Paving the Road to Re-election

Camille Boudot-Reddy* André Butler†

12th September 2023

Abstract

The prevailing view in the economic literature is that voters are particularly myopic, encouraging governments to leverage short-term re-election strategies. Under such conditions, public capital investment with long-term rewards – despite its central role in the process of sustained economic development – may be neglected. In the context of India's rural road construction programme, we evaluate the role which large-scale public infrastructure initiatives have on the electoral accountability mechanism. Using a *fuzzy* regression discontinuity design with newly-digitised village-level voting outcomes from the 2014 general election, we find evidence of electoral support attributed to the political alliance which spearheaded the programme. This support is sustained over two electoral cycles, with significant spillover effects in villages within 2 km of a newly built road. These political gains however, appear to be confined to incumbent candidates.

JEL Codes: H54, I38, O18, P43
Keywords: public infrastructure, voting, roads, India

*Birkbeck, University of London (c.boudot-reddy@bbk.ac.uk)
†Corresponding Author: University of Cambridge (ajb385@cam.ac.uk)

We are grateful to Manuel Bagues, Liang Bai, Antonella Bancalari, Maia Güell, and Tinghua Yu for helpful comments and conversations. We would also like to thank seminar participants at Birkbeck University of London, for their feedback. André Butler would like to thank the Royal Economic Society for their Academic Support Fund, which was used to compile the data for this paper. Khushboo Gupta, Mahantesh B S, and Alexis Reddy provided excellent research assistance.

1 Introduction

In 2017, 1 billion people lived more than 2 km from a paved road, 1.2 billion had no electricity, and 1 in 3 people lacked access to clean drinking water (UNICEF, 2020; World Bank Group, 2016, 2017). Analysis by the McKinsey Global Institute suggests that the world needs to invest an average of $3.3 trillion annually just to sustain current economic growth rates. Emerging economies, with the largest unexploited efficiency gains from greater infrastructure investment, account for 60% of that need (Gardner and Henry, 2021; Institute, ed, 2016). Despite the evident necessity to expand public infrastructure the scheduled investment to take place globally from 2015 to 2030 is only $2.2 trillion, resulting in what has been coined as the "global infrastructure gap" (Gardner and Henry, 2021). The role of financial support has received significant attention in the literature as a constraint to investment (Engel et al., 2022; Fay et al., 2021). In contrast, limited research has been devoted to understanding the political incentives for large government led infrastructure development initiatives.

Governments at all levels are known to make extensive use of well timed short-run re-election strategies. Evidence from vote-buying (Finan and Schechter, 2012), increasing budget surplus (Brender and Drazen, 2008) and programme expenditures (Brollo and Nannicini, 2012) in election years all suggest that these provide substantial rewards at the ballot box. This phenomenon may lead governments to turn towards short-run policies with immediate electoral returns at the expense of programmes whose benefits may only accrue in the medium to long-run. For instance, infrastructure development including transport, water, and electricity, may under such a democratic accountability mechanism receive sub-optimal investment.

In this paper, we provide causal evidence on the effect of exposure to a rural road building programme implemented by the Government of India – which brought paved roads to nearly 200,000 villages between 2000 to 2014 at a cost of almost $40 billion – on voting behaviour over two electoral cycles. While programmes aimed at improving public infrastructure have been implemented by governments around the world, evaluating the causal impact of access to these interventions on electoral outcomes is often challenging. This is partly due to the fact that governments are known to leverage their power to target public goods towards specific groups and locations (Banerjee and Somanathan, 2007; Burgess et al., 2015; Lehne et al., 2018; Mahadevan, 2019). In order to address this source of endogeneity and accurately ascertain a causal interpretation, we focus our analysis on

the construction of rural roads under the Pradhan Mantri Gram Sadak Yojana (Prime Minister's Village Road Programme, or PMGSY). A key feature of this programme implementation was to target the location of roads to villages with a population exceeding two discrete thresholds (500 and 1,000). We estimate that this policy guideline causes villages just above these thresholds to be 22 percentage points more likely to have received a road by the time of the 2014 general election. We exploit this exogenous source of variation in centrally planned road construction in a *fuzzy* regression discontinuity design.

We compile data on voting outcomes from close to one million polling-stations across 18 of India's most populated States for the 2014 general election. Using geocomputing techniques, we then match the location of polling-stations to villages. Importantly for our empirical approach this newly-digitised dataset enables us to leverage the *village* as our unit of analysis; corresponding to the administrative unit used by the government to target road construction across the country. We combine this information with data from the Population Censuses of 2001 and 2011, which describe village-level demographics and amenities. Information on road construction under the programme comes from the PMGSY portal.

The rural roads programme was initiated in 2000 with road construction taking-off in 2002. The Indian National Congress (INC) led coalition, the United Progressive Alliance (UPA), which won the 2004 general elections took control of implementing the programme. By the time of the 2014 general election, the UPA had been in power for two five-year mandates and overseen new road construction to nearly 200,000 villages across the country. It is *a priori* ambiguous as to how voters may respond to this public capital investment in rural roads. If content with improved access to rural roads from the PMGSY, they may choose to assign their votes to the incumbent UPA which largely spearheaded the programme. Alternatively, they may place responsibility with individual members of parliament irrespective of political alliance and hence reward their local elected representative. This effect however may be small or even reversed if voters do not perceive any benefits from the programme or believe that it was poorly implemented. In such a case, we may see an increase in the share of votes cast to the Bharatiya Janata Party (BJP) led coalition, the National Democratic Alliance (NDA), which was the main opposition during the period of programme implementation.

We estimate that the PMGSY significantly increased access to roads in treatment villages by approximately 22%. This significantly improved the integration of villages to the wider economy by increasing access to public bus services and private autorickshaws

by 15% and 13% respectively. Corroborating the work of Asher and Novosad (2020) – who evaluate the impact of the PMGSY on the structural transformation of the village economy – we find no significant effects on consumption. Yet despite limited observable economic benefits, we find that the direct beneficiaries of the rural road programme reward the political alliance which brought a road to their village. Specifically, a new road causes a statistically significant 7.3% increase in the share of votes allotted to the UPA. This vote dividend appears to be sustained even for roads built in the previous electoral cycle. We also find significant spillover effects in non-beneficiary villages within a 2km radius of a newly constructed road. When considering the voting-age population across both direct beneficiary and spill-over villages, we estimate the cost of the vote to be $507. Interestingly, villagers only appear to reward the UPA when the incumbent candidate stands for re-election. We interpret this as an attribution error by the electorate as incumbency does not alter the probability of receiving a new road. Nonetheless, this result does show a large electoral premium for fielding the same candidate when seeking re-election.

This paper is related to a large body of research dedicated to understanding how voters make decisions. The majority of this work has concentrated on the presence of political budget cycles. Explained using a simple retrospective voting behaviour under asymmetric information (Nordhaus, 1975), this literature suggests that voters use economic conditions immediately preceding elections as a signal of the government's ability (Healy and Lenz, 2014). In an analysis of 350 elections across 74 democracies, Brender and Drazen (2008) find that macroeconomic growth and budgetary surplus in an election year significantly improve the incumbent government's chance at re-election. This phenomenon has been found to be especially salient in the context of developing countries (Brender and Drazen, 2005; Shi and Svensson, 2006). Our paper extends this literature by considering the response of voters to government led infrastructure development initiatives which unlike business cycles, are based on medium to long-run impacts.

Recently, a number of studies have aimed to evaluate the electoral outcomes from large government led poverty-alleviation programmes. Broadly, evidence from conditional cash transfer (CCT) schemes suggest that voters value these initiatives and reward the incumbent.[1] For instance, Manacorda et al. (2011) find that beneficiary households from such a programme in Uruguay are 11 to 14% more likely to vote for the incumbent; with these results persisting even after the programme ends. In an evaluation of the Mexican Pro-

[1]For studies evaluating electoral returns from CCT schemes, refer to Baez et al. (2012) in Colombia, De La O (2013) in Mexico, Manacorda et al. (2011) in Uruguay, and Labonne (2013) in Philippines.

gresa programme, De La O (2013) find that participation mobilised beneficiaries to vote and reward the incumbent. Set apart from the context of CCTs, Zimmermann (2021) investigates electoral gains from India's public-works programme and finds that voter support declines with length of programme exposure as the electorate hold the government accountable for the quality of implementation. In this paper, we consider a different type of government led anti-poverty programme, that of infrastructure development. Importantly, unlike CCTs and the public-works programme, large-scale infrastructure initiatives do not involve hand-outs and hence may be harder for individuals to perceive the direct benefits and/or attribute these to the government. Recent work by Akbulut-Yuksel et al. (2023) evaluating the electoral gains to expressway construction in Turkey suggest that this infrastructure initiative significantly increased the vote share to the ruling party. Our study furthers this investigation by considering the electoral gains both over time as well as spatial spill-overs into non-beneficiary villages.

Finally, our paper fits closely with a growing literature documenting the value of democracies in driving economic growth (Acemoglu et al., 2019).[2] Specifically, in a global analysis of electrification using satellite imagery of night light, Min (2008) demonstrates that democratisation leads to a significant reduction in unelectrified population centres. In this paper, we seek to verify whether voters reward governments for promoting public capital investment. Given the mounting evidence highlighting the value of infrastructure in improving economic outcomes – construction of bridges in Nicaragua improved labour market participation (Brooks and Donovan, 2020), access to rural roads in India increased investment in human capital (Adukia et al., 2020), irrigation infrastructure through dams and canal networks significantly reduced poverty levels and promoted structural change (Duflo and Pande, 2007; Asher et al., 2022; Blakeslee et al., 2021) – there is a clear need to link the role which public infrastructure investment plays in the electoral accountability mechanism.

The rest of the paper is structured as follows. Section 2 describes the electoral system in India, as well as the policy implementation of the rural road building programme. Our data sources are explained in Section 3 and the empirical strategy including graphical evidence is presented in Section 4. Section 5 contains results on the impact of rural roads on the political economy. Finally, Section 6 concludes.

[2]Papers investigating the link between democracy and growth have focused on the role of the democratic process on growth (Acemoglu et al., 2019), public infrastructure investment (Min, 2008), curbing corruption (Ferraz and Finan, 2011), and policy choices (Besley et al., 2010).

2 Background

2.1 India's Political System and the 2014 General Election

India's electoral system is a first-past-the-post: in each parliamentary constituency the candidate with the most votes wins the seat within the lower house known as the *Lok Sabha*. The Election Commission of India (ECI) is an independent institution, responsible for ensuring a fair and transparent contest. The ECI has the power to hold political parties to account according to a strict code of conduct in the weeks prior to an election. This electoral code includes stipulations to ensure the incumbent does not hold a disproportionate advantage. For instance, the ECI could prohibit governments from implementing programmes that were not announced prior to the election being called. Furthermore, for the 2014 general election, a limit on election expenditure by an individual candidate was set to Rupees 7 million (equivalent to $88,000) in the largest states and to Rupees 5.4 million (equivalent to $68,000) in smaller states and union territories.

During the 2014 general election of India, there were two main political alliances that were declared before the day of the vote. Firstly, the incumbent United Progressive Alliance (UPA) which had won the two previous general elections (2004 and 2009). The UPA was led by the Indian National Congress (INC) and 10 smaller parties with mostly regional strongholds.[3] The second alliance was that of the main opposition National Democratic Alliance (NDA). The NDA was led by the Bharatiya Janata Party (BJP), alongside 22 smaller parties.[4] Alongside these main alliances, most parliamentary constituency are also contested by smaller parties and independent candidates. A total of 8,251 candidates contested the 543 elected *Lok Sabha* seats.

For administrative and security reasons the election to the 16th *Lok Sabha* was held in nine phases from April 7 to May 12 of 2014. As in each election monitors were randomly assigned to polling stations and only informed of their assignment the day before the

[3]The smaller UPA member parties for the 2014 general election include: Rashtriya Janata Dal, Nationalist Congress Party, Rashtriya Lok Dal, Jharkhand Mukti Morcha, Jammu & Kashmir National Conference, Mahan Dal, Indian Union Muslim League, Socialist Janata, Kerala Congress, Bodoland People's Front.

[4]The smaller NDA member parties for the 2014 general election included: Telugu Desam Party, Shiv Sena, Desiya Murpokku Dravida Kazhagam, Shiromani Akali Dal, Pattali Makkal Katchi, Marumalarchi Dravida Munnetra Kazhagam, Lok Janshakti Party, Rashtriya Lok Samta Party, Apna Dal, Haryana Janhit Congress, Swabhimani Paksha, Indhiya Jananayaga Katchi, Puthiya Needhi Katchi, Kongunadu Makkal Desia Katchi, All India N.R. Congress, Republican Party of India, Rashtriya Samaj Paksha, Revolutionary Socialist Party (Bolshevik), Kerala Congress (Nationalist), National People's Party, Naga People's Front, and Mizo National Front.

election, hence limiting susceptibility to manipulation. On the day of the election ballots were cast using electronic voting machines. The index finger of each voter was marked with indelible ink so as to avoid voter fraud.

There were 834 million registered voters, making it the largest election in history at the time. Approximately 66% of the electorate turned out to vote. The results of the election were announced on May 16. The BJP received 31% of the vote and won 282 seats, while its NDA coalition won a total of 336 seats. It was the first time since 1984 that a party had won enough seats to govern without the support of other parties. In contrast, the INC received only 19.3% of the vote and won 44 seats. The wider INC-led UPA coalition won a total of 59 seats, making it the worst-ever performance of this alliance since independence.

2.2 The Pradhan Mantri Gram Sadak Yojana

In the year 2000 the Indian government launched the Pradhan Mantri Gram Sadak Yojana (PMGSY), also known as the Prime Minister's Village Roads Scheme. Primarily, the purpose of this programme was to provide all-weather paved roads to unconnected villages across India. In practice however, the programme also upgraded low quality paved roads to already connected villages (see Figure A1 for a summary of road completion, disaggregated by new and upgraded roads, under the PMGSY between 2000 to 2014). Although the scheme was initiated under the BJP-led NDA alliance, the majority of roads were completed between 2004 and 2014 during which the INC-led UPA alliance had a majority rule of the *Lok Sabha*. During this period over 400,000 kilometres of road were constructed. The initiative benefited 185,000 villages of which 107,000 – encompassing a population of over 30 million people – had previously lacked an all-weather road. Funded by a combination of income from taxes, central government support, and loans from the Asian Development Bank and World Bank, the programme cost almost $40 billion.

Overseen by the federal Ministry of Rural Development, guidelines were issued by the National Rural Roads Development Authority on the selection of villages due to benefit from this programme. Using arbitrary thresholds based on the 2001 Population Census of India, construction of new roads were first targeted to villages with a population greater than 1000, followed by villages with a population greater than 500, and finally villages with a population greater than 250.[5] Implementation of the PMGSY was delegated to State

[5]The population thresholds were in fact set at the habitation level; a unit of aggregation below that of the village. However in practice habitation populations were pooled to the village level. We aggregate to the village level in order to closely match the implementation of the programme and because this aligns

governments allowing them to adapt these guidelines based on their specific requirements. For example, a State that had already connected all villages at a given target population could proceed directly to a smaller population threshold. Furthermore, the prioritisation guidelines also included the scope for practical concerns. For instance, smaller villages could be connected if they lay in the path of a priority village and groups of villages within 500 meters of each other could combine their populations to attain the eligibility thresholds.[6] As such, while a village's population relative to the threshold would have significantly influenced the probability of receiving a new road, the programme allocation was not definitive.

3 Data

In order to explore whether exposure to public infrastructure investment plays a role in the electoral accountability mechanism, we have digitised polling-station data on voting outcomes from India's 2014 general election. We combine this information with contemporaneous data on road construction under the PMGSY, as well as other features of the rural economy at the village-level. On average we observe our outcome variables six years following road completion, hence capturing the short to medium term impact of benefiting from the programme. This section describes in detail how we compiled the dataset, while Table 1 provides summary statistics describing the average village in our final sample prior to the onset of the PMGSY initiative.

3.1 Voting

Data on votes to the federal government for the 2014 general election come from the Election Commission of India (ECI). Documents – known as a Form-20 (see Figure A2 for an example) – are made publicly available by the Commission for each parliamentary constituency.[7] These documents contain information on the name of all candidates, their party affiliation, and the number of votes they received at each polling-station.[8] Unfortu-

with the data on all other outcomes of interest.

[6]For further details on the guidelines of this programme, refer to National Rural Roads Development Agency (2005).

[7]Links to the Form-20 documents for each State can be found here: https://eci.gov.in/statistical-report/link-to-form-20/

[8]There is also some information on the background of the candidate including gender, age, and caste. However this level of detail is not complete across all constituencies.

nately, information on the number of eligible voters is largely missing and hence we cannot include voter turnout as an outcome of interest in our analysis. We digitised data from all the Form-20s for eighteen of India's largest States; accounting for approximately 89% of the population.[9][10] This creates a dataset covering close to 800,000 polling-stations.

Since variation induced by the PMGSY programme is across villages it is essential to capture all voting outcomes at this aggregation level. We leverage the work of Susewind (2014) to obtain Global Position System (GPS) coordinates of all the polling-stations active during the 2014 general election.[11] A recent evaluation of whether resource scarcity enhances the scope for targeted spending in India by Mahadevan and Shenoy (2023) also uses these coordinates to aggregate voting data at the Gram Panchayat level.

Combining village boundary shapefiles offered by the Socioeconomic High-resolution Rural-Urban Geographic (SHRUG) Dataset on India (Asher et al., 2021), along with the GPS coordinates of polling-stations, we create a village-level match.[12] Specifically, we attribute vote counts to a village if the polling-station falls within the village boundary. For villages without polling-stations, we make use of a nearest neighbour assignment method and assign votes from a polling-station within a 2 km radius of a village. This criteria corresponds closely to the ECI guidelines stating that no voter should be more than 2 km away from a polling station. Furthermore, in a robustness test we show that our results are consistent when limiting the sample to villages with a polling-station within their boundary. Finally, we sum all the votes assigned to a village. We calculate our outcome variables to capture the share of votes cast to the main political parties, as well as the parliamentary constituency incumbent.

New road construction under the PMGSY programme began in the year 2000. Therefore, after receiving a new road voters were potentially exposed to two general elections

[9]Data from searchable PDFs was extracted using R – a software environment for statistical computing and graphics. Information from non-searchable PDFs was extracted using Transkribus – a platform for text recognition, image analysis, and structure recognition of historical documents. In the case of poor quality PDFs which could not be analysed using Transkribus, the data was manually digitised.

[10]The eighteen States for which voting data was digitised include: Andhra Pradesh, Assam, Bihar, Chhattisgarh, Gujarat, Haryana, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttar Pradesh, and West Bengal.

[11]GIS shapefiles for India's parliamentary constituencies including polling-station location co-ordinates are available from: https://pub.uni-bielefeld.de/record/2674065. We match the data from the Form-20 to the GPS coordinates based on a polling station number. To ensure that the match is as precise as possible, we drop data from parliamentary constituencies where the total number of polling stations is inconsistent between the two datasets (the number of polling booths matched in 80% of cases).

[12]The SHRUG datasets, codebooks, and references, can be found at: http://www.devdatalab.org/shrug. In this paper, we made use of the SHRUG, Version 2.0.

(in 2004 and 2009) before our observed voting outcome in the 2014 general election. Unfortunately, to the best of our knowledge, geo-located polling booth data for the 2004 and 2009 elections which would allow us to aggregate voting outcomes at the village level is not available. However in order to explore whether shifts in voting behaviour due to the programme are persistent over multiple elections or transitory, we conduct a heterogeneity analysis by splitting our treatment sample into villages that received a road before and after the 2009 general election.

3.2 Consumption

In order to investigate whether the rural roads programme had an effect on the local economy we consider a range of indicators aimed at capturing shifts in consumption. Night light, measured by satellites as the pixel luminosity in a geographic polygon, is now widely used as a proxy for economic activity when direct measures are otherwise unavailable (Henderson et al., 2011). We calculate the average of total night light for a village over a three-year period, from 2011 to 2013 (see Appendix B for a detailed discussion on the construction and use of this variable in economic studies). As a more direct measure, we leverage predicted consumption and poverty rate at the village-level, imputed using household micro-data collected by the India Human Development Survey-II in 2012 (refer to Appendix B for further detail on how these indicators are predicted and a discussion of the literature on the methodology). Finally, we look at an index of asset ownership as well as each individual major asset independently, as recorded by the Socio Economic Caste Census of India in 2012. Each of these indicator is available on the SHRUG.

3.3 Rural Roads

Information on rural road construction under the PMGSY programme was first scraped by Asher and Novosad (2020) from the official PMGSY portal in order to evaluate the impact of rural roads on the local development of the village economy.[13] We leverage this information, made publicly available by the authors on the SHRUG Dataset on India in order to identify villages treated by the programme. For each village connected by a new paved road the dataset also details the date for when the contract was awarded, as well as the date of completion, disaggregated by new versus upgraded roads. We are primarily

[13]Information on the programme guidelines, reports, and identities of newly connected villages can also be accessed from the official PMGSY portal: http://omms.nic.in/

interested in capturing the effect of benefiting from a *new* road, hence we limit our sample to villages that did not have a paved road in 2001. While evaluating the effect from a road upgrade would have also been interesting, we find that the construction of these did not strictly follow the policy rule to generate a discontinuity at the eligibility threshold and therefore are not included in our analysis.[14]

The PMGSY programme was initiated by the NDA alliance in 2000. This government however lost the 2004 elections and the majority of roads were therefore completed under the leadership of the UPA alliance between 2004 to 2014.[15] Consequently, we exclude villages that received a new road prior to 2005. Figure 1 reports the number of new roads built under the PMGSY between 2004 to 2014 among our sample of villages. By the 2014 general election of India, treated villages would have had access to a new road for an average of six years.

So as to verify whether the programme had an effect on transportation, we complement the information on road construction with data on road usage. Specifically, we consider the regular availability of transport services at the village by 2011. Recorded as part of the Population Census of India and made available on the SHRUG, we can measure the presence of buses (both private and public), vans, taxis, and auto-rickshaws.

3.4 Population

The Population Census of India, compiled as part of the SHRUG, provides information on village demographics. We use data from the 2001 Census to capture our assignment variable – village population. This enables us to replicate the precise policy rule used by the Government of India when targeting the construction of new rural roads across the country as part of the PMGSY programme. Accordingly, we consider villages with a population greater than the stipulated threshold as *treated* – that is, prioritised to receive a new road.

Asher and Novosad (2020) collaborated closely with the National Rural Roads Development Agency to identify State specific compliance with the PMGSY guidelines. Following their precedent, we restrict our sample for analysis to the six States that strictly adhered to the population thresholds (in parentheses) stipulated by the PMGSY: Chhattisgarh (500, 1000), Gujarat (500), Madhya Pradesh (500, 1000), Maharashtra (500), Orissa (500), and Rajasthan (500). Furthermore, we limit the sample of villages to those that had popula-

[14] Results from these tests are not reported in the manuscript, but are available on request.
[15] Figure A1 documents all roads built since 2000.

tions within the optimal bandwidth (84) from a treatment threshold. This gives us a final sample of 10,431 villages with non-missing information across all our outcomes of interest.

4 Empirical Approach

Estimating the impact of large-scale public infrastructure initiatives on voting behaviour is challenging due to concerns of endogeneity. Given the large budgets required for such programmes, the decision of where to allocate the investment is unlikely to be random. For instance, public goods may be targeted to the most productive villages. Or conversely, towards villages that are lagging economically. Any naive correlation estimates between public infrastructure investment and voting behaviour will therefore be an over or underestimate of the true effects. In order to identify exogenous variation in public capital investment, we exploit the Government of India PMGSY programme implementation guidelines aimed explicitly at expanding rural road construction across the country based on arbitrary population eligibility thresholds. This section details our proposed empirical approach – *fuzzy* Regression Discontinuity (RD) Design, alongside graphical evidence and estimation results corroborating the validity of this method.

4.1 Regression Discontinuity Design

The PMGSY programme used arbitrary village population thresholds to prioritise road construction (described in Section 2). These eligibility rules however, were not definitive. As such, we employ a $fuzzy$ RD design to estimate a change in voting behaviour caused by exposure to a new road built under the PMGSY government initiative. Specifically, we use the following two stage least squares specification with optimal bandwidth local linear regression (Gelman and Imbens, 2019; Calonico et al., 2020):

$$\begin{aligned} Road_{vds} = \gamma_0 + \gamma_1(pop_{vds} \geq T_s) + \gamma_2(pop_{vds} - T_s) \\ + \gamma_3(pop_{vds} - T_s).(pop_{vds} \geq T_s) + \nu X_{vds} + \mu_{dh} + \upsilon_{vds} \end{aligned} \tag{1}$$

$$\begin{aligned} Y_{vds} = \beta_0 + \beta_1 Road_{vds} + \beta_2(pop_{vds} - T_s) \\ + \beta_3(pop_{vds} - T_s).(pop_{vds} \geq T_s) + \sigma X_{vds} + \eta_{dh} + \varepsilon_{vds} \end{aligned} \tag{2}$$

Y_{vds} is the outcome of interest for village v in district d in state s. $Road_{vds}$ is a binary indicator which takes the value one if village v received a new road as part of the PMGSY programme before the year in which Y_{vsd} is measured (2013 for the voting outcomes and 2011 for consumption and transportation services). pop_{vds} is the 2001 village population, our assignment variable. T_s is the population eligibility threshold used by state s.[16] β_1 captures the average effect of receiving a new road on the outcome variable. All regressions use an optimal bandwidth of 84, calculated using a triangular kernel which places more weight on observations close to the threshold (Calonico et al., 2020). In a robustness test, we demonstrate that the results are consistent across a range of bandwidths and alternative kernels.

Control variables and fixed effects are not necessary for identification in an RD design, but do improve the efficiency of the estimation (Calonico et al., 2014; Imbens and Lemieux, 2008). We therefore include a vector of baseline village covariates – X_{vds} – as controls in our specification. Specifically we control for: village amenities (primary school, medical centre, electrification, distance from the closest urban centre), agricultural sector characteristics (total agricultural land area and share of agricultural land irrigated), and socio-economic indicators (share of the village population that are literate, belonging to a scheduled caste, own land, rely on subsistence farming, and with a HH income above Rs.250/month). Furthermore, we also include parliamentary constituency-population threshold fixed effects – μ_{dh} in Equation 1 and η_{dh} in Equation 2 – which are an interaction of constituency dummies with an indicator variable that takes the value one if village v is in a state where the highest population treatment threshold (1,000) was used.[17] In a robustness test, we show that the results are consistent when excluding these controls and fixed effects.

[16] As described in Section 2, while the PMGSY first targeted villages with a population above 1000, states were allowed use the lower threshold of 500 immediately if they had few villages of population over 1000 with no roads. Among our sample of the six states that closely complied to the programme, the thresholds used are: Chhattisgarh – 500 and 1,000, Gujarat – 500, Madhya Pradesh – 500 and 1,000, Maharashtra – 500, Orissa – 500, and Rajasthan – 500. Since the optimal bandwidth is less than 100, there is no overlap between the group of villages receiving roads under the 500 or 1000 threshold. This enables us to pool villages according to the population thresholds applied in each state. The lowest population eligibility threshold stipulated by the PMGSY guidelines was that of 250. However few villages of this group had received a road by 2014, hence we do not include this threshold in our analysis.

[17] Using district identifiers instead of parliamentary constituencies do not change our main results.

4.2 Impact of the PMGSY on Rural Roads

Causal inference in a *fuzzy* RD design is conditional on meeting three key assumptions. First, the conditional density of the assignment variable is continuously differentiable at the threshold. Second, baseline covariates are also balanced with conditional density continuously differentiable at the threshold. Third, there is a jump in the direct marginal effect of the treatment on the assignment variable at the threshold.

The first assumption is concerned with ruling out the possibility that villages could somehow manipulate their position at the threshold so as to be eligible for the programme. By plotting the density distribution of the village population in 2001 normalised at the eligibility threshold, Figure 2 demonstrates that there is no discontinuity in our assignment variable. This is formally corroborated by the McCrary test which estimates the log change in height between bins at the threshold and confirms that we cannot detect a significant discontinuity at that point (statistic of 0.04 with s.e. 0.044).

The second assumption attempts to address the concern that there may be village characteristics which are correlated to the treatment status. Table 1 presents summary statistics on village characteristics at baseline for our full sample (Column 1), as well as disaggregated for villages just below (Column 2) and those just above (Column 3) the threshold. We report the sample means for all our control variables which capture information on village amenities, the agricultural sector, and demographics (including socio-economic indicators). While there are average differences between villages below and above the threshold (Columns 4 and 5), we find no statistically significant discontinuity at that point when using the *fuzzy* RD specification (Columns 6 and 7). By plotting the relationship between the baseline control variables and the village population in 2001 normalised at the eligibility threshold, Figure 3 provides graphical evidence demonstrating no discontinuous changes across the treatment cut-off.

Finally the third assumption validates the treatment effect at the arbitrarily stipulated threshold. In Figure 4 we plot the share of villages that received a new road between 2004 and 2014 in each population bin normalised at the eligibility threshold. There is a clear substantial jump in the probability of being treated – receiving a new road – for villages with a population just above the threshold. This graphical evidence is further substantiated in Panel A of Table 2 which presents our first stage results from Equation 1 under a range of bandwidths from the programme population threshold. Our results suggest that crossing the stipulated PMGSY eligibility rule increases the probability of a

village receiving a new road by on average 22%, with a large amount of consistency between bandwidths. This result also holds among villages treated between 2009-2014 as well as those treated between 2004-2009, reported in Panel B and C respectively.

5 Results

5.1 Transport and the village economy

We begin by investigating whether India's national rural road construction programme changed provision of transportation services to the village – the most immediate benefit of being connected to the national road network. Table 3 presents regression discontinuity estimates on the impact of a new road on the availability of five major motorised transportation services, as recorded in the 2011 Population Census. For our full sample, reported in Panel A, we find that a new road causes a statistically significant 15% increase in the availability of public bus services; more than double that of the control group mean (11.7%) (Column 1). However, when we consider the timing of new road construction in Panel B and C, we see that increases in public bus services only manifests itself more than three years after the new roads are built. That is, we only estimate a statistically significant increase in public bus services in villages where the new road was built between 2004 and 2009. In contrast, there is no effect on provision of private bus services for the full sample (Column 2), however there is some marginal evidence that they do increase in the long run. We estimate a 17.9% increase in public bus services for roads built between 2004 and 2009, which is only significant at the 10% level. In terms of the other private transport services, while we do not find any significant shifts in the availability of taxis (Columns 3), there is evidence of an immediate substitution away from vans that may be better suited to unpaved roads (Column 4). Specifically, we estimate a large 42.7% drop in private vans in villages where a new road is less than 3 years old. Furthermore, we estimate a significant increase of 13.7% in the availability of autorickshaws – the least expensive form of private transport (Column 5). This effect is largest (28.9%) immediately after new roads are completed and before we see the increase in availability of more affordable public bus services. These results provide evidence that villages treated by the PMGSY initiative witnessed a consequential improvement in their integration to the wider economy.

Next, we examine whether access to a new road made people in these villages economically better off. Table 4 reports regression discontinuity estimates on four indicators of

consumption. Our results consistently suggest that there are no improvements in predicted consumption at the village-level. We estimate a statistically insignificant 1.8% increase in predicted consumption per capita; and can rule out a greater than 10% increase with 95% confidence (Column 1). Similarly, we find a statistically insignificant 1.9% drop in the village poverty rate; and can rule out a 9% decline with 95% confidence (Column 2). Evidence on asset ownership and night light are also not statistically significant and indicate well estimated null effects (Columns 3 and 4 respectively).[18] It would appear that a new road, despite providing a link to increased market opportunities, does not on average substantially improve the economic status of the population. Our results replicate those of Asher and Novosad (2020). Evaluating the impact of the PMGSY programme on the economic development of the village, the authors find that a new road does not appear to translate into any substantial improvements either on the aggregate economy or predicted consumption in the short to medium term.

5.2 Voting outcomes

How exposure to the PMGSY programme and access to a new road affects voter decision at the ballot box is *a priori* ambiguous. The lack of a meaningful economic impact – if real and not a ramification of the statistical power in our estimation – may mean that voters do not perceive any benefits from the scheme and hence do not incorporate this in their accountability mechanism. Alternatively, they may believe that the lack of any improvement in their status is the result of poor implementation and hence punish the incumbent government for a failed programme. Conversely, villagers may accrue benefits on other dimensions of welfare which are not captured by broad consumption indicators. Those in treated villages may simply value being connected to the national road network. In this context, voters may decide to reward the government for this new public good.

Table 5 presents regression discontinuity estimates on the impact of a new road on votes cast to political parties contesting India's 2014 general election. While the PMGSY initiative was launched in 2001 under the guidance of the NDA government at the time, the bulk of road construction took place under the leadership of the UPA which maintained a majority rule of the *Lok Sabha* from 2004 to 2014 (see Figure 1 for a summary of road completion alongside the timing of general elections). We find that a new road brought by

[18] In Table A1 we present regression discontinuity estimates on individual components of the asset index and find no statistically significant effects on the share of households owning these assets.

the PMGSY scheme causes a 6.5% increase in the share of votes allotted to the UPA leader – the INC – which is statistically significant that the 10% level (Panel A, Column 1). This increases marginally to a 7.3% advantage compared to the control group when considering all parties affiliated to the alliance (Panel A, Column 2; the discontinuity is represented graphically in Panel A of Figure 5). Interestingly, these shifts in voter behaviour are seen in villages where the road was built before the 2009 general election, wherein the electorate had already had an opportunity to express their views at the ballot box (Panel C, Columns 1 and 2). While the point estimates are similar for roads built during the current electoral cycle (Panel B, Columns 1 and 2) they are less precisely estimated, consequently we can not rule out a null effect with 90% confidence. Taken together, these results suggest that the electoral gains of public infrastructure investment are persistent over multiple elections or the electorate only express their views at the ballot box after sufficient time as elapsed for them to experience the real economic benefits of a new road. Finally, in the case of the NDA opposition (Column 4), as well as small parties and independent candidates (Column 5), we estimate small and insignificant decreases in their share of votes regardless of when the road was built. These results are consistent with recent work finding that voters are sensitive to government led poverty-alleviation schemes such as conditional cash transfers (Manacorda et al., 2011) and public-works programmes (Zimmermann, 2021) when evaluating the incumbent government at the ballot box.

5.3 Spillover effects

In this section we consider the possibility that roads do not only affect the voting behaviour in the beneficiary village but could also have significant spillover effects in nearby villages that benefit from being better connected to their local population centres. To do so, we examine voting outcomes in villages that did not receive a new road within a 1-4 km radius of a beneficiary village using the standard regression discontinuity specification (Equations 1 & 2). The point estimates and confidence intervals for these spillover effects are plotted in Figure 6. We can clearly see that the spillover voting effects in villages within 2 km are of a similar magnitude to those seen in the beneficiary villages (signified by zero km on the coefficient plot) and begin to fade away for villages further away. Based on this evidence Table 6 presents the spillover effects for villages within 2 km of beneficiary villages split by which election cycle the roads were built in. For roads built in the current electoral cycle we estimate an 11% increase in the share of votes to both the INC and its

wider alliance the UPA, both significant at the 10% level (Panel B, Columns 1 and 2). We also find that similar benefits persist for roads that were built in the previous electoral cycle (Panel C, Columns 1 and 2).

Given these large spillover effects our estimates of direct effects on beneficiary villages in Table 5 could be biased downwards relative to the total effects of new road provision. In order to explore this possibility we revisit our estimates of the direct effects on beneficiary villages excluding villages within a 2km radius from the control group. These results are presented in Table 7. We find that the point estimates for the share of votes to the INC and UPA were underestimated by 1.4 and 1 percentage points respectively (Panel A, Columns 1 and 2). Similar underestimates are seen for the coefficients when conditioning on the timing of road construction in Panels B and C. As a result we now estimate a statistically significant (at the 10% level) increase in the vote share to the UPA of 10.8%, when conditioning on roads built within the current electoral cycle (2009-2014). Furthermore, for roads built between 2004 and 2009, we also estimate a decline in the vote share to the BJP and NDA of 9.1% and 8.7% respectively, which are significant at the 10% level (Panel C, Columns 3 and 4). This decline roughly mirrors the gain in votes to the INC and UPA suggesting that marginal voters are shifting their vote from the BJP led alliance to the INC associated parties.

5.4 Incumbency effects

Under the PMGSY programme guidelines the award of new roads is based largely on an arbitrary population eligibility threshold and therefore should not be under the influence of individual members of parliament. Nonetheless, previous research suggests that voters have a tendency to vote for the status quo in good economic times (Bagues and Esteve-Volart, 2016). This phenomenon could create an electoral premium for incumbent candidates in locations where roads were build, despite the fact they should have had little to no influence over the decision. If this were the case, the ruling party may attempt to capitalise on this premium by fielding more incumbent candidates in locations where roads were built. To test this hypothesis we check for a discontinuity in incumbency at the population threshold. We find no evidence of targeting incumbency to capitalise on the PMGSY programme (point estimate 0.035, and standard error 0.061).

Having ruled out the possibility of political profiteering we test to see if the electorate do indeed erroneously reward the incumbent candidate. Table 8 presents regression dis-

continuity estimates on the impact of road construction on votes cast to the parliamentary constituency incumbent. Among our complete sample of villages treated between 2004 to 2014, reported in Panel A, we estimate that a new PMGSY road causes a statistically significant increase of 15.9% in the share of votes to an incumbent candidate standing for re-election (Column 1). Interestingly, a new candidate representing the incumbent party is not rewarded for the programme (Column 2). This result is consistent when considering voting behaviour over the current election cycle (Panel B, Columns 1 and 2) as well as the previous election (Panel C, Columns 1 and 2). We then disaggregate incumbency by political affiliation and find that the vote premium is targeted principally to the political alliance which delivered the programme. Having received a road between 2004 to 2014 increased the share of votes to a UPA incumbent candidate standing for re-election by 17.4% (Panel A, Column 3; the discontinuity is represented graphically in Panel D of Figure 5). Conversely, there is no effect on the vote share to a new UPA candidate (Panel A, Column 4)[19] or incumbents representing the opposition NDA coalition (Panel A, Columns 5 and 6)[20]. These results corroborate our previous investigation – Indian voters reward the political alliance which implemented the programme and this appears to be persistent over two election cycles.

5.5 Robustness

We examine the robustness of our results to alternative regression specifications and potential confounding factors. First, we conduct a placebo test by estimating the first stage and reduced form of our key voting outcome variables for a set of States[21] that did not adhere to the policy guidelines, as well as, villages close to the 1,000 threshold in States that used only the 500-person threshold (Gujarat, Maharashtra, Orissa and Rajasthan). Importantly, all these States continued to actively build roads during the study period. As reported in Table A2, we find no evidence of either a first stage or reduced form effects on any outcomes in the placebo sample. These results suggest that our main estimates are not picking up potential confounding factors, such as the electorate expressing an opinion on another existing or proposed policy with a similar eligibility criteria. Second, in Table

[19] Results on this variable for the 2009-2014 election cycle (Panel B, Column 4) should not be interpreted to provide causal estimates due to weak instruments.

[20] Results on this variable for the 2004-2009 election cycle (Panel C, Column 6) should not be interpreted to provide causal estimates due to weak instruments.

[21] These States include: Andhra Pradesh, Assam, Bihar, Jharkhand, Karnataka, Uttar Pradesh, and Uttarakhand.

A3, we present the results on our main voting outcomes for a range of bandwidths (60 to 100) as well as for both triangular and rectangular kernels. The results are very consistent across all specifications. Third, Table A4 reports results on our first stage and main voting outcomes when excluding the vector of baseline covariates and fixed effects. As expected, removing these variables increases the standard errors but does not meaningfully change any of the results. Finally, in Table A5 we show consistency in our estimates when limiting our sample to villages with a polling-station within their boundary.

6 Conclusion

Evidence from the economic voting literature has predominantly painted a picture of myopic voters that place excessive weight on measures immediately prior to elections when evaluating the incumbent. Encouraging governments to leverage short-term re-election strategies can be problematic. Specifically, this phenomenon may lead governments to neglect potentially ambitious public capital investment programmes. In this paper, we seek to investigate whether voters include large-scale infrastructure development initiatives in their electoral accountability mechanism at the ballot box.

In 2000, the Government of India launched its national rural road building programme. Fourteen years later, close to 200,000 villages had received a new paved rural road connecting it to the wider economy. This cost the government a total of $40 billion. An analysis of the medium-term returns to this investment however, do not appear to transform the rural economy (Asher and Novosad, 2020). Nevertheless, we find that the electorate incorporate this programme in their evaluation of the government when casting their vote. Indian voters appear to place significant value on their local public infrastructure and reward the political alliance which initiated the investment while also recognising the role of the local member of parliament in the implementation process.

Importantly, this work allows us to estimate the electoral returns to investment in public infrastructure. Focusing on the direct beneficiaries of the programme the average cost of a new road was $136,806 and the mean voting-age population in beneficiary villages was 1014.[22] We estimate that a new road increased the vote share to the UPA by 8.3% when excluding spillover villages from the control group, making the cost of a single vote $1,626 on average. However, we also estimate a similar increase in the vote share of 8.4% in villages within a 2 km radius of a new road. When we incorporate these into our

[22]The voting age in India is 18 years

calculations by taking the average voting-age population across all villages to be 883 and the mean number of villages within a 2 km radius to be 2.5, this then reduces the average cost of a vote to $507.

References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A Robinson, "Democracy does cause growth," *Journal of political economy*, 2019, *127* (1), 47–100.

Adukia, Anjali, Sam Asher, and Paul Novosad, "Educational investment responses to economic opportunity: evidence from Indian road construction," *American Economic Journal: Applied Economics*, 2020, *12* (1), 348–76.

Akbulut-Yuksel, Mevlude, Dozie Okoye, and Belgi Turan, "Expressway to Votes: Infrastructure Projects and Voter Persuasion," *The Economic Journal*, 2023, p. uead065.

Alkire, Sabina and Suman Seth, "Identifying BPL households: A comparison of methods," *Economic and Political Weekly*, 2013, pp. 49–57.

Asher, Sam, Alison Campion, Douglas Gollin, and Paul Novosad, "The long-run development impacts of agricultural productivity gains: Evidence from irrigation canals in India," *STEG Working Paper*, 2022.

_ **and Paul Novosad**, "Rural roads and local economic development," *American economic review*, 2020, *110* (3), 797–823.

_ **, Tobias Lunt, Ryu Matsuura, and Paul Novosad**, "Development research at high geographic resolution: An analysis of night-lights, firms, and poverty in India using the shrug open data platform," *The World Bank Economic Review*, 2021, *35* (4), 845–871.

Baez, Javier Eduardo, Adriana Camacho, Emily Conover, and Román Zárate, "Conditional cash transfers, political participation, and voting behavior," *World Bank Policy Research Working Paper*, 2012, (6215).

Bagues, Manuel and Berta Esteve-Volart, "Politicians' luck of the draw: Evidence from the Spanish Christmas lottery," *Journal of Political Economy*, 2016, *124* (5), 1269–1294.

Banerjee, Abhijit and Rohini Somanathan, "The political economy of public goods: Some evidence from India," *Journal of development Economics*, 2007, *82* (2), 287–314.

Bedi, Tara, Aline Coudouel, and Kenneth Simler, *More than a pretty picture: Using poverty maps to design better policies and interventions*, World Bank Publications, 2007.

Besley, Timothy, Torsten Persson, and Daniel M Sturm, "Political competition, policy and growth: theory and evidence from the US," *The Review of Economic Studies*, 2010, *77* (4), 1329–1352.

Blakeslee, David, Aaditya Dar, Ram Fishman, Samreen Malik, Heitor Pelegrina, and Karan Singh, "Irrigation and the spatial pattern of structural transformation in India," *Working Paper*, 2021.

Bleakley, Hoyt and Jeffrey Lin, "Portage and path dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Brender, Adi and Allan Drazen, "Political budget cycles in new versus established democracies," *Journal of monetary Economics*, 2005, *52* (7), 1271–1295.

_ and _, "How do budget deficits and economic growth affect reelection prospects? Evidence from a large panel of countries," *American Economic Review*, 2008, *98* (5), 2203–20.

Brollo, Fernanda and Tommaso Nannicini, "Tying your enemy's hands in close races: the politics of federal transfers in Brazil," *American Political Science Review*, 2012, *106* (4), 742–761.

Brooks, Wyatt and Kevin Donovan, "Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua," *Econometrica*, 2020, *88* (5), 1965–1997.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel, "The value of democracy: evidence from road building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–51.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell, "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 2020, *23* (2), 192–210.

_ , _ , and Rocio Titiunik, "Robust data-driven inference in the regression-discontinuity design," *The Stata Journal*, 2014, *14* (4), 909–946.

Duflo, Esther and Rohini Pande, "Dams," *The Quarterly Journal of Economics*, 2007, *122* (2), 601–646.

Elbers, Chris, Jean O Lanjouw, and Peter Lanjouw, "Micro-level estimation of poverty and inequality," *Econometrica*, 2003, *71* (1), 355–364.

Elvidge, Christopher D, Mikhail Zhizhin, Tilottama Ghosh, Feng-Chi Hsu, and Jay Taneja, "Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019," *Remote Sensing*, 2021, *13* (5), 922.

Engel, Eduardo, Ronald Fischer, and Alexander Galetovic, "Private Finance of Public Infrastructure," *Annual Review of Financial Economics*, 2022, *14*, 319–335.

Fay, Marianne, David Martimort, and Stephane Straub, "Funding and financing infrastructure: The joint-use of public and private finance," *Journal of Development Economics*, 2021, *150*, 102629.

Ferraz, Claudio and Frederico Finan, "Electoral accountability and corruption: Evidence from the audits of local governments," *American Economic Review*, 2011, *101* (4), 1274–1311.

Finan, Frederico and Laura Schechter, "Vote-buying and reciprocity," *Econometrica*, 2012, *80* (2), 863–881.

Gardner, Camille and Peter Blair Henry, "The Global Infrastructure Gap: Potential, Perils, and a Framework for Distinction," *Journal of Economic Literature*, 2021.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Harari, Mariaflavia, "Cities in bad shape: Urban geometry in India," *American Economic Review*, 2020, *110* (8), 2377–2421.

Healy, Andrew and Gabriel S Lenz, "Substituting the end for the whole: Why voters respond primarily to the election-year economy," *American Journal of Political Science*, 2014, *58* (1), 31–47.

Henderson, Vernon, Adam Storeygard, and David N Weil, "A bright idea for measuring economic growth," *American Economic Review*, 2011, *101* (3), 194–99.

Hentschel, Jesko, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi, "Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador," *The World Bank Economic Review*, 2000, *14* (1), 147–165.

Hodler, Roland and Paul A Raschky, "Regional favoritism," *The Quarterly Journal of Economics*, 2014, *129* (2), 995–1033.

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

Institute, McKinsey Global, ed., "Bridging Global Infrastructure Gaps," Technical Report, San Francisco 2016.

Labonne, Julien, "The local electoral impacts of conditional cash transfers: Evidence from a field experiment," *Journal of development economics*, 2013, *104*, 73–88.

Lehne, Jonathan, Jacob N Shapiro, and Oliver Vanden Eynde, "Building connections: Political corruption and road construction in India," *Journal of Development Economics*, 2018, *131*, 62–78.

Mahadevan, Meera, "The price of power: Costs of political corruption in Indian electricity," *Department of Economics, University of Michigan*, 2019.

_ **and Ajay Shenoy**, "The political consequences of resource scarcity: Targeted spending in a water-stressed democracy," *Journal of Public Economics*, 2023, *220*, 104842.

Manacorda, Marco, Edward Miguel, and Andrea Vigorito, "Government transfers and political support," *American Economic Journal: Applied Economics*, 2011, *3* (3), 1–28.

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

McKenzie, David J, "Measuring inequality with asset indicators," *Journal of Population Economics*, 2005, *18* (2), 229–260.

Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson, "Night-time light data: A good proxy measure for economic activity?," *PloS one*, 2015, *10* (10), e0139779.

Min, Brian, "Democracy and light: Electoral accountability and the provision of public goods," in "annual meeting of the Midwest Political Science Association" 2008.

National Rural Roads Development Agency, "Pradhan Mantri Gram Sadak Yojana – operations manual," Technical Report 2005.

Nordhaus, William D, "The political business cycle," *The review of economic studies*, 1975, *42* (2), 169–190.

O, Ana L De La, "Do conditional cash transfers affect electoral behavior? Evidence from a randomized experiment in Mexico," *American Journal of Political Science*, 2013, *57* (1), 1–14.

Shi, Min and Jakob Svensson, "Political budget cycles: Do they differ across countries and why?," *Journal of public economics*, 2006, *90* (8-9), 1367–1389.

Susewind, Raphael, "GIS shapefiles for India's parliamentary and assembly constituencies including polling booth localities," 2014.

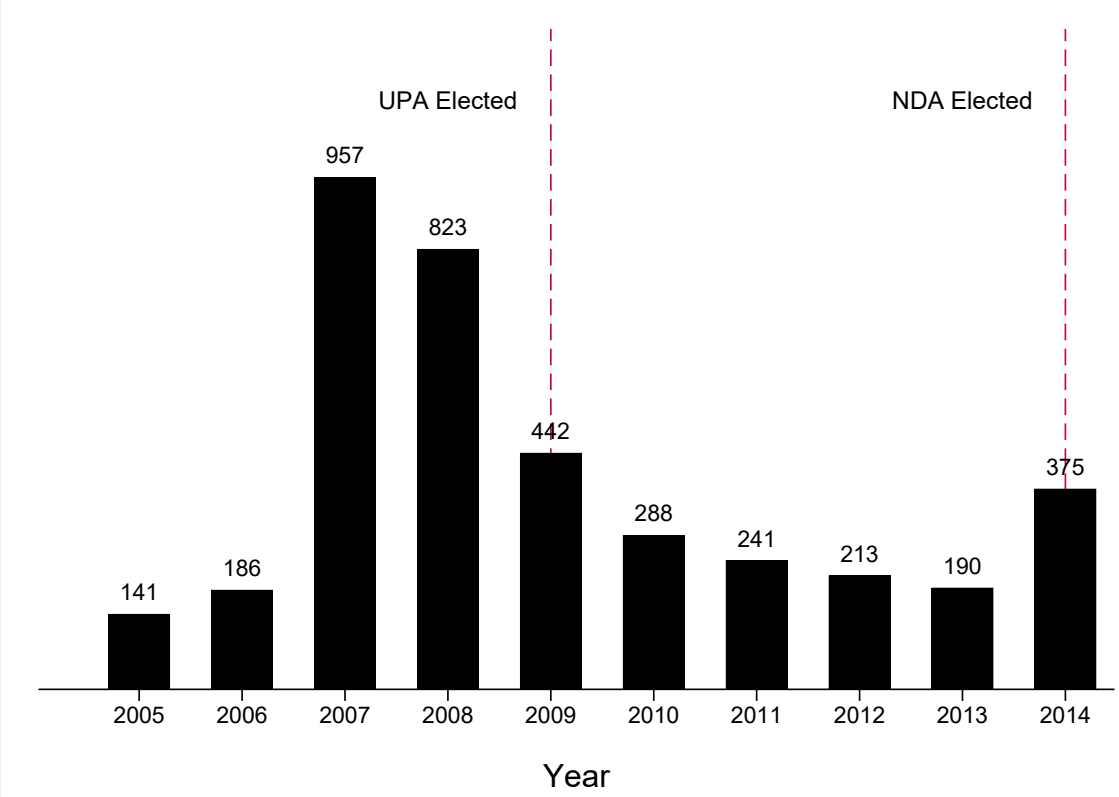
UNICEF, "Progress on drinking water, sanitation and hygiene: 2000-2017: Special focus on inequalities," Technical Report 2020.

World Bank Group, "Measuring rural access: Using new technologies," Technical Report 2016.

_ , "State of Electricity Access Report," Technical Report 2017.

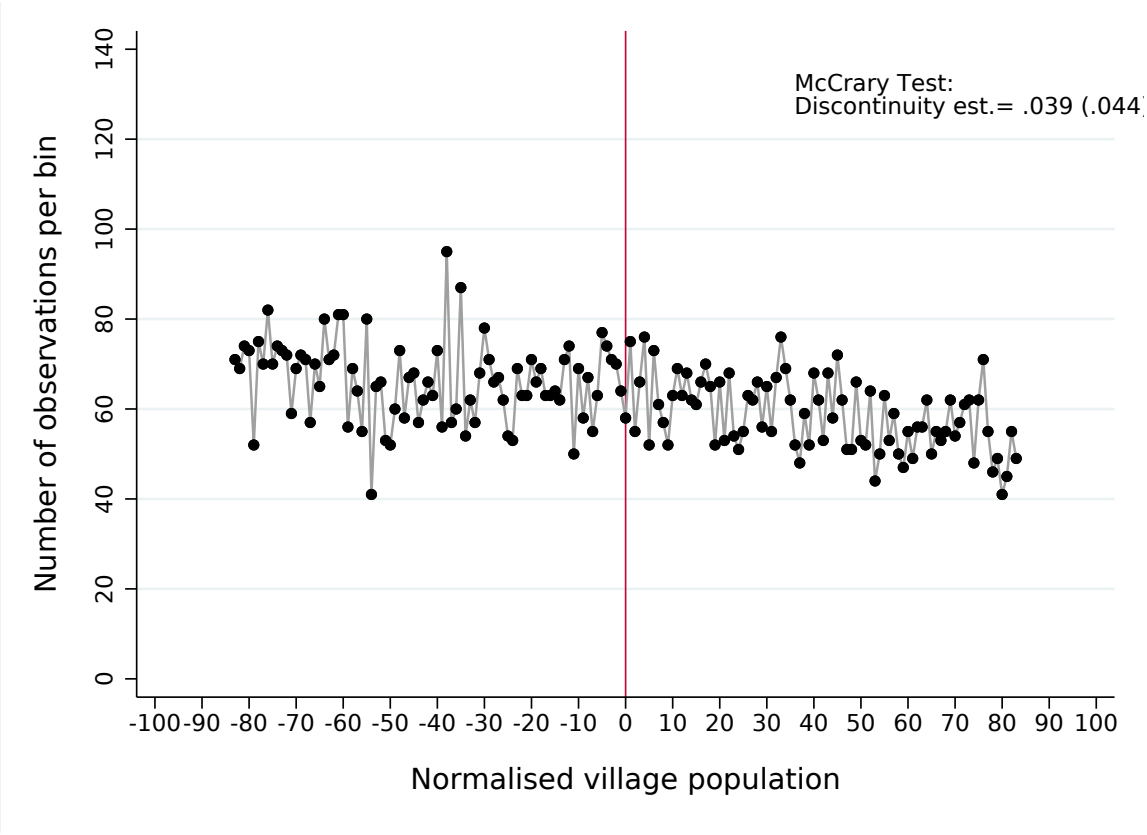
Zimmermann, Laura, "The Dynamic Electoral Returns of a Large Antipoverty Program," *Review of Economics and Statistics*, 2021, *103* (5), 803–817.

Figure 1: Roads completed under the PMGSY by year



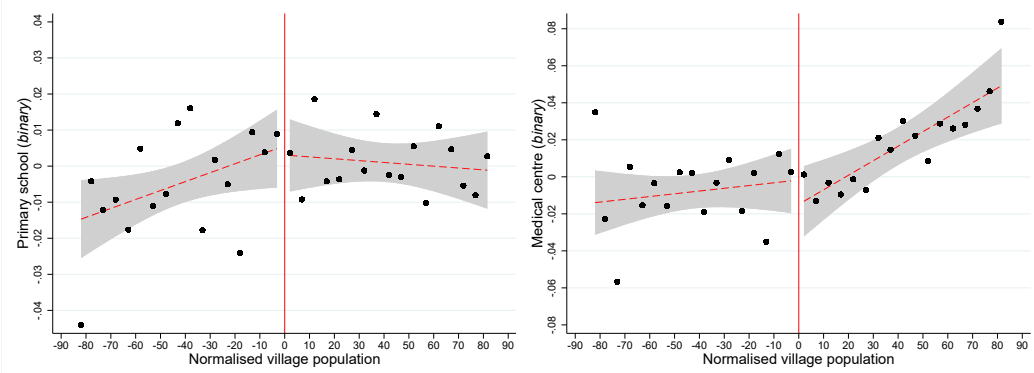
Notes: This figure shows road completion under the PMGSY initiative between 2005 to 2014, as well as the timing of general elections over this period. The bars represent the number of new roads completed under the PMGSY in each year (exact counts are reported above the bars). The dashed lines mark the years in which a general election was conducted, with the winning political coalition reported alongside. Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold).

Figure 2: Distribution of the assignment variable across the threshold



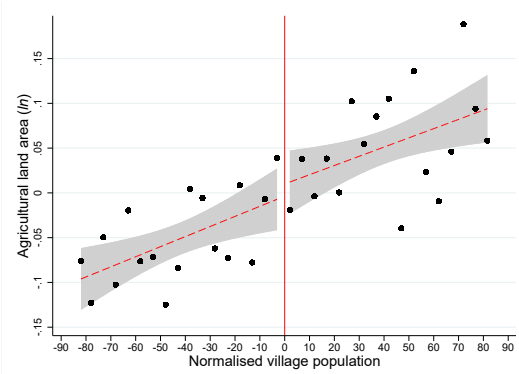
Notes: This figure plots the number of observations in each bin of the assignment variable. The assignment variable – village population – is normalised around the PMGSY eligibility threshold (either 500 or 1,000). The bin size is 0.25. A *fuzzy* RD design requires for the conditional density of the assignment variable to be continuously differentiable at the threshold. Following McCrary (2008), we report the coefficient and standard error for a discontinuity test which estimates the log change in height between bins at the threshold. Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold).

Figure 3: Balance of village characteristics at baseline

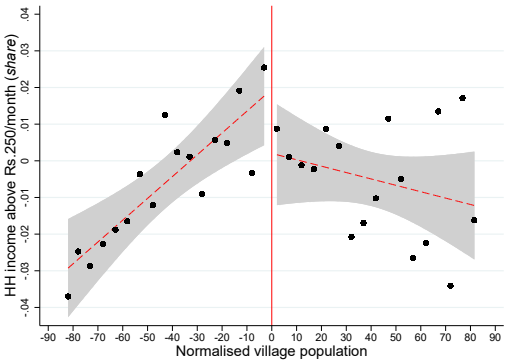


Panel A

Panel B



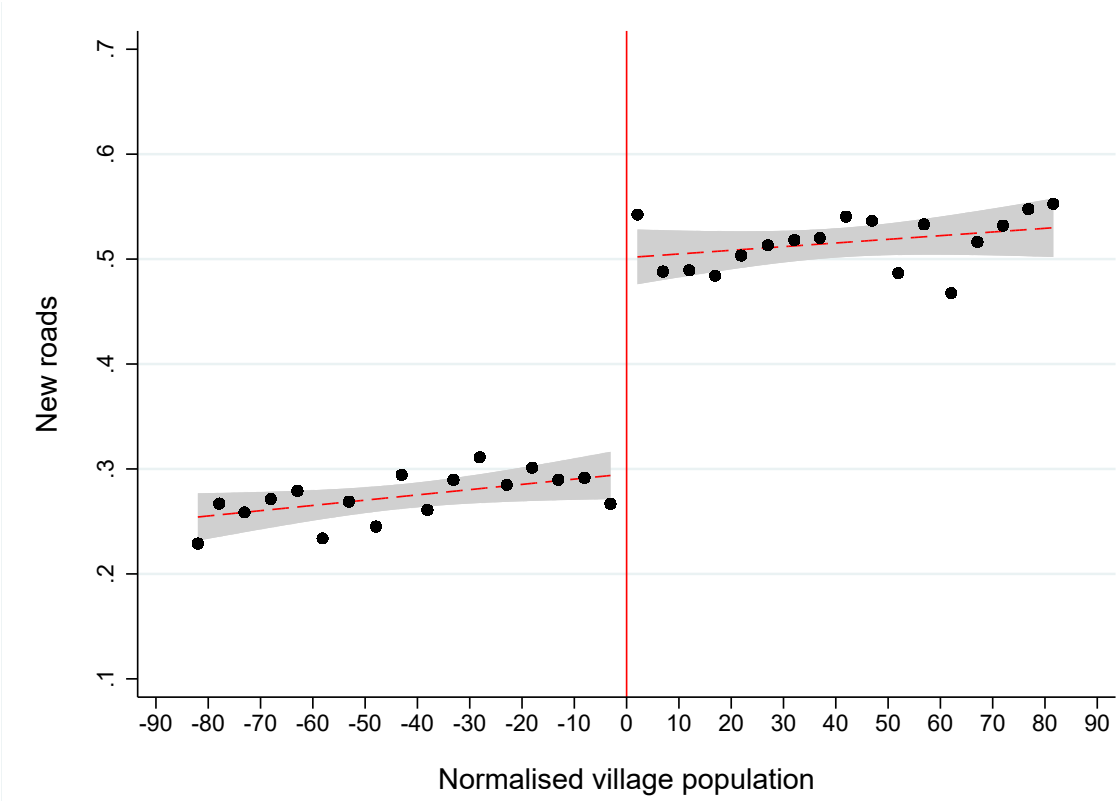
Panel C



Panel D

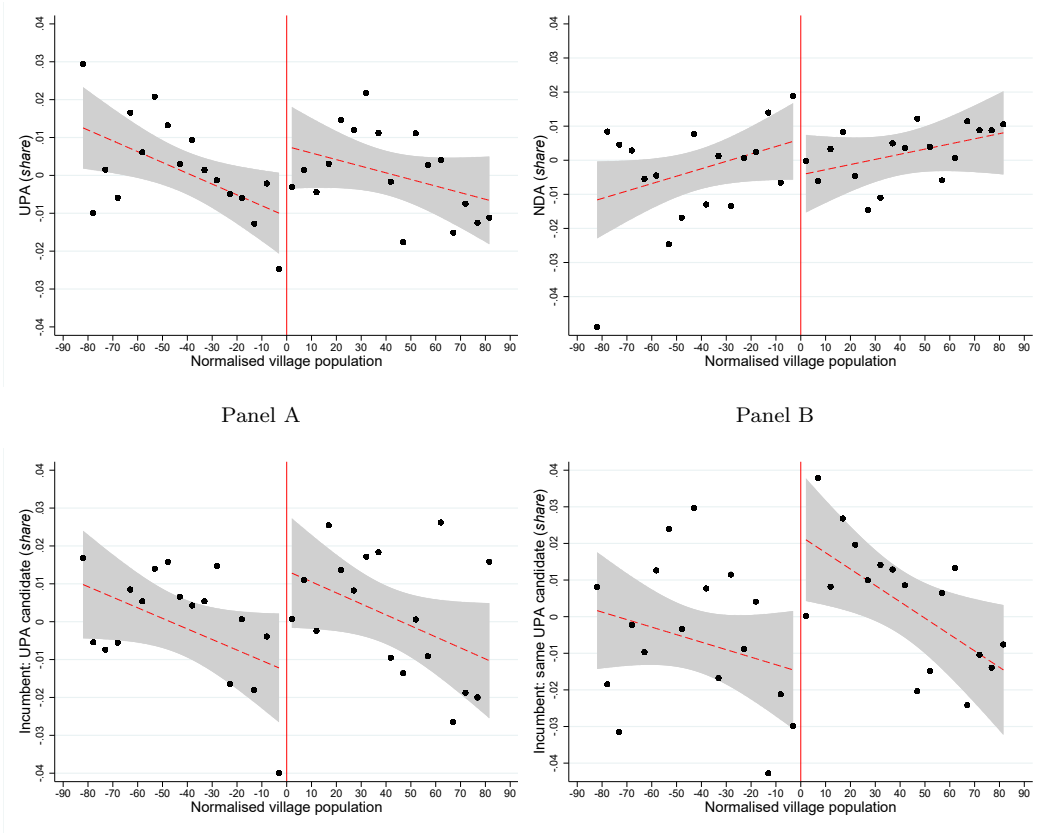
Notes: This figure plots the residualised village characteristics at baseline (after controlling for all variables in the main specification excluding the covariate of interest) against the assignment variable. The assignment variable – village population – is normalised around the PMGSY eligibility threshold (either 500 or 1,000). Points to the right of zero are above the treatment thresholds, while points to the left are below. We present graphical evidence on four of our baseline control variables. Each panel shows the mean values of the control variable in each bin of the assignment variable. The bin size is 5. The red dashed lines display predicted values of the regressions in the linear case allowing for a discontinuity at the threshold, with 95% confidence intervals displayed. Formal estimates of a discontinuity for these variables using a $fuzzy$ RD design specification are reported in Table 1. Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold).

Figure 4: First stage – Effect of PMGSY eligibility rules on road construction



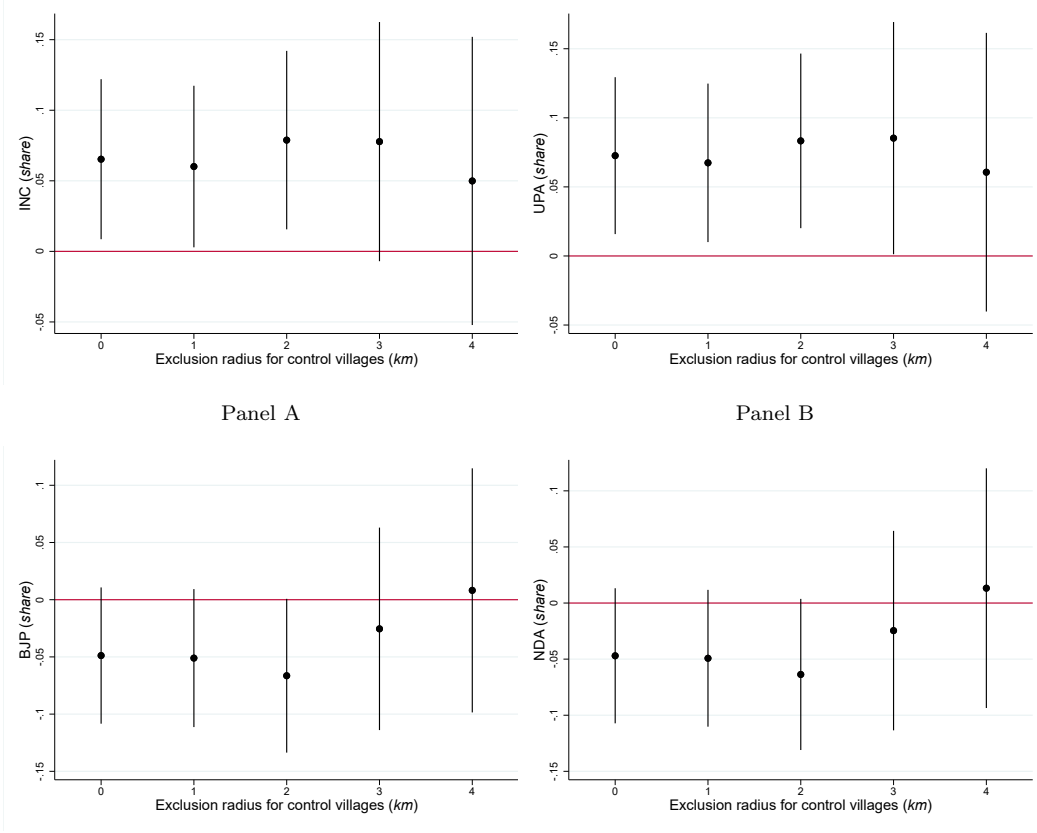
Notes: The figure plots the probability of getting a new road under the PMGSY by 2014 against the assignment variable. The assignment variable – village population – is normalised around the PMGSY eligibility threshold (either 500 or 1,000). Points to the right of zero are above the treatment thresholds, while points to the left are below. The figure shows the mean values of the variable of interest in each bin of the assignment variable. The bin size is 5. The red dashed lines display predicted values of the regressions in the linear case allowing for a discontinuity at the threshold, with 95% confidence intervals displayed. Formal estimates of a discontinuity for this variable using a $fuzzy$ RD design specification are reported in Table 2. Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold).

Figure 5: Reduced form – Effect of PMGSY road construction on voting behaviour



Notes: The figure plots residualised voting outcomes at the 2014 general election (after controlling for all variables in the main specification) against the assignment variable. The assignment variable – village population – is normalised around the PMGSY eligibility threshold (either 500 or 1,000). Points to the right of zero are above the treatment thresholds, while points to the left are below. We present graphical evidence on four of our voting outcomes. Each panel shows the mean values of the outcome variable in each bin of the assignment variable. The bin size is 5. The red dashed lines display predicted values of the regressions in the linear case allowing for a discontinuity at the threshold, with 95% confidence intervals displayed. Formal estimates of a discontinuity for these variables using a $fuzzy$ RD design specification are reported in Table 5 for Panels A and B, and Table 8 for Panels C and D. Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold).

Figure 6: Coefficient plots – Spillover effects of PMGSY road construction on voting behaviour



Panel A Panel B

Panel C Panel D

Notes: The figure plots point estimates and 90% confidence intervals for the effect of a new road on voting outcomes at the 2014 general election in villages within differing radii from the beneficiary village. We present voting outcomes for the INC, UPA, BJP, and NDA in Panels A, B, C, and D respectively. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). Using the final sample of 10,431 villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold). For each political party category, we restrict the sample to villages where a representative of that category was listed on the ballot. Confidence intervals are calculated using heteroskedasticity robust standard errors.

Table 1: Summary statistics and balance of village characteristics at baseline

Variable	Full sample (1)	Below threshold (2)	Over threshold (3)	Difference in means (4)	p-value on difference (5)	RD estimate (6)	p-value on RD estimate (7)
Panel A: Amenities							
Primary school (*binary*)	0.954	0.949	0.960	0.011	0.01	-0.001	0.98
Medical centre (*binary*)	0.159	0.148	0.172	0.025	0.00	-0.046	0.49
Electricity (*binary*)	0.424	0.408	0.441	0.034	0.00	0.129	0.16
Distance from nearest town (*km*)	26.949	26.902	27.002	0.100	0.82	-4.704	0.24
Panel B: Agricultural sector							
Agricultural land area (*ln*)	5.149	5.096	5.209	0.112	0.00	0.027	0.84
Agricultural land irrigated (*share*)	0.280	0.275	0.286	0.011	0.06	0.048	0.35
Panel C: Demographics							
Literacy (*share*)	0.457	0.453	0.461	0.007	0.02	0.005	0.85
Scheduled caste (*share*)	0.142	0.141	0.143	0.002	0.54	-0.010	0.75
Landownership (*share*)	0.737	0.738	0.736	-0.003	0.53	0.019	0.66
Subsistence agriculture (*share*)	0.434	0.438	0.430	-0.007	0.18	0.035	0.47
HH income above Rs.250/month (*share*)	0.759	0.758	0.760	0.002	0.76	-0.049	0.32
N	10431	5514	4917				

Notes: The table presents summary statistics and balance tests for village characteristics measured at baseline. Data on these baseline covariates is obtained from the 2001 Population Census of India (variables covering village amenities, agricultural sector indicators, share of the population that are literate, and belonging to a scheduled caste) and the 2002 Below Poverty Line Census (variables on demographics including share of the population that own land, rely on subsistence agriculture, and with a HH income above Rs.250/month). Columns 1-3 show the unconditional mean for all villages, villages below the population threshold, and villages above the population threshold respectively. Column 4 presents the difference in means between Columns 2 and 3. Column 5 shows the p-value for the difference in means. Column 6 reports the regression discontinuity estimates capturing the effect of being above the population threshold on the baseline covariate. The specification includes baseline village-level controls (with the covariate of interest omitted from the vector of controls) and constituency-cutoff fixed effects (see Section 4 for details). Finally Column 7 presents the p-value for the regression discontinuity estimates. Using the final sample of villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold) treated between 2004 to 2014.

Table 2: First stage – Effect of PMGSY eligibility rules on road construction

	Bandwidth of the population threshold					
	± 60	± 70	± 80	± 90	± 100	± 110
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New roads completed between 2004 and 2014						
New road	0.232***	0.227***	0.223***	0.220***	0.219***	0.219***
	(0.022)	(0.020)	(0.019)	(0.018)	(0.017)	(0.016)
F-Statistic	114.53	128.11	141.11	152.23	167.10	184.25
N	7742	9027	10305	11572	12894	14166
R^2	0.19	0.19	0.18	0.18	0.18	0.18
Panel B: New roads completed between 2009 and 2014						
New road	0.160***	0.157***	0.153***	0.149***	0.148***	0.148***
	(0.022)	(0.020)	(0.019)	(0.018)	(0.017)	(0.016)
F-Statistic	52.74	59.05	63.94	68.11	74.51	80.97
N	6046	7082	8096	9098	10099	11099
R^2	0.16	0.16	0.15	0.15	0.15	0.15
Panel C: New roads completed between 2004 and 2009						
New road	0.193***	0.190***	0.189***	0.186***	0.186***	0.186***
	(0.022)	(0.020)	(0.019)	(0.018)	(0.017)	(0.016)
F-Statistic	78.68	89.21	99.86	108.98	120.00	133.03
N	6378	7436	8481	9552	10662	11725
R^2	0.22	0.21	0.21	0.21	0.21	0.21

Notes: This table presents first stage estimates on the effect of being above the PMGSY eligibility threshold on the probability of a village being treated – receiving a new road. Panel A presents result where the dependent variable is an indicator variable that takes the value one if the village has received a PMGSY road between 2004 and 2014. Panel B presents results for villages treated between the 2009 and 2014 general elections, while Panel C presents results for villages treated between the 2004 and 2019 general elections. Results for villages within 60 of the population threshold (440-560 for the low threshold and 940-1060 for the high threshold) are presented in Column 1. Columns 2-6 expand the sample to include villages within 70, 80, 90, 100, and 110 of the population threshold. Using the sample of villages from the six compliant States and with no paved road at baseline. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). Heteroskedasticity robust standard errors are reported in parenthesis below the point estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Impact of new road on transportation

	Public bus (*binary*) (1)	Private bus (*binary*) (2)	Taxi (*binary*) (3)	Van (*binary*) (4)	Autorickshaw (*binary*) (5)
Panel A: New roads completed between 2004 and 2011					
New road	0.151**	0.076	-0.001	-0.069	0.137***
	(0.065)	(0.080)	(0.053)	(0.072)	(0.052)
F-statistic	135.930	135.930	135.930	135.930	135.930
N	10431	10431	10431	10431	10431
Panel B: New roads completed between 2009 and 2011					
New road	0.117	0.163	-0.058	-0.427***	0.289***
	(0.119)	(0.146)	(0.091)	(0.124)	(0.106)
F-statistic	59.985	59.985	59.985	59.985	59.985
N	8223	8223	8223	8223	8223
Panel C: New roads completed between 2004 and 2009					
New road	0.208**	0.179*	0.034	0.005	0.183***
	(0.088)	(0.105)	(0.068)	(0.094)	(0.069)
F-statistic	94.830	94.830	94.830	94.830	94.830
N	9080	9080	9080	9080	9080
Control mean	0.117	0.204	0.070	0.157	0.051
Control SD	0.321	0.403	0.255	0.363	0.219

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on transportation facilities to the village by 2011. We consider five categories of transport services – public buses, private buses, taxis, vans, and autorickshaws, reported in Columns 1 to 5 respectively – measured by an indicator variable which takes the value one if the service is present within the village and zero otherwise. Panel A includes all new PMGSY roads built between 2004 and 2011. Panel B refers to villages which received a new road between 2009 and 2011, while Panel C only includes villages with a new road built between the 2004 and 2009 general elections. Using the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4: Impact of new road on consumption

	Consumption per capita	Poverty rate	Household assets	Night light
	(ln)	$(share)$	$(index)$	(ln)
	(1)	(2)	(3)	(4)
Panel A: New roads completed between 2004 and 2012 (2014 for night light)				
New road	0.018	-0.019	-0.020	0.064
	(0.043)	(0.036)	(0.172)	(0.178)
F-statistic	70.05	69.38	135.930	142.064
N	10431	10431	10431	10431
Panel B: New roads completed between 2009 and 2012 (2014 for night light)				
New road	-0.079	0.079	-0.083	-0.284
	(0.082)	(0.067)	(0.314)	(0.294)
F-statistic	30.77	31.36	59.985	65.794
N	8222	8219	8223	8223
Panel C: New roads completed between 2004 and 2009				
New road	0.022	-0.028	-0.042	0.183
	(0.058)	(0.047)	(0.227)	(0.236)
F-statistic	48.91	49.85	94.830	99.123
N	9079	9078	9080	8594
Control mean	2.657	0.477	-0.006	12.163
Control SD	0.282	0.219	1.006	22.068

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on consumption by 2012 (2014 for night light). We consider three direct measures of consumption: imputed log consumption per capita (Column 1), share of the population living below the poverty line (Column 2; poverty line is set at Rs.31/day), and a household asset ownership index (Column 3; calculated as the village-level average of the primary component of indicator variables for all household assets captured in the Socio Economic Caste Census of 2012). Additionally, we rely on measures of night light luminosity from satellite images as a proxy for consumption. We calculate the average of total night light over a three year period (Column 4). Panel A includes all new PMGSY roads built between 2004 and 2012 (in the case of night light this is measured until 2014). Panel B refers to villages which received a new road between 2009 and 2012 (2014 in the case of night light), while Panel C only includes villages with a new road built between the 2004 and 2009 general elections. Using the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Summary statistics for consumption per capita and night light are reported on the level form of the variable. Heteroskedasticity robust standard errors are presented in parenthesis, except for consumption and poverty which report bootstrapped standard errors. * significant at 10% ** significant at 5% *** significant at 1%.

Table 5: Impact of new road on share of votes to political parties

	United Progressive Alliance		National Democratic Alliance		Small parties & independents
	INC	All parties	BJP	All parties	
	(1)	(2)	(3)	(4)	(5)
Panel A: New roads completed between 2004 and 2014					
New road	0.065*	0.073**	-0.049	-0.047	-0.030
	(0.035)	(0.035)	(0.036)	(0.037)	(0.027)
F-statistic	140.523	141.498	145.003	140.942	140.337
N	7339	7538	7316	7554	7564
Panel B: New roads completed between 2009 and 2014					
New road	0.087	0.102	-0.058	-0.046	-0.058
	(0.067)	(0.065)	(0.066)	(0.066)	(0.047)
F-statistic	50.376	53.208	53.814	53.331	53.241
N	5487	5657	5461	5670	5679
Panel C: New roads completed between 2004 and 2009					
New road	0.079*	0.086**	-0.062	-0.060	-0.033
	(0.041)	(0.042)	(0.045)	(0.046)	(0.034)
F-statistic	110.070	107.663	110.923	106.099	106.067
N	6147	6330	6133	6348	6354
Control mean	0.357	0.359	0.518	0.517	0.126
Control SD	0.194	0.194	0.203	0.203	0.155

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on the share of votes to political parties in the 2014 general election of India. Panel A includes all new PMGSY roads built between the 2004 and 2014 general elections. Panel B reports results for villages treated between the 2009 and 2014 general elections, while Panel C presents results for villages treated between the 2004 and 2009 general elections. We consider three political party categories: UPA (Column 1-2), NDA (Column 3-4), and small parties and independent candidates unaffiliated to any alliance (Column 5). For the two leading alliances, we report votes to both the leading party (INC for the UPA and BJP for the NDA) as well as all parties in the coalition. For each political party category, we restrict the sample to villages where a representative of that category was listed on the ballot. Using the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Table 6: Impact of new road on share of votes to political parties in villages within 2km of a PMGSY road

	United Progressive Alliance		National Democratic Alliance		Small parties & independents
	INC	All parties	BJP	All parties	
	(1)	(2)	(3)	(4)	(5)
Panel A: New roads completed between 2004 and 2014					
New road	0.080**	0.084**	-0.079**	-0.075*	-0.013
	(0.038)	(0.037)	(0.040)	(0.040)	(0.030)
F-statistic	110.345	111.564	112.304	110.767	110.406
N	5967	6136	5943	6151	6161
Panel B: New roads completed between 2009 and 2014					
New road	0.113*	0.118*	-0.096	-0.080	-0.038
	(0.063)	(0.062)	(0.063)	(0.063)	(0.044)
F-statistic	48.091	50.761	50.065	50.630	50.648
N	4119	4259	4092	4271	4280
Panel C: New roads completed between 2004 and 2009					
New road	0.080*	0.084*	-0.090*	-0.085*	-0.005
	(0.044)	(0.045)	(0.048)	(0.049)	(0.037)
F-statistic	87.051	85.100	86.575	83.615	83.653
N	4776	4929	4761	4946	4952
Control mean	0.358	0.360	0.514	0.514	0.129
Control SD	0.191	0.191	0.199	0.199	0.153

Notes: This table presents *fuzzy* RD estimates on the spillover effect of a new road on the share of votes to political parties in villages within 2km of a PMGSY for the 2014 general election of India. Panel A includes all new PMGSY roads built between the 2004 and 2014 general elections. Panel B reports results for villages treated between the 2009 and 2014 general elections, while Panel C presents results for villages treated between the 2004 and 2009 general elections. We consider three political party categories: UPA (Column 1-2), NDA (Column 3-4), and small parties and independent candidates unaffiliated to any alliance (Column 5). For the two leading alliances, we report votes to both the leading party (INC for the UPA and BJP for the NDA) as well as all parties in the coalition. For each political party category, we restrict the sample to villages where a representative of that category was listed on the ballot. Using the neighbouring village within 2km from the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Table 7: Impact of new road on share of votes to political parties, excluding control villages within 2km of the treatment group

	United Progressive Alliance		National Democratic Alliance		Small parties & independents
	INC	All parties	BJP	All parties	
	(1)	(2)	(3)	(4)	(5)
Panel A: New roads completed between 2004 and 2014					
New road	0.079**	0.083**	-0.066	-0.064	-0.024
	(0.039)	(0.039)	(0.041)	(0.041)	(0.031)
F-statistic	109.065	110.283	110.983	109.490	109.129
N	5972	6141	5948	6156	6166
Panel B: New roads completed between 2009 and 2014					
New road	0.100	0.108*	-0.078	-0.063	-0.046
	(0.064)	(0.063)	(0.064)	(0.064)	(0.046)
F-statistic	47.712	50.375	49.677	50.244	50.260
N	4120	4260	4093	4272	4281
Panel C: New roads completed between 2004 and 2009					
New road	0.097**	0.100**	-0.091*	-0.087*	-0.019
	(0.045)	(0.046)	(0.050)	(0.050)	(0.038)
F-statistic	86.040	84.103	85.546	82.626	82.664
N	4780	4933	4765	4950	4956
Control mean	0.357	0.359	0.517	0.516	0.128
Control SD	0.193	0.194	0.202	0.202	0.154

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on the share of votes to political parties in the 2014 general election of India where control villages with 2 km of PMGSY roads are excluded. Panel A includes all new PMGSY roads built between the 2004 and 2014 general elections. Panel B reports results for villages treated between the 2009 and 2014 general elections, while Panel C presents results for villages treated between the 2004 and 2009 general elections. We consider three political party categories: UPA (Column 1-2), NDA (Column 3-4), and small parties and independent candidates unaffiliated to any alliance (Column 5). For the two leading alliances, we report votes to both the leading party (INC for the UPA and BJP for the NDA) as well as all parties in the coalition. For each political party category, we restrict the sample to villages where a representative of that category was listed on the ballot. Using the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold and outside a 2 km radius of the treated group) is reported. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Table 8: Impact of new road on votes to the constituency incumbent

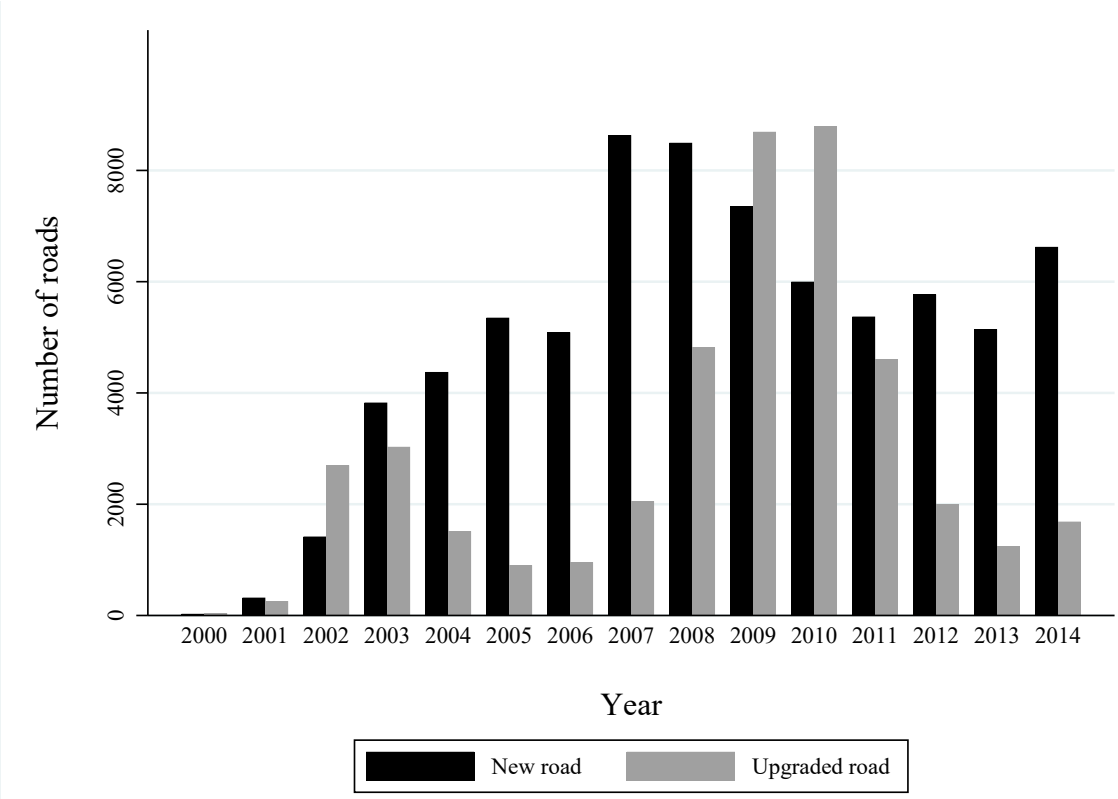
	All candidates		UPA candidate		NDA candidate	
	Same	New	Same	New	Same	New
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New roads completed between 2004 and 2014						
New road	0.159***	0.000	0.174***	0.007	0.064	0.028
	(0.049)	(0.058)	(0.053)	(0.059)	(0.080)	(0.142)
F-statistic	98.129	52.055	69.350	56.685	33.664	8.998
N	6502	3754	3000	1339	1501	1337
Panel B: New roads completed between 2009 and 2014						
New road	0.238***	0.009	0.249**	0.062	0.092	-0.005
	(0.086)	(0.094)	(0.099)	(0.220)	(0.145)	(0.128)
F-statistic	42.303	22.666	24.261	3.795	13.962	14.992
N	5215	2873	2216	790	1238	1138
Panel C: New roads completed between 2004 and 2009						
New road	0.232***	-0.019	0.218***	0.009	0.124	-0.126
	(0.070)	(0.074)	(0.069)	(0.058)	(0.107)	(0.405)
F-statistic	67.632	40.191	49.723	64.584	24.604	1.627
N	5476	2977	2545	1141	1277	1055
Control mean	0.421	0.421	0.366	0.402	0.545	0.515
Control SD	0.214	0.214	0.196	0.214	0.197	0.180

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on the share of votes to the constituency incumbent in the 2014 general election of India. Panel A includes all new PMGSY roads built between the 2004 and 2014 general elections. Panel B reports results for villages treated between the 2009 and 2014 general elections, while Panel C presents results for villages treated between the 2004 and 2009 general elections. We disaggregate incumbency by political affiliation – whether the candidate represents the UPA (Columns 3-4) or the NDA (Column 5-6). Furthermore, we disaggregate incumbency by whether or not the incumbent is the same individual or a new individual (representing the incumbent party). The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Using the final sample of villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Appendices

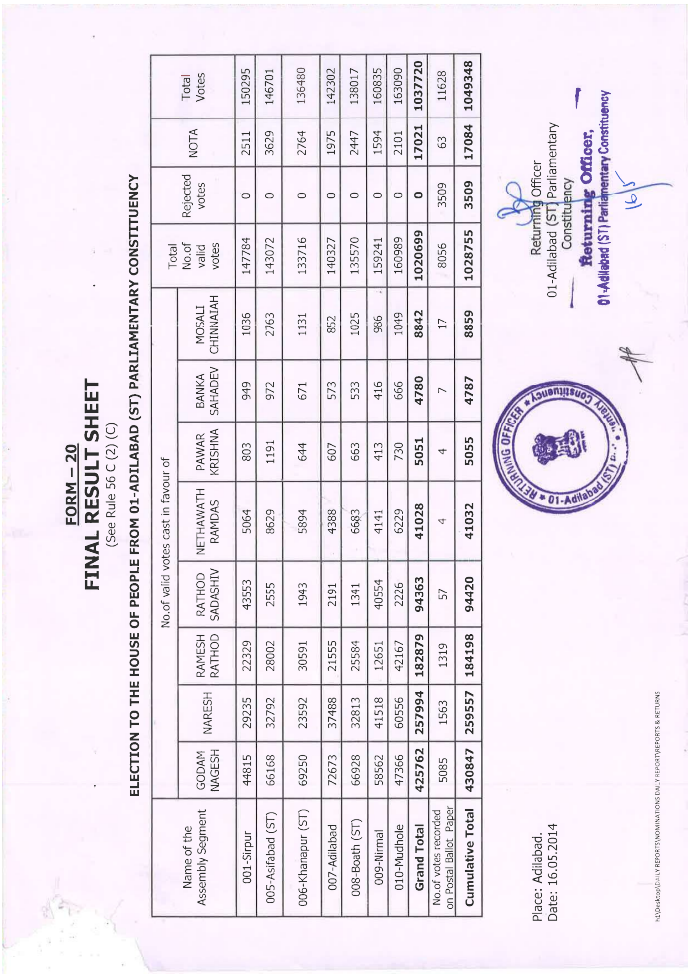
A Additional Tables and Figures

Figure A1: New and upgraded roads completed under the PMGSY by year



Notes: This figure shows road completion under the PMGSY initiative between 2000 to 2014, disaggregated by whether the roads were newly built or upgraded. The bars represent the number of roads completed under each category by year. Based on the sample of all 185,000 villages that benefited from the programme during this period.

Figure A2: Example of Form20



FORM – 20

FINAL RESULT SHEET

(See Rule 56 C (2) (C)

ELECTION TO THE HOUSE OF PEOPLE FROM 01-ADILABAD (ST) PARLIAMENTARY CONSTITUENCY

Name of the Assembly Segment	No.of valid votes cast in favour of								Total No.of valid votes	Rejected votes	NOTA	Total Votes
	GODAM NAGESH	NARESH	RAMESH RATHOD	RATHOD SADASHIV	NETHAWATH RAMDAS	PAWAR KRISHNA	BANKA SAHADEV	MOSALI CHINNAIAH				
001-Sirpur	44815	29235	22329	43553	5064	803	949	1036	147784	0	2511	150295
005-Asifabad (ST)	66168	32792	28002	2555	8629	1191	972	2763	143072	0	3629	146701
006-Khanapur (ST)	69250	23592	30591	1943	5894	644	671	1131	133716	0	2764	136480
007-Adilabad	72673	37488	21555	2191	4388	607	573	852	140327	0	1975	142302
008-Boath (ST)	66928	32813	25584	1341	6683	663	533	1025	135570	0	2447	138017
009-Nirmal	58562	41518	12651	40554	4141	413	416	986	159241	0	1594	160835
010-Mudhole	47366	60556	42167	2226	6229	730	666	1049	160989	0	2101	163090
Grand Total	**425762**	**257994**	**182879**	**94363**	**41028**	**5051**	**4780**	**8842**	**1020699**	**0**	**17021**	**1037720**
No.of votes recorded on Postal ballot Paper	5085	1563	1319	57	4	4	7	17	8056	3509	63	11628
Cumulative Total	**430847**	**259557**	**184198**	**94420**	**41032**	**5055**	**4787**	**8859**	**1028755**	**3509**	**17084**	**1049348**

Place: Adilabad.
Date: 16.05.2014

Returning Officer
01-Adilabad (ST) Parliamentary Constituency

Returning Officer,
01-Adilabad (ST) Parliamentary Constituency

Notes: Extract from a Form20 document.

Table A1: Impact of new road on ownership of assets

	Solid house	Land	Mechanised equipment	Refrigerator	Vehicle	Phone
	(*share*)	(*share*)	(*share*)	(*share*)	(*share*)	(*share*)
	(1)	(2)	(3)	(4)	(5)	(6)
New road	0.000	0.006	0.004	0.001	0.003	0.057
	(0.047)	(0.042)	(0.014)	(0.015)	(0.027)	(0.051)
F-statistic	135.930	135.930	135.840	135.930	135.930	135.930
N	10431	10431	10430	10431	10431	10431
Control mean	0.224	0.572	0.041	0.037	0.142	0.447
Control SD	0.265	0.248	0.077	0.083	0.156	0.298

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on household asset ownership by 2012. We consider the share of households in the village that own each of the following assets: solid house (Column 1), land (Column 2), mechanised farm equipment (Column 3), refrigerator (Column 4), vehicle (Column 5), and phone (Column 6). The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Using the final sample of villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold) treated between 2004 to 2014. Heteroskedasticity robust standard errors are presented in parenthesis. * significant at 10% ** significant at 5% *** significant at 1%.

Table A2: First stage and reduced form estimates on voting behaviour for the main and placebo samples

	First stage	Reduced form			
	Above population threshold	Political party		Incumbent	
		UPA	NDA	UPA	NDA
	(*binary*)	(*share*)	(*share*)	(*share*)	(*share*)
	(1)	(2)	(3)	(4)	(5)
Panel A: Main sample					
New road	0.223***	0.019**	-0.012	0.048***	0.016
	(0.019)	(0.009)	(0.009)	(0.013)	(0.021)
F-statistic	142.06				
N	10431	7538	7554	3000	1501
Control mean	0.274	0.36	0.52	0.366	0.545
Control SD	0.446	0.19	0.20	0.196	0.197
Panel B: Placebo sample					
New road	0.001	0.008	-0.017	0.022	-0.004
	(0.019)	(0.010)	(0.010)	(0.015)	(0.026)
F-statistic	0.00				
N	8140	7113	6845	2525	1274
Control mean	0.253	0.25	0.40	0.262	0.416
Control SD	0.435	0.23	0.23	0.220	0.238

Notes: This table presents first stage estimates on the effect of being above the PMGSY eligibility population threshold on the probability of a village being treated – receiving a new road, as well as reduced form estimates on the effect of a new road on voting behaviour in the 2014 general election of India, for varying samples. Panel A reports results on the final sample of villages used in the main analysis (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold) treated between 2004 to 2014. Panel B reports results on a placebo sample of villages from States that did not comply to the policy guidelines, as well as villages close to the 1,000 threshold in states that used only the 500 person threshold. We report this robustness test on four of our voting outcomes measured as the share of votes to the UPA (Column 2), NDA (Column 3), parliamentary constituency UPA (Column 4) and NDA (Column 5) incumbent individual. For each of these voting outcomes we restrict the sample to villages where a representative of that category was listed on the ballot. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression, the mean and SD of the control group (villages with population below the PMGSY eligibility threshold) is reported. Heteroskedasticity robust standard errors are reported below point estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table A3: Impact of new road on voting behaviour by kernel and bandwidth

	Political party		Incumbent	
	UPA	NDA	UPA	NPA
	(*share*)	(*share*)	(*share*)	(*share*)
	(1)	(2)	(3)	(4)
Panel A: Triangular kernel				
± 60	0.072*	-0.059	0.194***	0.051
	(0.040)	(0.042)	(0.062)	(0.092)
± 80	0.069**	-0.045	0.167***	0.062
	(0.035)	(0.037)	(0.053)	(0.082)
± 100	0.062*	-0.034	0.142***	0.071
	(0.032)	(0.034)	(0.048)	(0.078)
Panel B: Rectangular kernel				
± 60	0.077*	-0.053	0.096**	0.106
	(0.040)	(0.042)	(0.042)	(0.070)
± 80	0.061*	-0.026	0.096**	0.106
	(0.034)	(0.036)	(0.042)	(0.070)
± 100	0.047	-0.008	0.096**	0.106
	(0.029)	(0.031)	(0.042)	(0.070)

Notes: This table presents *fuzzy* RD estimates on the effect of a new road on voting behaviour in the 2014 general election of India, for varying bandwidths and kernel specifications. Panel A shows results using a triangular kernel, while Panel B uses a rectangular kernel weighting. In each case, we show robustness to three different bandwidth choices (60, 80, 100) for the sample of villages from the six compliant States with no paved road at baseline and treated between 2004 to 2014. We report this test on four of our voting outcomes measured as the share of votes to the UPA (Column 1), NDA (Column 2), parliamentary constituency incumbent (Column 3), parliamentary constituency UPA (Column 3) and NDA (Column 4) incumbent individual. For each of these voting outcomes we restrict the sample to villages where a representative of that category was listed on the ballot. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). Heteroskedasticity robust standard errors are presented in parentheses and p-values in brackets. * significant at 10% ** significant at 5% *** significant at 1%.

Table A4: First stage and RD estimates on voting behaviour when excluding controls and fixed effects

	First stage	Regression Discontinuity			
	Above population threshold	Political party		Incumbent	
		UPA	NDA	UPA	NDA
	(*binary*)	(*share*)	(*share*)	(*share*)	(*share*)
	(1)	(2)	(3)	(4)	(5)
Panel A: Excluding baseline controls					
New road	0.223***	0.074**	-0.047	0.178***	0.073
	(0.019)	(0.035)	(0.037)	(0.054)	(0.080)
F-statistic	135.84	132.465	132.478	65.487	33.800
N	10431	7538	7554	3000	1501
Panel A: Excluding baseline controls and fixed effects					
New road	0.212***	0.066*	-0.039	0.191***	0.064
	(0.020)	(0.040)	(0.042)	(0.065)	(0.077)
F-statistic	111.32	101.240	100.502	49.275	34.808
N	10431	7538	7554	3000	1501
Control mean	0.27	0.359	0.517	0.366	0.545
Control SD	0.45	0.194	0.203	0.196	0.197

Notes: This table presents first stage estimates on the effect of being above the PMGSY eligibility population threshold on the probability of a village being treated – receiving a new road, as well as *fuzzy* RD estimates on the effect of a new road on voting behaviour in the 2014 general election of India, when varying controls. Panel A reports results when excluding the vector of baseline controls from the main specification. Panel B reports results when excluding both the vector of baseline controls as well as the constituency-cutoff fixed effects. We report this robustness test on four of our voting outcomes measured as the share of votes to the UPA (Column 2), NDA (Column 3), parliamentary constituency UPA (Column 4) and NDA (Column 5) incumbent individual. For each of the voting categories, we restrict the sample to villages where a representative of that category was listed on the ballot. Using the final sample of villages (villages from the six compliant States, with no paved road at baseline, and a population within the optimal bandwidth (84) of the eligibility threshold) treated between 2004 to 2014. For each regression we report the mean and SD of the control group (villages with population below the PMGSY eligibility threshold). Heteroskedasticity robust standard errors are reported below point estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table A5: First stage and RD estimates on voting behaviour for the sample of villages with polling-stations within their boundary

	First stage	Regression Discontinuity			
	Above population threshold	Political party		Incumbent	
		UPA	NDA	UPA	NDA
	(*binary*)	(*share*)	(*share*)	(*share*)	(*share*)
	(1)	(2)	(3)	(4)	(5)
New road	0.238***	0.099**	-0.082	0.094	0.049
	(0.028)	(0.050)	(0.053)	(0.065)	(0.104)
F-statistic	71.65	70.978	69.160	37.258	17.448
N	4401	3028	3036	1273	1273
Control mean	0.286	0.358	0.514	0.367	0.539
Control SD	0.452	0.195	0.206	0.196	0.187

Notes: This table presents first stage estimates on the effect of being above the PMGSY eligibility population threshold on the probability of a village being treated – receiving a new road, as well as *fuzzy* RD estimates on the effect of a new road on voting behaviour in the 2014 general election of India, for the reduced sample of villages with a polling-station. We report this robustness test on four of our voting outcomes measured as the share of votes to the UPA (Column 2), NDA (Column 3), parliamentary constituency UPA (Column 4) and NDA (Column 5) incumbent individual. For each of the voting categories, we restrict the sample to villages where a representative of that category was listed on the ballot. Using the reduced sample of villages (villages from the six compliant States, with no paved road at baseline, a population within the optimal bandwidth (84) of the eligibility threshold) treated between 2004 yo 2014 with a polling-station within the village boundary. The specification includes baseline village-level controls and constituency-cutoff fixed effects (see Section 4 for details). For each regression we report the mean and SD of the control group (villages with population below the PMGSY eligibility threshold). Heteroskedasticity robust standard errors are reported below point estimates. * significant at 10% ** significant at 5% *** significant at 1%.

B Appendix: Data

B1 Consumption

Most developing countries do not collect detailed information on income or consumption as part of their censuses. As such, estimates of these economic indicators at a high geographic resolution are often unavailable at regular time intervals. Policy makers (especially the World Bank) and researchers have therefore recently relied on a method developed by Elbers et al. (2003) which uses an imputation rule derived from a household survey to generate small-area estimates of consumption in census data (Bedi et al., 2007). In a comparison of methods, McKenzie (2005) show that this prediction method through auxiliary surveys most accurately predicts non-durable consumption. Hentschel et al. (2000), demonstrate that this method produces unbiased estimates of poverty.

Since the early 1990s the Government of India has conducted national socioeconomic censuses collecting information at both the individual and household level on caste, occupation, earnings, and assets, in order to determine the eligibility of households into various welfare schemes (Alkire and Seth, 2013). In 2012, the fourth such Socio Economic Caste Census (SECC) was implemented.[23] In that year, the India Human Development Survey-II (IHDS-II) was also conducted. It recorded direct measures of household consumption, as well as equivalent questions to the SECC on household assets and earnings.[24] Following the methodology of Elbers et al. (2003), Asher et al. (2021) use the IHDS-II data to predict household level consumption in the SECC dataset. Specifically, the researchers first estimate regressions of total household consumption on dummy variables of assets and earnings in the IHDS-II.[25] Coefficients from these regressions are then used to impute household level consumption values in the SECC. Finally, based on these household level values the researchers generate village level statistics for mean predicted consumption per capita and the share of the population below the poverty line.[26] Bootstrap estimates of these village level indicators are made available by the research team on the Socioeconomic High-resolution Rural-Urban Geographic (SHRUG, Version 1.5) open data platform for

[23]Information on the census can be found on the SECC website:https://secc.gov.in/welcome. Though the Government initially made the raw data public, only aggregated information is now available on the website.

[24]Information and data related to this survey can be found on the platform of Data Sharing for Demographic Research:https://www.icpsr.umich.edu/web/pages/DSDR/index.html

[25]These are the exact same variables as those recorded in the SECC. They include: type of roof and wall material, number of rooms, ownership of phone, house, vehicle, land, kisan credit card, and refrigerator, as well as the highest individual income in the household.

[26]The official poverty line for rural India is set at Rs.27/day, based on the Planning Commission's Tendulkar Committee Report in 2014.

India.[27] We take these 1000 bootstrapped variables for predicted consumption per capita (for the purpose of the regression, these variables are log transformed) and share of the population below the poverty line, and run an additional bootstrap process on our main sample of villages when estimating the effect of access to rural roads on these indicators. As outlined in the work of Elbers et al. (2003), this bootstrapping process is required to obtain correct standard errors and p-values on our estimates.

Specific to our setting of Indian villages, Asher et al. (2021) provide three validation tests for the bootstrap estimates of consumption used in our analysis. First, the distribution of the consumption estimates at the village level matches broadly to that found in two national surveys conducted at the same time and at the same geographic level (IHDS-II and the National Sample Survey-2012). Second, there is a strong covariance between the district level predicted consumption estimates and those in the original household survey (IHDS-II). Third, by identifying how each component used in the imputation rule affects the difference in average consumption between the estimates and the original survey (IHDS-II), the researchers find that the transformation of asset ownership to consumption assumes a similar relationship across datasets. These findings provide confirmation that the predicted consumption estimates are valid proxies of the direct survey measures.

B2 Night Light

As an additional proxy for consumption, we leverage remote sensing imagery on Night-Time Light (NLT) at the village level across India. Initiated by the work of Henderson et al. (2011), NTL has since become a widely used proxy for economic activity. Researchers have adopted night-time luminosity to effectively capture GDP growth (Henderson et al., 2011), cross-sectional GDP (Bleakley and Lin, 2012), urbanisation (Harari, 2020), public expenditure (Hodler and Raschky, 2014), and employment (Mellander et al., 2015). In an analysis of Indian villages, Asher et al. (2021) find that night light is a highly statistically significant proxy for a range of development outcomes including - population, employment, per capita consumption, and electrification.

Night-time luminosity data is made available by the U.S. National Oceanographic and Atmospheric Administration (NOAA). Data come from the Visible Infrared Imaging Radiometer Suite (VIIRS) instruments aboard the joint NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites. This remotely sensed data provide global daily measurements, at a spatial resolution of 375 and 750 meters (depending on the band), within the visible to near-infrared portion of the spectrum: 400-900

[27] For detailed information on consumption data using the SHRUG open data platform, please refer to Asher et al. (2021). The dataset, including codebooks and references, can be found at:http://www.devdatalab.org/shrug

nanometers (nm). This spectral range is ideal for exploring night time luminosity as many human-made light sources provide spectral responses in this range. A description of the satellite instrumentation, data collection, and processing methods for NTL is detailed in the work of Elvidge et al. (2021). Asher et al. (2021) leverage similar data to verify the effectiveness of night-time luminosity as a proxy for development indicators at the village level in India. These data are aggregated to villages and towns across the country and made available by the research team on the Socioeconomic High-resolution Rural-Urban Geographic (SHRUG, Version 2.0) open data platform for India.[28]

[28]For detailed information on NTL data using the SHRUG open data platform, please refer to Asher et al. (2021). The dataset, including codebooks and references, can be found at:http://www.devdatalab.org/shrug