 (3.651)	7.682 (3.631)	-0.2924 (0.2692)
Start of term	-0.0060 (0.0057)		-0.0083 (0.0066)	0.133 (0.054)		0.184 (0.062)	
Start of term*winner (Start of term*MLAsamename in cols 4-6)	0.0021 (0.0072)		0.0023 (0.0084)	-0.092 (0.054)		-0.111 (0.054)	
End of term		0.0174 (0.0087)	0.0138 (0.0086)		0.032 (0.082)	-0.000 (0.085)	
End of term*winner (End of term*MLAsamename in cols 4-6)		-0.0103 (0.0092)	-0.0093 (0.0094)		-0.134 (0.057)	-0.156 (0.059)	
Politically irrelevant*post announcement*winner							0.0144 (0.0218)
Constituency controls	X	X	X	X	X	X	X
Candidate controls	X	X	X				X
Road-level controls				X	X	X	
State fixed effects	X	X	X	X	X	X	X
Agreement year fixed effects	X	X	X	X	X	X	X
N	5,572	4,188	4,188	1,542	1,542	1,542	4,658

Note: Terms in the sample for columns 1-3 are disaggregated by time: first year, last year, and remaining period. Terms in the sample for column 7 are disaggregated by time (pre- and post- the announcement of delimitation) and spatially (based on intra-constituency boundaries introduced by delimitation). The RD bandwidth for columns 1-3 and 7 is 6.2% and the bandwidth for columns 4-6 is 2.7%. Standard errors are clustered at the election level in columns 1-3 and 7 and at the contractor level in columns 4-6. Controls for columns 1-3 and 7 are the same as in Table 3. Controls for columns 4-6 are the same as in Table 4. Regression in column 7 includes all lower-order interactions (not reported). All regressions include a constant. Tables A8-A10 of the online appendix show results for a range of bandwidths.

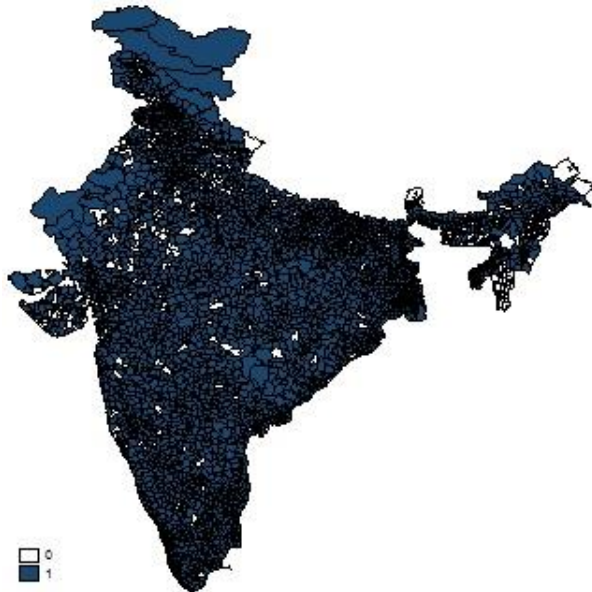
**Table 7: E-Procurement**

$\Delta$ Share of same name contractors <sub>t</sub>	Whole Sample		Margin of Victory <6.2%		Margin of Victory <5%		Margin of Victory <2.5%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Winner	0.0092 (0.0060)	0.0113 (0.0067)	0.0278 (0.0107)	0.0273 (0.0117)	0.0242 (0.0119)	0.0234 (0.0134)	0.0273 (0.0169)	0.0292 (0.0190)
Margin	-0.0274 (0.0266)	-0.0073 (0.0318)	-0.4330 (0.2043)	-0.5064 (0.2121)	-0.4493 (0.2745)	-0.5726 (0.2980)	-1.1868 (0.8038)	-1.6060 (0.9125)
Margin*winner	0.0420 (0.0407)	-0.0235 (0.0461)	0.1885 (0.2684)	0.3052 (0.2833)	0.4848 (0.3641)	0.7512 (0.3832)	1.9803 (1.0067)	2.6059 (1.1523)
E-procurement	-0.0052 (0.0039)	-0.0073 (0.0059)	-0.0015 (0.0056)	-0.0114 (0.0090)	-0.0029 (0.0059)	-0.0134 (0.0096)	-0.0055 (0.0085)	-0.0116 (0.0131)
E-procurement*winner	0.0033 (0.0058)	0.0038 (0.0064)	0.0032 (0.0086)	0.0070 (0.0093)	0.0021 (0.0093)	0.0052 (0.0102)	0.0031 (0.0125)	0.0053 (0.0139)
AC controls		X		X		X		X
Candidate controls		X		X		X		X
State fixed effects		X		X		X		X
Election year fixed effects		X		X		X		X
N	9,990	8,612	5,324	4,665	4,562	3,996	2,596	2,260

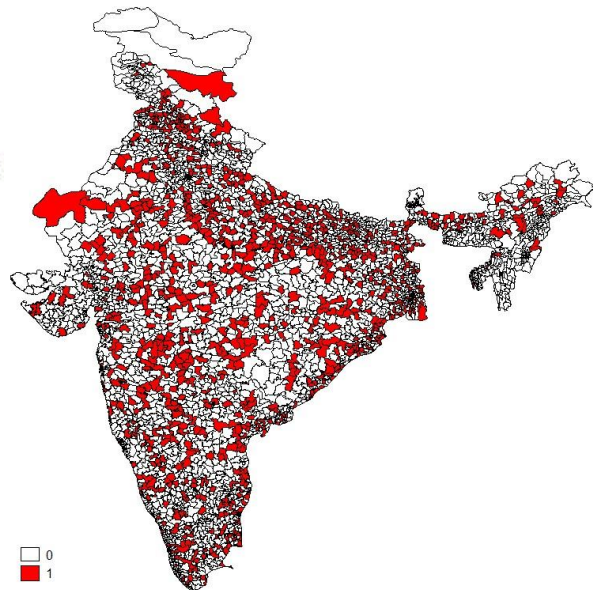
Note: Local linear regression estimates. Standard errors are clustered at the election-level. Variables are defined either in the text or in the note for table 1. AC controls include: Reserved seat, Road count<sub>t-1</sub>, Mean population, Mean SC/ST population, Mean connectivity, Mean road length<sub>t-1</sub>. Candidate controls: age, gender, incumbency, former-runner-up status.

# MAPS

Map 1: Constituencies in Sample

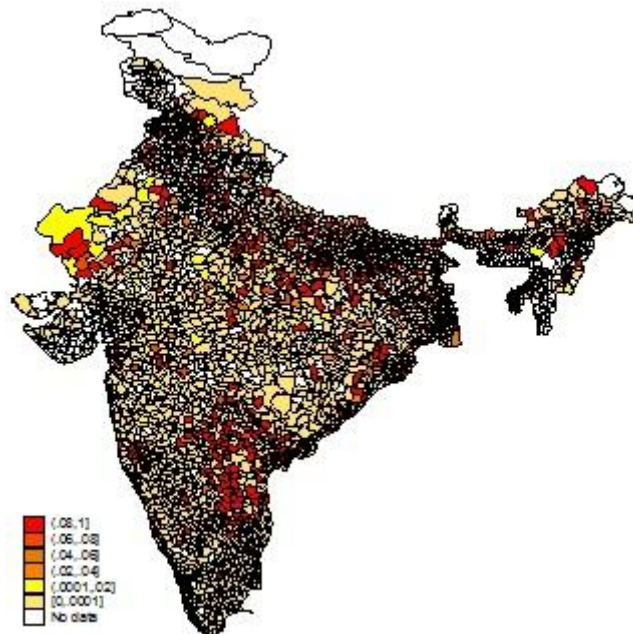


Map 2: Constituencies with close elections



Note: The constituencies shown on maps are based on the pre-2008 delimitation. Post delimitation data is assigned to pre-delimitation boundaries. In Map 1 all constituencies with consecutive electoral terms with PMGSY road construction are shaded blue. In Map 2 constituencies with at least one election with a margin of victory lower than 6.2% (derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011)) are shaded red.

Map 3: Variation in the baseline share of same name contractors



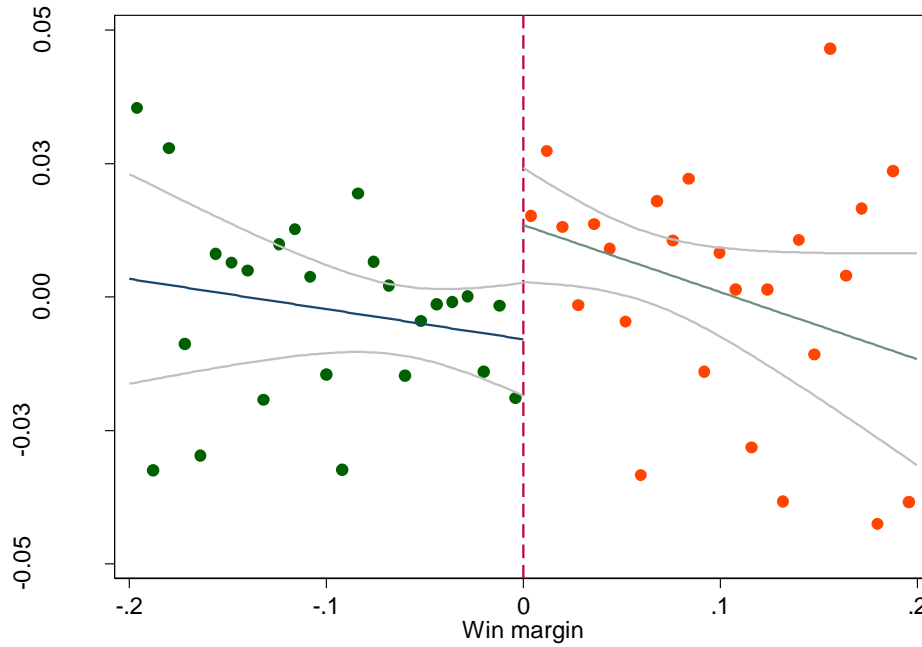
Note: The constituencies shown on maps are based on the pre-2008 delimitation. Post delimitation data is assigned to pre-delimitation boundaries. Map 3 plots  $share_{ijt}$  for all constituencies in the sample. Darker shades indicate a higher value of  $share_{ijt}$ .



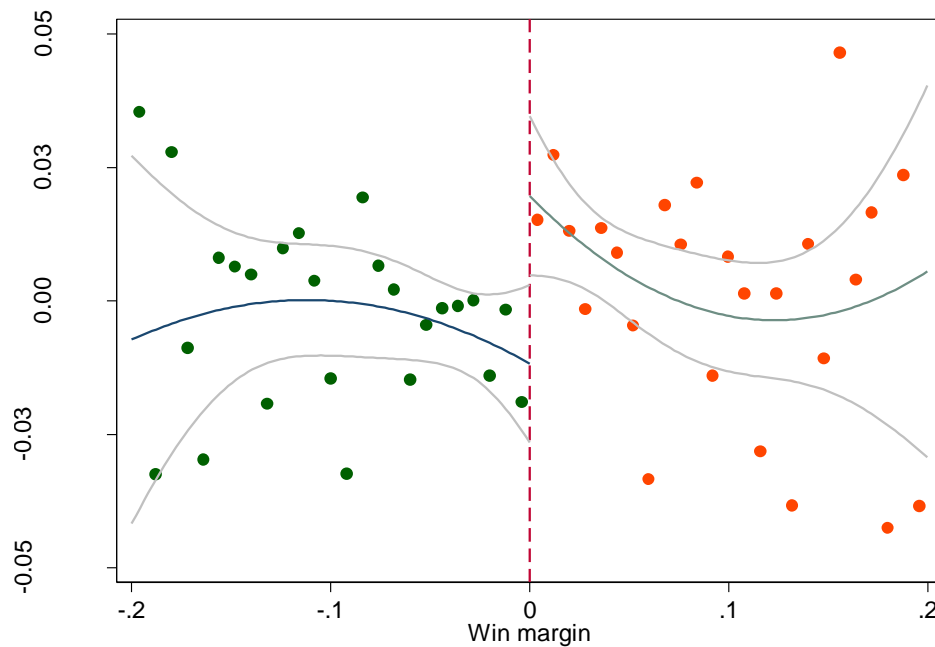
# FIGURES

**Figure 2: Graphical depiction of RD**

**Change in share of same name contractors – linear fit**

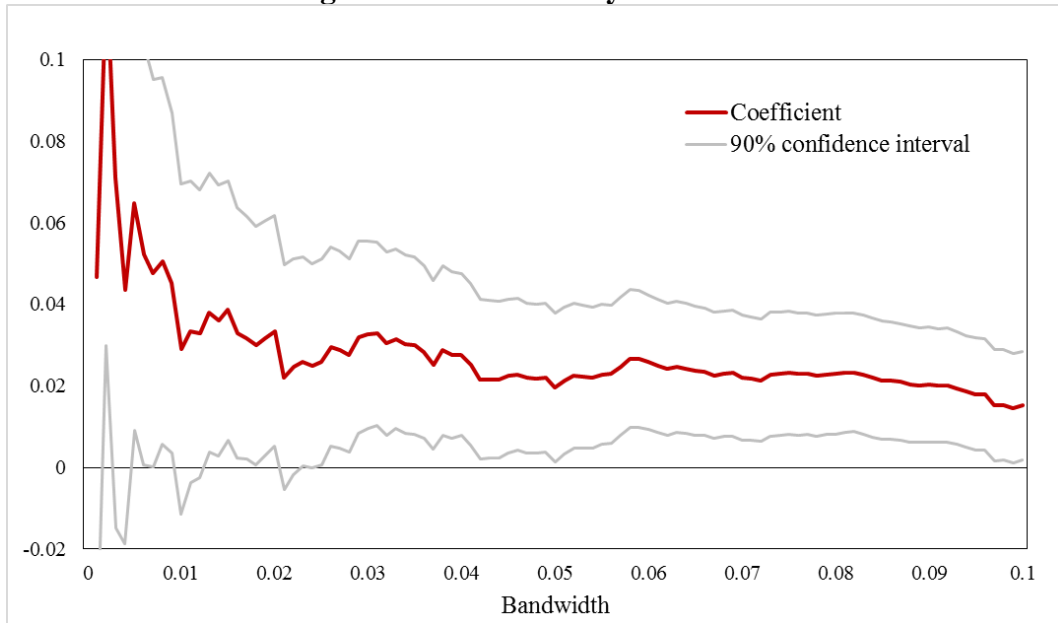


**Change in share of same name contractors – quadratic fit**



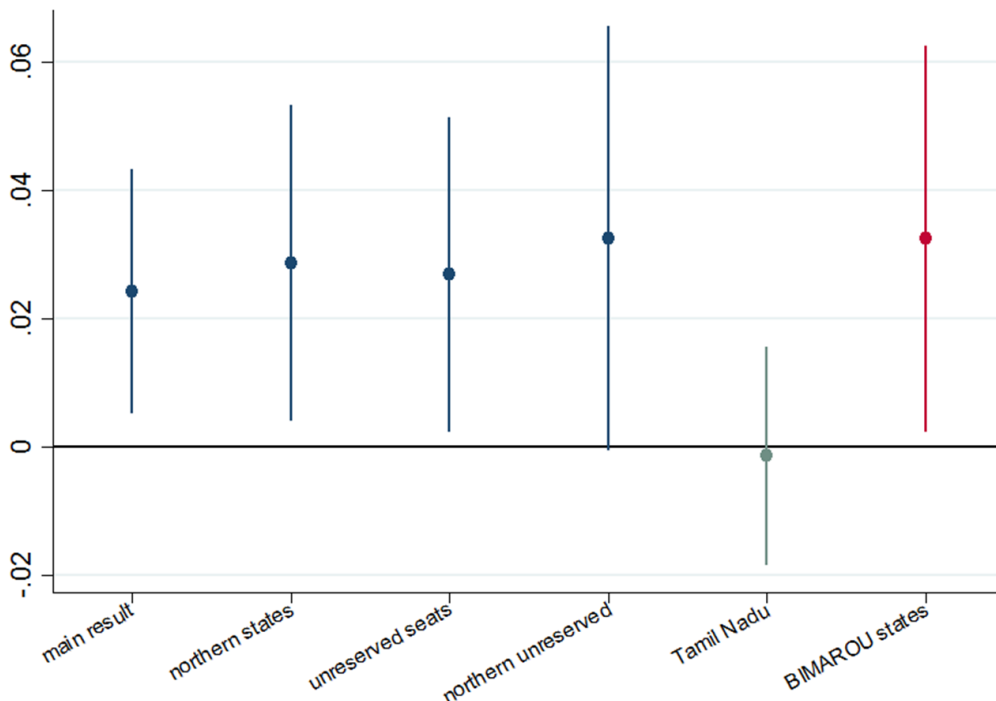
Note: Lines fitted separately on the samples left and right of the cut-off. 5% confidence intervals plotted in grey. Each marker represents a bin of 100 observations.

**Figure 1: Main effect by bandwidth**



Note: The chart plots the coefficient for winner in our main specification (equivalent to Table 3, columns 2, 4, 6, and 8) with the full set of candidate and constituency controls as well as state and year fixed effects.

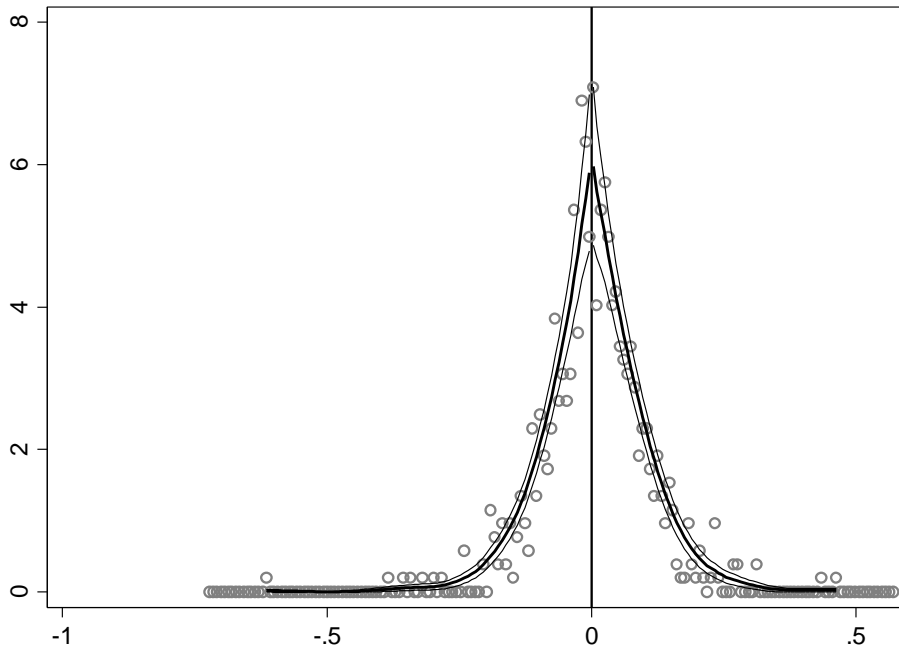
**Figure 3: heterogeneous effect**



Note: Chart plots the coefficient for our main specification at the 6.2% bandwidth, for a range of samples: (1) the full sample; (2) northern states, which include Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Madhya Pradesh, Punjab, Odisha, Rajasthan, Uttarakhand, Uttar Pradesh and West Bengal; (3) constituencies not reserved for Scheduled Caste or Scheduled Tribe candidates; (4) the intersection of (2) and (3); (5) the state of Tamil Nadu; and (6) BIMAROU states (see footnote 23).

# FOR ONLINE PUBLICATION:

**Figure A1: graphical depiction of the McCrary test**



Note: This figure plots the McCrary test. The running variable in this analysis is continuously distributed by construction. The test is performed on an alternative version of the margin variable: the margin of victory for the candidate with the higher level of  $share_{ijt-1}$ .

**Table A1: Sources of election data**

Source	Years covered	No. of elections	No. of candidates	Candidate-level variables used in the sample
ECI digitised data	2005-2014	7,328	80,323	name, vote share, gender, party
Bhavani (2012)	1977-2012	31,422	300,087	name, vote share, gender, party
Empowering India	1951-2015	19,715	196,935	assets, education, age
National Election Watch	2004-2015	8,944	73,200	assets, liabilities, education, criminal charges

Note: ECI digitised data refers to a subset of the ECI data that are available online at [eci.nic.in](http://eci.nic.in). Bhavani (2012) is a dataset kindly made public by Rikhil R. Bhavnani. Empowering India and National Election Watch are NGOs. Their data is accessible at [www.empoweringindia.com](http://www.empoweringindia.com) and [myneta.info](http://myneta.info) respectively. Where a variable is listed twice in the fifth column, this is due to incomplete time series or missing values that are filled in by drawing on multiple datasets.

**Table A2: Local linear regression RD - levels**

Share of same name contractors <sub>t</sub>	Whole Sample		Margin of Victory <6.2%		Margin of Victory <5%		Margin of Victory <2.5%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Winner	0.0085 (0.0032)	0.0116 (0.0047)	0.0121 (0.0054)	0.0161 (0.0079)	0.0097 (0.0059)	0.0082 (0.0089)	0.0149 (0.0077)	0.0187 (0.0111)
Margin	0.0119 (0.0212)	-0.0084 (0.0250)	-0.0891 (0.1120)	-0.1013 (0.1443)	-0.0298 (0.1443)	0.0437 (0.1889)	-0.7048 (0.4017)	-0.8124 (0.4440)
Margin*winner	-0.0364 (0.0349)	0.0114 (0.0435)	0.0632 (0.1802)	0.1070 (0.2385)	0.0770 (0.2476)	0.2981 (0.3435)	1.0360 (0.6630)	1.1712 (0.7479)
Incumbent		0.0085 (0.0040)		0.0140 (0.0056)		0.0155 (0.0062)		0.0106 (0.0079)
Runner-up in previous election		0.0067 (0.0048)		0.0020 (0.0067)		0.0043 (0.0073)		0.0026 (0.0103)
Female candidate		-0.0192 (0.0049)		-0.0174 (0.0071)		-0.0207 (0.0070)		-0.0115 (0.0118)
Age		0.0000 (0.0002)		-0.0001 (0.0002)		-0.0001 (0.0002)		0.0000 (0.0003)
AC controls		X		X		X		X
State fixed effects		X		X		X		X
Election year fixed effects		X		X		X		X
N	15,208	7,068	8,202	3,921	6,964	3,365	3,816	1,880

Note: Local linear regression estimates. Standard errors are clustered at the election-level. Variables are defined either in the text or in the note for table 1. AC controls include: Reserved seat, Road count<sub>t-1</sub>, Mean population, Mean SC/ST population, Mean connectivity, Mean road length<sub>t-1</sub>.

**Table A3: Local linear regression RD – Alignment with Chief Minister**

$\Delta$ Share of same name contractors <sub>t</sub>	Whole Sample		Margin of Victory <6.2%		Margin of Victory <5%		Margin of Victory <2.5%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Winner	0.0054 (0.0057)	0.0065 (0.0060)	0.0191 (0.0099)	0.0192 (0.0103)	0.0134 (0.0109)	0.0146 (0.0116)	0.0206 (0.0150)	0.0226 (0.0160)
Margin	0.0046 (0.0305)	0.0165 (0.0342)	-0.3143 (0.1879)	-0.2717 (0.1853)	-0.2873 (0.2425)	-0.2541 (0.2544)	-0.9884 (0.6741)	-1.1450 (0.7454)
Margin*winner	-0.0119 (0.0420)	-0.0439 (0.0451)	0.0589 (0.2718)	0.0106 (0.2679)	0.3182 (0.3529)	0.2247 (0.3713)	1.4494 (0.8854)	1.7467 (0.9771)
Aligned	-0.0053 (0.0061)	-0.0053 (0.0067)	-0.0007 (0.0077)	0.0012 (0.0083)	-0.0028 (0.0082)	-0.0026 (0.0089)	-0.0062 (0.0110)	-0.0042 (0.0122)
Aligned*winner	0.0103 (0.0077)	0.0099 (0.0084)	0.0158 (0.0105)	0.0150 (0.0109)	0.0171 (0.0112)	0.0174 (0.0117)	0.0139 (0.0151)	0.0105 (0.0162)
AC controls		X		X		X		X
Individual controls		X		X		X		X
State fixed effects		X		X		X		X
Election year fixed effects		X		X		X		X
N	8,108	7,290	4,392	4,012	3,756	3,432	2,101	1,915

Note: Local linear regression estimates. Standard errors are clustered at the election-level. Variables are defined either in the text or in the note for table 1. AC controls include: Reserved seat, Road count<sub>t-1</sub>, Mean population, Mean SC/ST population, Mean connectivity, Mean road length<sub>t-1</sub>.

**Table A4: Cost estimates with additional controls (2.7% bandwidth)**

Dependent variable: Ln(sanctioned cost/km)	(1)	(2)	(3)	(4)	(5)
MLAsamename	0.083 (0.042)	0.084 (0.042)	0.078 (0.042)	0.083 (0.042)	0.079 (0.042)
Margin	-5.298 (1.957)	-5.276 (1.955)	-4.970 (1.968)	-5.349 (1.964)	-4.998 (1.972)
Margin* MLAsamename	5.562 (2.830)	5.536 (2.828)	5.310 (2.771)	5.596 (2.817)	5.316 (2.750)
Bridge		0.258 (0.272)			0.267 (0.265)
Altitude			-0.068 (0.033)		-0.069 (0.034)
Ruggedness				14.798 (13.753)	15.849 (12.941)
Road level controls	X	X	X	X	X
State fixed effects	X	X	X	X	X
Agreement year fixed effects	X	X	X	X	X
N	1,394	1,394	1,394	1,394	1,394

Note: Standard errors clustered at the contractor level to account for intra-contractor correlation of the error term at the road level. The bandwidth of 2.7% is derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2012) All regressions include the same set of road level controls as Table 6.

**Table A5: Parametric regression discontinuity estimated on full sample**

$\Delta\text{Share}_t$	Linear		Quadratic Polynomials		Cubic Polynomials	
	(1)	(2)	(3)	(4)	(5)	(6)
Winner	0.0094 (0.0052)	0.0094 (0.0055)	0.0138 (0.0062)	0.0138 (0.0067)	0.0190 (0.0075)	0.0193 (0.0082)
Margin	-0.0001 (0.0299)	0.0271 (0.0345)	-0.0226 (0.0609)	-0.0058 (0.0719)	-0.1004 (0.1209)	-0.1562 (0.1506)
Margin*Winner	-0.0059 (0.0414)	-0.0504 (0.0456)	-0.0740 (0.0870)	-0.0998 (0.0994)	-0.1247 (0.1604)	-0.0253 (0.1890)
Margin <sup>2</sup>			-0.0764 (0.1540)	-0.1130 (0.1921)	-0.6165 (0.6727)	-1.2326 (0.9282)
Margin*Winner <sup>2</sup>			0.3834 (0.2689)	0.3849 (0.2869)	1.8161 (1.0541)	2.0317 (1.2413)
Margin <sup>3</sup>					-0.8224 (0.8418)	-1.8595 (1.3418)
Margin*Winner <sup>3</sup>					-0.5365 (1.0901)	1.0619 (1.5242)
Constituency-level controls		X		X		X
Candidate-level controls		X		X		X
State fixed effects		X		X		X
Election year fixed effects		X		X		X
N	8,116	7,068	8,116	7,068	8,116	7,068

Note: Estimated by OLS. Standard errors are clustered at the election-level. Variables are defined either in the text or in the note for table 1. The candidate and constituency level controls in columns 2, 4, 6 and 8 are the same as in column (2) of Table 3. These controls and state and election year fixed effects are not reported. All regressions include a constant.



**Table A6: 2<sup>nd</sup> vs 3<sup>rd</sup> Placebo Tests**

$\Delta\text{Share}_t$	All 2 <sup>nd</sup> vs 3 <sup>rd</sup> with margin < 6.2%		2 <sup>nd</sup> vs 3 <sup>rd</sup> where margin between 1 <sup>st</sup> and 2 <sup>nd</sup> < 6.2%		2 <sup>nd</sup> vs 3 <sup>rd</sup> where both margins < 6.2%		2 <sup>nd</sup> vs 3 <sup>rd</sup> where margin between 1 <sup>st</sup> and 2 <sup>nd</sup> < 6.2% and margin between 2 <sup>nd</sup> and 3 <sup>rd</sup> < 14%	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2 <sup>nd</sup> place	-0.0170 (0.0130)	-0.0085 (0.0144)	0.0025 (0.0090)	0.0029 (0.0099)	-0.0019 (0.0181)	0.0092 (0.0198)	-0.0069 (0.0130)	-0.0049 (0.0144)
Margin (2 <sup>nd</sup> vs 3 <sup>rd</sup> )	-0.0137 (0.3041)	-0.1024 (0.3328)	-0.0003 (0.0301)	0.0510 (0.0411)	-0.2903 (0.4384)	-0.5531 (0.4660)	0.1415 (0.1276)	0.1602 (0.1495)
Margin*2 <sup>nd</sup> place	0.2128 (0.3748)	0.1908 (0.4268)	-0.0169 (0.0389)	-0.1412 (0.0616)	0.2809 (0.5037)	0.3856 (0.6039)	-0.1219 (0.1705)	-0.2099 (0.2019)
Incumbent		-0.0007 (0.0120)		0.0061 (0.0081)		-0.0015 (0.0146)		0.0046 (0.0116)
Runner-up in previous election		0.0034 (0.0139)		0.0161 (0.0100)		0.0033 (0.0192)		0.0201 (0.0142)
Female candidate		-0.0189 (0.0185)		-0.0080 (0.0111)		-0.0097 (0.0214)		-0.0136 (0.0188)
Age		-0.0002 (0.0004)		-0.0002 (0.0003)		-0.0003 (0.0006)		-0.0005 (0.0004)
AC controls		X		X		X		X
State fixed effects		X		X		X		X
Election year fixed effects		X		X		X		X
N	2,120	1,760	4,228	3,408	1,300	1,090	2,158	1,812

Note: the regressions in this table are equivalent to those in Table 3, except that these compare 2<sup>nd</sup> placed candidates to 3<sup>rd</sup> placed candidates. In columns (1) and (2) the sample is restricted to candidates whose difference in vote shares was smaller than 6.2% (as in the main specification). Columns (3) and (4) compare 2<sup>nd</sup> and 3<sup>rd</sup> placed candidates in the elections that make up the sample for the main specification. The sample for columns (5) and (6) is the intersection of the previous two, including candidates from elections where both the gap between 1<sup>st</sup> and 2<sup>nd</sup> and between 2<sup>nd</sup> and 3<sup>rd</sup> was smaller than 6.2%. Columns (7) and (8) are the same, except that the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011) is used to derive the optimal bandwidth. Standard errors are clustered at the election level. All regressions include a constant.

**Table A7: Road-level regression discontinuity - estimates for quality**

Dependent variable:	Days overrun			Ratio: actual cost to sanctioned cost			Failed inspection			Road missing		
	<5%	<3%	<2.5%	<5%	<3.5%	<2.5%	<5%	<4.3%	<2.5%	<5%	<4.4%	<2.5%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Margin of victory:												
MLAsamename	-68.8 (66.2)	-33.5 (87.4)	-74.7 (98.1)	-0.0314 (0.0343)	-0.0492 (0.0402)	-0.0234 (0.0520)	0.1477 (0.0720)	0.1415 (0.0771)	0.0945 (0.1011)	0.1034 (0.0475)	0.1144 (0.0542)	0.2124 (0.0880)
Margin	-451.6 (1,800.2)	-508.6 (3,965.8)	-590.3 (3,966.1)	0.9680 (1.1826)	2.0571 (1.7202)	1.2800 (2.8749)	-0.9762 (2.1643)	-0.4117 (2.7463)	-0.4000 (5.0564)	-0.3571 (0.9319)	-1.4931 (1.2592)	-5.3082 (2.5858)
Margin*MLAsamename	1,340.2 (2,721.4)	-1,174.6 (5,195.0)	3,241.8 (6,145.5)	-0.8965 (1.4272)	-1.6319 (2.1579)	-1.5568 (3.5434)	-4.4654 (2.9691)	-4.9992 (3.7064)	-2.5460 (6.5991)	-2.3055 (1.6350)	-1.0067 (1.9612)	-1.1932 (4.5318)
Months since election	1.9 (1.1)	2.8 (1.3)	3.9 (1.5)	0.0008 (0.0004)	0.0006 (0.0005)	0.0010 (0.0006)	0.0016 (0.0014)	0.0014 (0.0015)	-0.0006 (0.0018)	-0.0003 (0.0005)	-0.0002 (0.0006)	0.0010 (0.0009)
Ln(length)	36.6 (17.9)	43.8 (20.2)	51.5 (21.4)	0.0178 (0.0089)	0.0180 (0.0100)	0.0153 (0.0136)	0.0118 (0.0231)	0.0137 (0.0239)	0.0092 (0.0300)	-0.0184 (0.0097)	-0.0188 (0.0106)	-0.0222 (0.0148)
Reserved seat	26.5 (73.7)	102.0 (92.6)	112.7 (96.0)	0.0479 (0.0621)	0.0575 (0.0808)	0.1153 (0.1047)	0.0338 (0.0576)	0.0589 (0.0607)	0.1177 (0.0770)	0.0219 (0.0353)	-0.0149 (0.0276)	-0.0240 (0.0441)
Mean population of habitations	-23.0 (15.9)	-28.5 (19.7)	-17.3 (20.7)	-0.0012 (0.0048)	0.0000 (0.0054)	0.0067 (0.0071)	0.0219 (0.0238)	0.0256 (0.0247)	-0.0226 (0.0360)	-0.0006 (0.0044)	-0.0000 (0.0044)	0.0047 (0.0061)
SCST share of habitations	-15.9 (62.3)	-55.3 (66.7)	-111.4 (69.8)	-0.0032 (0.0237)	-0.0136 (0.0304)	-0.0223 (0.0421)	-0.1604 (0.0771)	-0.1550 (0.0796)	-0.0528 (0.1149)	-0.0008 (0.0188)	-0.0017 (0.0193)	-0.0224 (0.0339)
State fixed effects	X	X	X	X	X	X	X	X	X	X	X	X
Agreement year FE	X	X	X	X	X	X	X	X	X	X	X	X
N	1,604	1,139	940	1,933	1,570	1,139	820	758	482	870	796	490

Note: Standard errors clustered at the contractor level to account for intra-contractor correlation of the error term at the road level. The optimal bandwidths of 3% (for days overrun), 3.5% (for cost overruns), 4.3% (for failed inspections), and 4.4% (for missing roads) are derived from the optimal bandwidth choice rule of Imbens and Kalyanaraman (2011). All regressions include a constant.

**Table A8: RD test for electoral cycles in preferential allocation**

$\Delta\text{Share}_t$	Start of term heterogeneity		End of term heterogeneity		Start and end of term heterogeneity	
	(1)	(2)	(3)	(4)	(5)	(6)
Winner	0.0249 (0.0081)	0.0288 (0.0089)	0.0264 (0.0084)	0.0246 (0.0108)	0.0260 (0.0087)	0.0236 (0.0113)
Margin	-0.1926 (0.1952)	-0.1764 (0.2095)	-0.1891 (0.1953)	0.0323 (0.2605)	-0.1898 (0.1952)	0.0282 (0.2603)
Margin*winner	-0.2357 (0.3363)	-0.3650 (0.3590)	-0.2411 (0.3363)	-0.4818 (0.4445)	-0.2400 (0.3363)	-0.4740 (0.4442)
Start of term	-0.0058 (0.0053)	-0.0060 (0.0057)			-0.0035 (0.0052)	-0.0083 (0.0066)
Start of term* winner	0.0021 (0.0069)	0.0021 (0.0072)			0.0010 (0.0070)	0.0023 (0.0084)
End of term			0.0078 (0.0058)	0.0174 (0.0087)	0.0063 (0.0057)	0.0138 (0.0086)
End of term* winner			-0.0034 (0.0069)	-0.0103 (0.0092)	-0.0030 (0.0070)	-0.0093 (0.0094)
Constituency Controls		X		X		X
Candidate Controls		X		X		X
State fixed effects		X		X		X
Agreement year fixed effects		X		X		X
N	6,266	5,572	6,266	4,188	6,266	4,188

Note: Standard errors clustered at the election level. All estimates conducted on 5% bandwidth. All regressions include a constant. For this analysis there are potentially three observations per electoral term: the value of  $\text{Share}_t$  for the first 12 months after an election, the value of  $\text{Share}_t$  for the last 12 months before the next election, and the value of  $\text{Share}_t$  over the remaining term. For constituencies where no roads were built in one of these periods, the number of observations will be less than three.

**Table A9: RD test for electoral cycles in cost**

Ln(sanctioned cost/km)	Start of term heterogeneity		End of term heterogeneity		Start and end of term heterogeneity	
	(1)	(2)	(3)	(4)	(5)	(6)
MLAsamename	0.311 (0.110)	0.137 (0.053)	0.298 (0.111)	0.146 (0.054)	0.332 (0.116)	0.163 (0.056)
Margin	-9.933 (4.921)	-7.107 (2.797)	-10.489 (5.054)	-7.053 (2.808)	-10.135 (4.897)	-7.063 (2.776)
Margin*MLAsamename	7.022 (6.935)	7.866 (3.654)	7.601 (7.034)	7.486 (3.651)	7.462 (6.918)	7.682 (3.631)
Start of term	0.176 (0.102)	0.133 (0.054)			0.222 (0.107)	0.184 (0.062)
Start of term* MLAsamename	-0.237 (0.104)	-0.092 (0.054)			-0.256 (0.109)	-0.111 (0.054)
End of term			-0.019 (0.098)	0.032 (0.082)	-0.009 (0.104)	-0.000 (0.085)
End of term* MLAsamename			-0.106 (0.105)	-0.134 (0.057)	-0.151 (0.110)	-0.156 (0.059)
Road level controls	X	X	X	X	X	X
State fixed effects		X		X		X
Agreement year fixed effects		X		X		X
N	1,542	1,542	1,542	1,542	1,542	1,542

Note: Standard errors clustered at the contractor level. All estimates conducted on 5% bandwidth. The set of road-level controls is the same as in Table 4 and Table 5. All regressions include a constant.

**Table A10: RD test for heterogeneity based on “political relevance”**

$\Delta$ Share:	Full sample with interactions		Margin of victory <6.2%		Margin of victory <5%		Margin of victory <2.5%	
	(1)	(2)					(3)	(4)
Winner	0.0114 (0.0057)	0.0138 (0.0061)	0.0300 (0.0101)	0.0279 (0.0109)	0.0267 (0.0106)	0.0297 (0.0116)	0.0226 (0.0140)	0.0200 (0.0151)
Margin	0.0139 (0.0282)	0.0124 (0.0305)	-0.1183 (0.1729)	-0.0961 (0.1882)	-0.1533 (0.2215)	-0.1654 (0.2394)	-0.2873 (0.6242)	-0.2439 (0.7029)
Margin*winner	-0.0138 (0.0445)	-0.0187 (0.0485)	-0.3593 (0.2589)	-0.2924 (0.2692)	-0.0746 (0.3233)	-0.1029 (0.3526)	0.2786 (0.8525)	0.6273 (0.9580)
Politically irrelevant	0.0028 (0.0052)	0.0078 (0.0067)	-0.0060 (0.0079)	0.0034 (0.0102)	-0.0030 (0.0086)	0.0106 (0.0112)	0.0045 (0.0107)	0.0157 (0.0145)
Post announcement	0.0047 (0.0070)	0.0062 (0.0085)	0.0045 (0.0097)	0.0077 (0.0114)	0.0067 (0.0107)	0.0120 (0.0125)	-0.0033 (0.0171)	-0.0008 (0.0196)
Politically irrelevant*winner	-0.0069 (0.0082)	-0.0103 (0.0088)	-0.0109 (0.0119)	-0.0151 (0.0129)	-0.0043 (0.0124)	-0.0160 (0.0136)	-0.0054 (0.0162)	-0.0142 (0.0177)
Post announcement* winner	-0.0151 (0.0103)	-0.0141 (0.0108)	-0.0165 (0.0145)	-0.0131 (0.0150)	-0.0150 (0.0158)	-0.0136 (0.0163)	0.0051 (0.0230)	0.0072 (0.0238)
Politically irrelevant*post announcement	-0.0032 (0.0099)	-0.0062 (0.0114)	0.0103 (0.0143)	0.0016 (0.0165)	0.0097 (0.0158)	-0.0025 (0.0181)	0.0123 (0.0220)	0.0034 (0.0249)
Politically irrelevant*post announcement* winner	0.0063 (0.0147)	0.0060 (0.0156)	0.0110 (0.0206)	0.0144 (0.0218)	-0.0049 (0.0219)	0.0070 (0.0231)	-0.0168 (0.0295)	-0.0080 (0.0315)
Constituency and candidate level controls		X		X		X		X
State fixed effects		X		X		X		X
Agreement year fixed effects		X		X		X		X
N	9,774	8,462	5,234	4,658	4,532	4,063	2,580	2,327

Note: Observations at the MLA level. Standard errors are clustered at the election-level. The term-level sample is disaggregated, allowing for multiple observations per term. The sample for Columns 3 and 4 is restricted (i) to areas that did not remain part of the same constituency after delimitation and (ii) to the time period between the announcement of the delimitation reform and the first election under the new delimitation.