Political Power-Sharing, Firm Entry, and Economic Growth: Evidence from Multiple Elected Representatives*

Harsha Dutta† Pulak Ghosh‡ Arkodipta Sarkar§ Nishant Vats¶

September 13, 2024

Abstract

We examine the effect of political power-sharing on local economic activity. This effect depends on the relative importance of the risks associated with unchecked power and the potential efficiency gains or losses arising from checks and balances. Our research design exploits a geographic discontinuity design due to the haphazard overlap of electoral and administrative boundaries that generates quasi-random variation in the number of politicians governing adjacent regions. We supplement this design using an episode of electoral delimitation that allows us to exploit within-region variation in the number of politicians. We find increasing the number of politicians governing an area can lead to new firm creation, lower unemployment, and greater real economic activity. Our results suggest that non-aligned multiple politicians enhance state efficiency by imposing checks and balances on each other, leading to lower regulatory obstacles, less cronyism, and improved provision of public infrastructure, creating an economically favourable environment for firm creation.

*We thank Isha Agarwal, Sumit Agarwal, Sam Asher, Marianne Bertrand, Robin Burgess, Utpal Bhattacharya, Maxime Bonelli, Emanuele Colonnelli, Eliana La Ferrara, Stefano Fiorin, Raymond Fisman, Anand Goel, Todd Gormley, Allan Hsiao, Yan Ji, Elisabeth Kempf, Stefan M. Lewellen, Ron Masulis, Lyndon Moore, Abhiroop Mukherjee, Deniz Okat, Imran Rasul, Amit Seru, Amir Sufi, Margarita Tsoutsoura, Robert Vishny, Vikrant Vig, Christopher Woodruff, Guo Xu, Alminas Zaldokas and Luigi Zingales for helpful comments and suggestions. We are thankful to the seminar participants at the BREAD Conference, Chicago Brownbag Seminar, Inter-finance PhD Seminar, HKUST, NSE-NYU Conference on Indian Capital Markets, NUS, TADC, WEFIDEV, EFA, CEPR/IFS/UCL/BREAD/TCD Workshop in Development Economics, Finance Down Under. We thank the NSE-NYU Stern Initiative on the Study of Indian Capital Markets for financial support. We take responsibility for all errors

†Harsha Dutta is at the Hong Kong University of Science and Technology. eMail:: sshdutta@connect.ust.hk

‡Pulak Ghosh is at Indian Institute of Management Bangalore. eMail: pulak.ghosh@iimb.ac.in

§Arkodipta Sarkar is at the National University of Singapore. e-Mail: asarkar@nus.edu.sg

¶Nishant Vats is at Olin Business School, Washington University - St. Louis. eMail: vats@wustl.edu

1 Introduction

"Power tends to corrupt; absolute power corrupts absolutely."

-Lord Acton, April 1887 (in a letter to Bishop Creighton)

A key aspect of institutional design is the extent of unchecked power delegated to its leaders or politicians (Aghion, Alesina, and Trebbi, 2004). Specifically, finding the right balance between preventing government coercion and promoting economic performance has been a central principle in designing optimal institutions (Hayek, 1960). In this context, our paper examines a specific aspect of the political-institutional framework that facilitates power-sharing, which involves multiple elected representatives (or multiple politicians). Specifically, we test if increasing the number of politicians governing an area adds economic value. Studying this relationship is imperative in understanding the implications of political institutions that foster power-sharing among politicians, a setup commonly observed in decentralized governance systems. For instance, in the US, each state is represented by two senators. Similarly, Switzerland employs a seven-member executive council to lead its federal administration, and globally, coalition governments led by multiple political parties are on the rise.

The relationship between multiple politicians and economic growth is theoretically ambiguous. On the one hand, multiple politicians can negatively affect local economic conditions by increasing corruption due to the presence of multiple grabbing hands, higher holdup problem, and coordination failure or free-rider problems discussed in the common agency literature.[1] On the other hand, multiple politicians can improve local economic conditions by reducing power concentration and imposing checks and balances, leveraging diverse skills and division of labor, and reducing corruption.[2] Despite the theoretical ambiguity, the empirical evidence is limited, primarily due to issues of endogeneity associated with the emergence of political institutions that foster power-sharing among politicians, as discussed in Aghion, Alesina, and Trebbi (2004).

[1]See Shleifer and Vishny (1993) for a discussion on multiple grabbing hand problem. See Grossman and Hart (1986), Hart and Moore (1990), Blanchard and Kremer (1997), and Olken and Barron (2009), among others for a discussion on the holdup problem. See Bernheim and Whinston (1986), Dixit (1997), and Dixit, Grossman, and Helpman (1997), among others for a discussion on common agency problem.

[2]See Holmstrom (1978), Rasmusen and Ramseyer (1994), Persson, Roland, and Tabellini (1997), Bardhan (2002), and Aghion, Alesina, and Trebbi (2004) among others for a discussion of these issues.

We make progress on this discussion by directly estimating the effect of multiple politicians in shaping the environment that facilitates firm entry and consequently fosters economic growth using a large natural experiment in India. We exploit the nationwide haphazard overlap of electoral and administrative boundaries that create a quasi-random variation in the number of politicians governing an area. Our central finding is that increasing the number of politicians governing an area increases firm entry and economic growth. Broadly we find that the presence of multiple politicians may be associated with improved state efficiency. Particularly, we note that the positive relationship between economic growth and multiple politicians is stronger when these politicians have a greater cost of collusion or greater incentives to impose checks and balances on each other. We show that such an institutional environment can foster firm entry by reducing regulatory obstacles and cronyism as well as improving public good provision. Furthermore, we observe reduced demand for government unemployment assistance in these areas, suggesting that the development of the private sector can reduce the need for such aid.

An ideal experiment to study this question would involve randomly assigning the number of politicians responsible for governing an administrative unit. We leverage a unique setting in India that approximates such a scenario, which arises from the haphazard overlap between electoral boundaries and administrative boundaries. The electoral boundaries are drawn by a powerful non-partisan commission with the objective of equalizing populations across all electoral constituencies within a state. When defining electoral boundaries, this commission does not take into account the administrative boundaries at the block level, which is the lowest administrative unit in India. As a result, electoral boundaries may intersect with administrative units, leading to some administrative units being entirely subsumed within an electoral constituency, while others are divided or split across multiple electoral constituencies. This incongruity in the alignment of these two boundaries creates a quasi-random variation in the number of politicians governing a specific area.

This setup offers credible identification for three main reasons. First, irrespective of whether a block is entirely or partially within an electoral constituency, it falls under the jurisdiction of the politician elected from that constituency. These politicians share equal and joint responsibility for governing the block. We refer to the blocks governed

by multiple politicians as split blocks and those governed by a single politician as unsplit blocks. Specifically, if a block is entirely contained within an electoral constituency, it is overseen by a single politician. Conversely, if a block is divided among multiple electoral constituencies, it is governed by multiple politicians. Second, in 2008, there was a change in the electoral boundaries, leading to a change in the spatial distribution of politicians governing a block. This change provides us with the opportunity to perform two types of analyses – a simple geographic regression discontinuity analysis conducted after 2008, and a differences-in-discontinuity analysis that allows us to examine pre-trends. Third, Iyer and Reddy (2013) notes that the 2008 electoral boundaries established by the election commission are mostly politically neutral and largely free from gerrymandering. The political neutrality of the 2008 delimitation episode decreases the probability of manipulation and strengthens the argument that the units near the boundary can serve as a valid counterfactual reference point.

Moreover, we focus on new firm entry as our key dependent variable for two reasons. First, firm entry is an important determinant of aggregate productivity growth and local employment (McMillan and Woodruff, 2003; Woodruff, 2018). Second, new firm creation depends on the local regulatory environment (Johnson, McMillan, and Woodruff, 2002), which is shaped by local politicians through their control of the local bureaucratic machinery (Besley, Burgess, Khan, and Xu, 2022). Therefore, new firm entry provides for a natural setup for evaluating the effect of multiple politicians on local governance.

We begin by employing geographic regression discontinuity (RD) design that combines the haphazard overlap between electoral boundaries and administrative boundaries with granular geo-referenced micro-data on firm entry after 2008. Specifically, we compare units located close to the boundary separating split blocks from unsplit blocks. Comparing units, villages hereafter, on either side of the boundary, we find the number of new firms is 3% higher and nightlights intensity is 7% higher in villages just inside a split block relative to villages in unsplit blocks just outside the boundary. Furthermore, we document that the treatment effect increases monotonically with the number of politicians governing the split blocks.

Additionally, villages just inside a split block are associated with 5% higher employment and a reduced demand for unemployment benefits, which is reflected in the lower

application rates of the National Employment Guarantee Scheme (NREGA). The lower NREGA application rates align with the findings of Gulzar and Pasquale (2017), who document lower NREGA disbursement rates in areas governed by multiple politicians. They argue that this result suggests poor implementation of the program in such areas. However, we argue that the lower NREGA disbursement may be driven by lower work-day claims or applications. This interpretation is consistent with our other results that regions with multiple politicians exhibit greater new firm entry and higher employment, manifesting as lower demand for unemployment assistance. Therefore, our findings suggest that the development of the private sector can reduce the need for government unemployment assistance.

Our RD estimation relies on two assumptions. First, we assume that all relevant factors exhibit a smooth variation at the block boundaries. To support this assumption, we analyze the 2001 Census data and find that villages located just within a split block closely resemble villages in unsplit blocks located just outside the boundary across several dimensions, including population, access to education and health, banking, infrastructure, and geography. Second, our analysis hinges on the assumption that there was no selective sorting, i.e., villages did not have the ability to choose whether they would be part of a split or unsplit block. The political neutrality of the 2008 delimitation episode and the haphazard overlap of the two boundaries reduces the likelihood of such sorting. Nevertheless, we address this concern by focusing on a subset of boundaries – separating split and unsplit blocks – which are less susceptible to manipulation. Specifically, we obtain similar results when we focus on boundaries that follow straight-line patterns or boundaries defined by geographical features, such as rivers and river basins. Overall, these tests imply that the villages located in unsplit blocks just outside the boundary of a split block are a valid counterfactual to villages situated just inside the split block.

Next, we exploit both the spatial and temporal variation generated by our natural experiment. India redrew its electoral boundaries in 2008 after three decades, resulting in widespread changes to electoral boundaries and the distribution of the number of politicians governing a block. Specifically, we use a differences-in-discontinuity design to examine the effect of the exogenous change in the number of politicians governing a block

following the 2008 delimitation.[3] This test adds to our analysis in two ways. First, this test allows us to document the effect of an area moving from being governed by a single politician to multiple politicians and vice versa. Second, this setup allows us to include village fixed effects, which relaxes the requirement for the distribution of the number of politicians across villages to be independent of their pre-2008 characteristics. As long as these differences are time-invariant, this selection is accounted for by village fixed effects. Moreover, we show similar pre-trends for firm entry across the treatment and the control group before the 2008 delimitation.

Our estimate indicates the transition of a block from being governed by a single politician to multiple politicians is associated with a 1.1%-1.6% increase in the entry of new firms, and the transition of a block from being governed by multiple politicians to a single politician is associated with a 0.7%-1.0% reduction in firm entry. This result indicates the effect of changing the number of politicians is symmetric. Specifically, we observe an increase in firm entry when a block switches from being unsplit to split and a decrease in firm entry when a block switches from being split to unsplit.

Next, we investigate the mechanism that underlies the positive relationship between number of politicians and firm entry. We propose that multiple politicians may enhance state efficiency by imposing checks and balances on each other. Checks and balances serve as a safeguard against the abuse of power, ensuring that no single entity becomes overly dominant and that the government operates in a manner that ensures accountability and transparency. Consequently, this can lead to lower regulatory obstacles, less cronyism, and improved provision of public infrastructure, creating an economically favorable environment for the entry of new firms.

The effectiveness of checks and balances depends on the degree of collusion among multiple politicians. For example, checks and balances are low when the group of multiple politicians can collude with each other. To assess this mechanism, we use two measures – non-alignment in political parties and non-alignment in caste. Non-alignment in political parties and the caste of these multiple politicians can reduce collusion by increasing the

[3]We note that although the delimitation announcement took place in 2008, the actual changes in assembly constituency boundaries occurred in the subsequent state election. However, employing an event study design using the actual election year after 2008 would violate the no-anticipation assumption. We argue that a part of the impact of the boundary changes is internalized at the time of the announcement, as firms are forward-looking agents. Therefore, an empirical design exploiting the announcement of new electoral boundaries is more appropriate. Section 4.3 provides more details of the analysis.

incentives for deviation or the cost of collusion. For instance, politicians from different parties may find it difficult to find common ground or cooperate due to ideological differences or political animosity (Spenkuch, Teso, and Xu, 2023; Kempf and Tsoutsoura, 2021). Moreover, politicians from different political parties may have stronger incentives to check the dominance of each other. Similarly, differences in caste can reduce collusion due to lower cultural proximity, leading to reduced trust or higher information frictions (Fisman, Paravisini, and Vig, 2017). In essence, the two measures capture the idea that for collusion to occur, the net benefits from such collaboration must outweigh the benefits from deviation.

We test this hypothesis by exploiting the differences in political parties and the caste of politicians in split blocks. We document greater entry of new firms in split blocks where multiple politicians are not aligned in terms of political party or caste lineage. These results suggest that a potential channel through which multiple politicians may positively affect firm entry is by imposing checks and balances on each other, especially when collusion between them is difficult.

We provide further evidence to support the checks and balances hypothesis. Using a close-election regression discontinuity design, we document lower private returns to holding public office – a measure of rent extraction (Fisman, Schulz, and Vig, 2014) – for a politician that has a greater share of the population residing in split blocks. Additionally, we document that areas with multiple politicians exhibit cost-effective provision of roads; greater provision of infrastructure; greater confidence in the local politician and other aspects of local state machinery, such as police and the village governance body; greater confidence in the ability of the state to provide goods such as schools and hospitals relative to the private sector; and a higher likelihood of incumbent's reelection. Lastly, using survey data, we observe that politicians are more actively involved in overseeing the activities of local bureaucrats in split blocks relative to bureaucrats in unsplit blocks within their constituency. Overall, these results suggest that rent-seeking ability is lower in the presence of multiple politicians and they improve the overall state efficiency.

We explore several dimensions of cross-sectional heterogeneity in industry and firm characteristics to evaluate the pathways underlying the higher entry of firms in blocks with multiple politicians. First, we find that multiple politicians increase firm entry by re-

ducing regulatory bottlenecks – manifested through the higher entry of firms in regulated industries. Relatedly, we provide direct evidence showing the projects announced in split blocks take less time to receive regulatory approvals. This finding is further reflected in the higher entry of small firms because regulatory or procedural bottlenecks often act as barriers to the entry of small firms. Second, we find that firms operating in industries less susceptible to cronyism are more likely to enter in split blocks. This suggests that the higher entry of firms in areas with multiple politicians is unlikely to be driven by the grease in the wheel channel. These results indicate that the presence of multiple politicians can influence the costs of entry by enhancing the efficiency of regulatory procedures carried out by local bureaucrats, who are accountable to them.

Related Literature: Our findings contribute to the understanding of the relationship between economic environment and the design of political institutions.[4] We contribute to this literature by focusing on a certain aspect of political institutional design – the presence of multiple politicians. Specifically, we provide well-identified empirical evidence showing that increasing the number of politicians governing an area can add economic value by imposing checks and balances on each other. The importance of checks and balances in political institutions dates back to Montesquieu (1751) and Kant (1795) and has been discussed extensively in the works of Barro (1973), Ferejohn (1986), Persson, Roland, and Tabellini (1997, 2000), Bardhan (2002), Aghion, Alesina, and Trebbi (2004), and Acemoglu, Robinson, and Torvik (2013), among others. We add to this literature by documenting that checks and balances imposed by multiple politicians can serve as a safeguard against the abuse of power. This ensures that no single entity becomes overly dominant and that the government operates in a manner that ensures accountability and transparency.

We contribute to the literature examining the impact of constraints, especially regulatory costs, on limiting firm entry.[5] For instance, Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002) and Klapper, Laeven, and Rajan (2006) find that costly regulations restrict the creation of new firms. On a broader note, Woodruff (2018) presents a detailed review of the literature on the constraints faced by small and growing businesses

[4]There is a large literature on this broad topic. See Alesina and Rosenthal (1996), Tsebelis (1999), Tsebelis (2002), Glaeser, La Porta, Lopez-de Silanes, and Shleifer (2004), Aghion, Alesina, and Trebbi (2004), and Acemoglu and Robinson (2015), among others.

[5]Djankov (2009) presents a detailed discussion of the literature examining the role of regulatory constraints in firm entry.

in the context of developing markets. This paper adds to the literature by documenting a relatively unexplored barrier to firm entry – political checks and balances. We document that political checks and balances may prevent power abuse, ensuring government accountability and transparency. Consequently, this creates an economically favorable environment with lower regulatory obstacles, reduced cronyism, and improved public infrastructure, encouraging new firms to enter the market.

This study also contributes to our understanding of the role of bureaucrats in the efficient functioning of a state (see Besley, Burgess, Khan, and Xu (2022) for a comprehensive review). Bureaucrats are often perceived as exploitative agents who prioritize the interests of powerful industrialists (Tullock, 1965; Peltzman, 1976; Djankov, La Porta, Lopez-de Silanes, and Shleifer, 2002) that can potentially crowd out the entry of other firms (Stigler, 1971; Zingales, 2017). Moreover, politicians can exploit bureaucrats for personal gains by controlling their appointments, transfers, and promotions, often resulting in adverse economic outcomes (Iyer and Mani (2012), Xu (2018), Colonnelli, Prem, and Teso (2020), Akhtari, Moreira, and Trucco (2022), among others). We contribute to the existing body of knowledge by demonstrating that political arrangements involving multiple politicians can act as a system of checks and balances. This arrangement enables effective monitoring of bureaucrats' performance, resulting in better provision of public goods and the establishment of an economic environment conducive to the entry of new firms and overall economic development. Essentially, our findings indicate how multiple politicians can effectively participate in overseeing bureaucrats and improving governance outcomes.

Our paper also contributes to the literature on the common agency problem. This literature has primarily focused on issues of lack of coordination, duplication of effort, and free-rider problems to hypothesize the negative outcomes of multiple principals (see Bernheim and Whinston (1986), Dixit (1997), Dixit, Grossman, and Helpman (1997), Mezzetti (1997), Peters (2001), Laussel and Le Breton (2001), Bergemann and Välimäki (2003), among others). However, we document that multiple principals, in our case multiple politicians, can have a positive effect through better management of their subordinates, in our case local bureaucrats.

Closest to our study is the work of Gulzar and Pasquale (2017). Like them, we use a close border design to compare villages across split and unsplit blocks. Gulzar and

Pasquale (2017) focus on unemployment assistance from India's *National Rural Employment Guarantee Scheme* (NREGA) and find unemployment assistance is lower in areas where bureaucrats answer to multiple politicians, and they argue that this result suggests poor implementation of the program in such areas. This paper contributes by examining the relationship between multiple politicians and economic activity in the private sector – firm entry, employment, and consequently economic development. This allows us to present an alternative perspective to Gulzar and Pasquale (2017). Our results indicate regions with multiple politicians have greater new firm entry and higher employment, manifesting as lower demand for unemployment assistance. We directly validate the lower demand for unemployment assistance by examining the NREGA applications data. We show that 5% lower workday claims or applications for unemployment benefits filed in split blocks relative to unsplit blocks. Therefore, our findings suggest that the development of the private sector can reduce the need for government unemployment assistance. Moreover, we add to the identification strategy of Gulzar and Pasquale (2017) by using the boundary reforms in 2008 to employ a differences-in-discontinuity design allowing us to examine pre-trends.

The remainder of the paper proceeds as follows: Section 2 discusses the institutional background on which the empirical investigation of this paper is based. Section 3 provides a brief description of the data. Section 4 lays down the empirical strategy and the main results. Section 5 provides the mechanism and section 6 concludes.

2 Administrative and Electoral Structure of India

The administrative structure of India is designed as a nested hierarchy of country subdivisions. This hierarchy includes six tiers of government, starting from the federal government of the Republic of India, followed by states, divisions, districts, subdistricts, and finally, the local governance units also referred to as blocks (see appendix Figure A.1). These blocks include smaller towns and villages, which are this paper's primary unit of analysis. The policy implementation at the block level is under the jurisdiction of local bureaucrats who are accountable to local state-level politicians. Indian voters elect representatives at the federal and state level to the Parliament and state assembly, respec-

tively. State-level politicians, referred to as members of the Legislative Assembly (MLA), represent their constituents at the State Legislative Assembly and are elected based on a first-pass-the-post rule. The jurisdiction of local politicians includes all blocks, wholly or partially, contained within the boundaries of their electoral constituency. While policies are collectively designed at the state level, each representative is responsible for its implementation within their constituency through local bureaucrats.

In this section, we will describe this administrative structure of India and in the process highlight the multifaceted relationships among three key actors: entrepreneurs contemplating firm establishment, Block Development Officers (BDOs) wielding regulatory authority, and politicians overseeing local governance structures.

2.1 Local Bureaucrats

The regulatory framework for firm entry in India is characterized by an interplay between national and local governmental bodies. This multi-tiered system, where both central and local authorities have important roles, shapes how businesses are established and operated. The interaction between these different levels of governance affects various aspects of firm entry. For example, business registration, which involves both national and local procedures, takes 68 days and costs around 7% of property value in India. This differs from other OECD high-income economies, which often have different regulatory structures. Similarly, resolving commercial disputes, a process that frequently involves multiple government levels can take more than 3 years through local courts in India—three times longer than in comparable countries. These timelines for registration and dispute resolution, along with associated costs, reflect the complex interactions between different governmental levels in India's business regulatory environment.

National-level requirements for a firm starting a business are the essential certificates that they need to operate like Certificate of Incorporation, PAN, and GST registration. Meanwhile, at the local governments, typically administered by bureaucrats like BDOs, play a crucial role in shaping the business landscape. These local authorities are responsible for issuing key permits such as Shop and Establishment Licenses and Trade Licenses, which significantly impact firm entry and operations. Trade licenses, in particular, serve as

a critical regulatory tool, ensuring businesses comply with local laws and operate within prescribed standards.

BDOs, as key local administrative officials, often have discretionary power in interpreting and applying local regulations. They can impact the speed and ease of obtaining these key permits. They can influence licencing process in several ways – expedite or delay the processing of license applications, interpret local bylaws and regulations, potentially affecting which businesses are allowed to operate in certain areas, have the authority to conduct inspections and enforce compliance, which can impact both new and existing businesses, and provide guidance on local requirements, potentially easing the process for some applicants while creating barriers for others. Moreover, BDOs and other local bodies are instrumental in enforcing sector-specific regulations, such as Food Safety and Standards Authority of India (FSSAI) licenses for food businesses or pollution control clearances for manufacturing units. Their interpretation and application of these regulations can vary, leading to differences in the ease of firm entry across different blocks or districts.

2.2 Electoral Boundaries:

MLA are elected at the assembly constituency level. The boundaries of an assembly constituency are drawn by an independent, non-partisan, and powerful commission known as the Delimitation Commission of India.[6] The explicit goal of the Delimitation Commission is to redraw electoral boundaries such that the population is equalized across all assembly constituencies within a state (GOI, 2004).Although these electoral boundaries were drawn after each decennial census, this process was halted in 1977 until after 2001.[7]

India began the process of redrawing the boundaries of electoral constituencies in 2002, based on the 2001 census of India. On 4 January 2008, the Cabinet Committee on Political Affairs (CCPA) decided to implement the order from the Delimitation Commission following a notice from the Supreme Court of India in December 2007. The President

[6]The voting members of the Delimitation Commission of India include a retired Supreme Court judge, the Chief Election Commissioner of India, and respective state Election Commissioners. Other members include 10 elected representatives from each state acting as associate members, but they do not have voting power on the final decisions of the Commission. The Commission's final decision cannot be challenged in any court of law or be modified by the Parliament or the respective State Legislative Assemblies.

[7]The Constitution (42nd Amendment) Act of 1976 froze all electoral boundaries until after the 2001 Census. This Constitutional amendment was enacted during the period of national emergency in India.

finally approved the order from the Delimitation Commission on 19 February 2008. Iyer and Reddy (2013) argue that the 2008 delimitation resulted in widespread changes to the electoral boundaries since this was the first redrawing of electoral boundaries in three decades, and this process of redistricting, by and large, achieved its goal of population equalization and mainly was politically neutral. Moreover, the Constitution (Eighty-fourth Amendment) Act of 2001 restricts changing of the electoral boundaries until after the year 2026.[8]

2.3 Overlap of Electoral and Administrative Boundaries:

The governance is administered at the block level, while the representatives are elected at the assembly constituency level. The explicit goal of the Delimitation Commission is to redraw electoral boundaries such that the population is equalized across all assembly constituencies within a state. The administrative boundaries of a block are not considered during the design of assembly constituencies. This results in each state being divided by two sets of boundaries - administrative block boundaries and assembly constituency boundaries. Appendix Figure A.2 shows a representative example of the electoral and administrative boundaries of the state of Karnataka.

A block, whether wholly or partially, inside an assembly constituency comes under the jurisdiction of the MLA. Note that although a block can be under the jurisdiction of multiple MLA from different assembly constituencies, individuals can only vote for candidates in one assembly constituency, depending on the village or town they reside in. Hence, the multiple MLAs governing a block do not directly compete with each other in elections. As a result of this overlap, blocks are either under the jurisdiction of multiple politicians (split blocks) or a single politician (unsplit blocks), as shown in appendix Figure A.3. The block-level local bureaucrats are accountable to all the MLAs in charge of the block, regardless of whether a block is wholly or partially included in the assembly constituency. Figure 1 presents a microcosm of our setting using an example of adjacent split and unsplit blocks of Morshi and Warud in the district of Aamravati, Maharashtra. The haphazard overlap of assembly constituency and block

[8]The redrawing of electoral boundaries is closely tied to the schedule of the Indian census. The schedule of the decennial Indian census implies the electoral boundaries implemented in 2008 will be intact until at least 2031.

boundaries generates the variation in the number of politicians governing a block. Figure A.4 presents the geographic distribution of split and unsplit blocks in our sample. Split blocks constitute 40% of blocks in our sample, and do not appear to be geographically concentrated. Split blocks account for 50% of the total 2001 population and 40% of the total geographic area. This is the primary source of heterogeneity we exploit in this paper.

2.4 BDOs and MLAs

The relationship between politicians and bureaucrats is characterized by mutual interdependence, significantly influencing governance and resource allocation. This symbiotic nexus typically operates through two main channels, which are not mutually exclusive.

Firstly, there exists a quid pro quo dynamic between politicians and bureaucrats. In this arrangement, bureaucrats leverage their positions to support aligned politicians in several ways: they may implement policies that favor certain voter groups, they can manage public services to generate goodwill among constituents, and they might use their administrative influence to bolster political support. In return, politicians offer bureaucrats several benefits: enhanced job security, protection from unwanted transfers, and shielding from administrative uncertainties. This reciprocal arrangement aligns with theories of bureaucratic discretion and political patronage, as discussed by (Wilson, 2019; Calvo and Murillo, 2004)

Secondly, the relationship can manifest as collaborative rent-seeking behavior. In this scenario: politicians and bureaucrats work together to extract benefits from public resources or regulatory processes. There is often an implicit agreement to share any illicitly obtained gains. This is in line with bureaucratic corruption and political rent-seeking proposed by (Shleifer and Vishny, 1993; Bardhan, 1997)

3 Data

This section describes the main data sources used for the analysis in this paper.

3.1 Data on Firm Entry

We utilize novel data from the Ministry of Corporate Affairs (MCA), Government of India, covering approximately 750,000 for-profit private firms established between 2003 and 2016. This dataset is used by the Indian government to calculate private sector value-added and is representative of the entire formal private sector. It includes firms' registration dates and precise operational addresses, enabling us to construct a spatio-temporal dataset of private sector activity in India from 2003 to 2016.

To achieve this, we first estimate the geographic coordinates of each pincode using the addresses of corresponding post offices and then assign these coordinates to each firm based on the pincode of their registered address. Appendix B discusses the details of the methodology and potential measurement errors associated with this approach, with the analysis indicating that these errors are unlikely to significantly affect our results.

Subsequently, we aggregate the number of firms at the desired spatial resolution, such as village or unit level, using these spatial coordinates. The spatial and temporal distribution of new firms is presented in Appendix Figure C.1a and Appendix Table C.1.

3.2 Identifying Split Blocks

We obtain the geospatial data on administrative blocks from MLInfomap. We obtain geospatial data on electoral boundaries for the 2008 and the 1977 delimitation from DataMeet. We overlap these shapefiles to identify split and unsplit blocks before and after the 2008 delimitation. Following the logic that a block that is either wholly or partially located in an electoral/assembly constituency is under the jurisdiction of the politician, we define a block as a split block if it is a part of more than one assembly constituency. Specifically, to account for any errors in the drawing of electoral and administrative boundaries, we require each partition to occupy at least 10% of the geographic area of the block.

Appendix Figures A.2 and A.3 present a pictorial representation of the exercise described above for the state of Karnataka. Specifically, Appendix Figures A.2a and A.2b illustrate the administrative and electoral boundaries for Karnataka, respectively. Appendix Figure A.3a shows the overlap between electoral and administrative boundaries

and Figure A.3b presents the split and unsplit blocks formed by the overlap. Figure 2 illustrates a case of adjacent split and unsplit blocks in our sample.

Appendix Figure A.4 presents the geographic distribution of split and unsplit blocks in our sample. Split blocks make up 39.43% of blocks in our sample and do not appear to be geographically concentrated. The split blocks account for 50% of the total 2001 population and 40% of the total geographic area. Moreover, 83.08% of the split blocks are governed by two politicians, three politicians govern 14.83% of split blocks, and more than three politicians govern the remaining 2.09% of the split blocks (see Appendix Figure A.5).

3.3 Exogenous Block Boundaries

We further refine the identification in our RD design to address potential issues of selective sorting around the boundaries separating split and unsplit blocks in two ways: (1) identifying boundaries defined by geographic features - in this case, river segments, and (2) identifying straight-line-like boundaries.

To define boundaries determined by rivers and river basins, we utilize high-resolution geospatial data on river segments in India from HydroATLAS. We identify boundary segments that align with the course of a river segment by overlaying river segments onto block boundaries. Specifically, we require the length of the river segment to be at least 80% of the length of the boundary segment and the river is large enough to be a territorial delimiter; that is, we require the average discharge along the river segment to be greater than or equal to 200 m^3/s or cumecs. This definition results in the inclusion of 3.02% of boundaries and 5.74% of blocks from the original sample. Figure 3 presents representative examples illustrating block boundaries determined by a river segment.

Second, we identif straight-line-like boundaries by calculating the Sinuosity Index (SI) for shared boundary segments between neighboring blocks.[9] Boundaries with an SI value less than 1.05 (more straight-like) are included, totaling 6.24% of boundaries and 10.68% of blocks. Appendices show examples of these boundary types. Appendix

[9]Sinuosity index (SI) is defined as the ratio of the actual length of the shared boundary segment to the Euclidean distance between its endpoints. Border segments with sinuosity closer to 1 are more straight-line-like.

figures C.2 and C.3 present a representative example of a straight-line-like and a squiggly boundary segment, respectively.

3.4 Other Data Sources

NREGA Application Data. We obtain data on individual-level applications filed for unemployment relief under the National Rural Employment Guarantee Scheme (NREGA) from 2016 to 2020 from the Ministry of Rural Development. The data provides information on the number of days for which each applicant seeks unemployment relief. To the best of our knowledge, we are the first to collect and use the application-level NREGA data.

Block Development Officer Survey. We use survey data on the opinions of local bureaucrats – Block Development Officers (BDOs) – from the states of Bihar and Kerela. The states conducted the cross-sectional survey from January-March of 2022. We use three key survey responses. First is the opinion of BDOs on the management by the politicians. The response is noted on a three-point scale where 1 indicates a lower degree of management and 3 indicates a greater degree of management. Second is the opinion of BDOs on their job difficulty. The response is noted on a ten-point scale where 1 indicates a low level of difficulty and 10 indicates a high level of difficulty. Third is the BDO's response on the availability of resources – the number of computers, permanent staff, and temporary staff. Additionally, the survey provides demographic information on BDOs, such as their education, gender, age, experience, and the block's name.

We analyze various other datasets: global nighttime light data from Li, Zhou, Zhao, and Zhao (2020) combining DMSP and VIIRS, demographic info from the 2001 and 2011 Population Census of India, election results, and Pradhan Mantri Gram Sadak Yojana (PMGSY) data from SHRUG (see Asher, Lunt, Matsuura, and Novosad (2021) for discussion on the SHRUG dataset). Additional data sources include the Swacch Bharat Mission (SBM) records, Outkast Python package which is trained on last names and states of 140 million Indians from the 2011 Socio-Economic Caste Census to predict caste of local politicians using their last names and states, IHDS-I and IHDS-II for measuring institutional confidence, Capital Expenditure (CapEx) data from CMIE for regulatory approval

time, and financial details of State Assembly Election candidates from ADR post-2008 delimitation.

4 Impact of Multiple Politicians

The primary objective of this paper is to identify the effect of multiple politicians on local economic growth, in general, and the entry of firms, in particular. The haphazard overlap of administrative block boundaries and the electoral constituency boundaries allows us to identify blocks under the supervision of one politician (unsplit block) and blocks under the supervision of multiple politicians (split blocks). Blocks entirely subsumed within an electoral constituency are classified as unsplit blocks. By contrast, blocks that span more than one electoral constituency are classified as split blocks - split between multiple politicians.

4.1 Empirical Strategy

The empirical strategy hinges on comparing two administrative regions that are similar in all attributes but differ in the number of politicians at the helm of their administrative affairs. To do so, we follow two identification strategies. First, we employ a cross-sectional geographic regression discontinuity design by examining the differences in the outcome variable on either side of the boundary separating a split block and an unsplit block. Second, we employ a differences-in-discontinuity design by exploiting the 2008 delimitation of electoral constituencies. The redrawing of electoral constituency boundaries resulted in converting some unsplit blocks into split blocks and vice versa. The natural experiment of delimitation allows us to examine changes in differences across the border between two blocks in cases where delimitation changes the number of politicians in a block while keeping them fixed in the contiguous block.

4.1.1 Spatial Regression Discontinuity

First, we estimate the causal effect of multiple politicians on the outcomes of interest using a regression discontinuity design. We compare outcomes of interest in a village located

just inside a split block relative to villages on the other side of the boundary, within an unsplit block. We pool the outcome variables in a village across years, since 2008, and estimate the average causal effect of multiple politicians using the following regression discontinuity specification:

$$Y_{v(v\in b(B)} = \gamma \cdot Split_b + f(distance_v) + \beta \cdot X_v + \phi_B + \varepsilon_{v(v\in b)} \ \ for\ v \in bw \tag{1}$$

where $Y_{v(v\in b(B)}$ is the outcome of interest in a village v located within administrative block b belonging to a pair of split and unsplit blocks that share a common boundary B. $Split_b$ is a binary variable that takes a value of 1 if a block spans multiple electoral constituencies and consequently has multiple politicians in charge. X_v refers to a vector of village-level covariates such as population, area, distance to district headquarters, and compactness; ϕ_B is boundary fixed effect; and $f(location_v)$ is the RD polynomial, which controls for smooth functions of the geographic location for village v. Our baseline specification is a local linear polynomial in the distance of the village to the boundary B estimated separately on each side of the boundary, as suggested by Calonico, Cattaneo, and Titiunik (2014), Cattaneo, Idrobo, and Titiunik (2019), and Gelman and Imbens (2019). Following Dell and Querubin (2018), we use a triangular weighting kernel. We estimate our specification for different bandwidths bw, taking values of 5, 10, 20, and 50 km on either side of the boundary. Figure 2 illustrates this geographic RD design.[10] We check robustness to using various other forms of the RD polynomial, kernel, and bandwidths in section 4.2.2.

Our coefficient of interest is γ, the effect of being just inside a split block, and governed by multiple politicians, on our outcome of interest. The interpretation of γ as a causal effect of multiple politicians requires two identifying assumptions. First, all relevant factors before the 2008 delimitation varied smoothly at the block boundaries. This assumption implies the villages located close to the boundary are similar in terms of observables and unobservables before being sorted into split and unsplit blocks in 2008. Second, no selective sorting occurs across the RD threshold; that is villages cannot select in or out of a split block. This assumption ensures a village, within a narrow bandwidth of

[10]Figure 2 illustrates a case of adjacent split and unsplit blocks in our sample. The geographic area shaded in red, is the administrative block of Morshi, in the district of Amaravati, Maharashtra. This block is split between the two assembly constituencies of Morshi and Teosa. The geographic area shaded in blue is the administrative block of Warud, an unsplit block, in the district of Amaravati, Maharashtra.

the boundary, is arbitrarily allocated into a split or an unsplit block. The two assumptions, taken together, ensure the villages, located in unsplit blocks, just outside the boundary of a split block are a valid counterfactual to villages situated just inside the split block.

The two identifying assumptions are likely to be true for our empirical design. As discussed in section 2, a block is defined as either split or unsplit based on the 2008 delimitation, which was largely politically neutral, devoid of gerrymandering, and focused primarily on equalizing populations across electoral constituencies (Iyer and Reddy, 2013). These electoral constituencies are large relative to the villages in the narrow bandwidth. Hence, villages along the boundaries did not have the ability to negotiate whether their blocks would be governed by single or multiple politicians.

Moreover, Table 1 provides empirical support for the smoothness requirement by presenting a balance test across multiple socio-economic dimensions, demonstrating the similarity of villages on both sides of the boundaries. We address concerns of selective sorting by focusing our analysis on a subset of (1) boundaries that are close to a straight line because squiggly boundaries can indicate selective sorting (Polsby and Popper, 1991; Alesina, Easterly, and Matuszeski, 2011), and (2) boundaries defined by salient geographic characteristics, in this case, rivers and river basins.

4.1.2 Time Series Difference in Spatial Discontinuity

Next, we examine the effect of time-series variation in the number of politicians governing a block by exploiting the exogenous delimitation of electoral constituencies in 2008. The natural experiment allows us to control for time-invariant unobservables across villages by including village fixed effects and time-varying heterogeneity across adjacent block pairs through boundary $\times$ year fixed effects. Specifically, the natural experiment allows us to investigate the variations that occur across the border between two blocks when the delimitation process alters the number of politicians governing a block while keeping it unchanged in the neighboring block. We estimate the following differences-in-discontinuity

regression specifications:

$$Y_{v(v\in b(B),t)} = \gamma \cdot (Unsplit \rightarrow Split)_b \times Post_t + f(distance_v) + \beta_v + \phi_{Bt} + \varepsilon_{v(v\in b,t)} \ \forall \ v \in bw \tag{2}$$

$$Y_{v(v\in b(B),t)} = \gamma \cdot (Split \rightarrow Unsplit)_b \times Post_t + f(distance_v) + \beta_v + \phi_{Bt} + \varepsilon_{v(v\in b,t)} \ \forall \ v \in bw \tag{3}$$

In regression specification 2, $(Unsplit \rightarrow Split)_b$ is a binary variable that takes a value of 1 for blocks that transitioned from being unsplit to split between multiple politicians following the 2008 delimitation. We compare these unsplit to split blocks with contiguous blocks that remained unsplit throughout. In regression specification 3, $(Split \rightarrow Unsplit)_b$ is a binary variable that takes a value of 1 for blocks that transitioned from being split to unsplit following the 2008 delimitation. We compare these split to unsplit blocks with contiguous blocks that remained split throughout. $Post_t$ is a dummy variable taking the value of 1 for the years following the 2008 delimitation. ϕ_{Bt} controls for the time-varying boundary fixed effect that allows γ to be estimated from the differences between split and unsplit blocks sharing the same boundary. β_v denotes the village fixed effect which allows the estimation to come from the time-series variation of the same village and consequently the administrative block switching its split status. Finally, the inclusion of the RD polynomial $f(distance_v)$ allows γ to capture the difference in geographic discontinuity over time.

4.2 Spatial Discontinuity Results

This section examines the spatial difference in the propensity of firms to enter and local economic activity between split and unsplit blocks. To do this, we aggregate the primary variables – number of new firms and night light intensity – at the village level for the years post delimitation, i.e. 2008-2016. Figure 4a and 4b present the RD plots for firm entry and nightlight intensity, respectively, with distance to the boundary as the running variable. We observe a discontinuous increase in firm entry and nightlight intensity at the boundary when moving from an unsplit block to a split block.

Table 2 reports estimates for specification 1 using the natural logarithm of 0.001 plus firm entry and nightlight intensity as the dependent variables in Panel A and B, respectively. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4),

respectively. Across all specifications, the coefficient of interest is positive and statistically significant. Our estimates indicate that villages just inside the split block exhibit 3% higher firm entry and 7% higher nightlight intensity relative to villages in unsplit blocks just outside the boundary. The estimate in Panel B is higher as it may represent the aggregate effect of multiple politicians on the formal as well as the informal economy. Appendix Table D.1 reports cross-sectional heterogeneity based on the number of politicians in a split block. The magnitude of the coefficients increases in the number of politicians.

Using census data, we find 6% higher employment in villages just inside a split block than in villages in unsplit blocks just outside the boundary (see Panel A of Table 3). Lastly, we document 5% lower applications for unemployment benefits filed under National Rural Employment Guarantee Scheme (NREGA), suggesting lower demand for unemployment benefits (see Panel B of Table 3). This lower demand for NREGA is consistent with the Gulzar and Pasquale (2017), who show that multiple politicians are associated with lower unemployment benefits.

4.2.1 Exploiting Exogenous Block Boundaries

Although the electoral constituency borders are drawn by a non-partisan committee to equate population across electoral constituencies, there might be a concern that the boundary separating two blocks could be drawn to place certain villages in split or unsplit blocks. We address this concern in two ways, using a subsample of boundaries defined by rivers and a subsample of straight-line-like boundaries.

First, we restrict the sample to boundaries between split and unsplit blocks defined by geographic features, in this case, rivers, and river basins. The intuition of this estimation is that boundaries defined by geographic features are likely to be devoid of selective sorting. Figure 3 presents representative examples illustrating block boundaries determined by a river segment. Panel A of Table 4 reports the results using a sample of units partitioned by boundaries defined by rivers and river basins. The estimate is positive and statistically significant indicating that our baseline results are unlikely to be plagued by selective sorting of villages into split and unsplit blocks.

Second, we restrict estimation to a sample of split and unsplit blocks partitioned by straight-line-like borders. Appendix figures C.2 and C.3 present a representative exam-

ple of a straight-line-like and a squiggly boundary segment, respectively. The intuition behind this estimation is that straight-line-like boundaries are more likely to be arbitrary than squiggly ones as discussed in Polsby and Popper (1991) and Alesina, Easterly, and Matuszeski (2011). We define a boundary as a straight line if the Sinuosity Index (SI) measure for the boundary is less than 1.05. Panel B of Table 4 reports the results restricting sample to units partitioned by boundaries that follow a straight line. We find qualitatively similar results using the sub-sample of straight-line-like boundaries.[11]

4.2.2 Robustness

We conduct various robustness tests to ensure our results are not driven by: a particular econometric specification, specific sample or transformation of the dependent variable, differences across villages along the boundary and covariates, firm exit, spatial auto-correlation, and spurious correlation.

First, we verify our results are robust to alternative econometric specifications. Appendix Table E.1 shows the results are robust to using a local quadratic polynomial instead of a local linear polynomial. The results are robust to using a uniform kernel as shown in appendix Table E.2. Appendix Table E.3 shows that the results are robust to modification of $f(.)$ in specification 1 to be a bivariate function of the latitude and the longitude of the village. Appendix Figure E.1 shows the results are robust to choosing any bandwidth between 2 km and 50 km. Appendix Table E.4 reports that the results are robust to inference based on Conley (1999) standard errors, which allows for spatial dependence of an unknown form, with cutoff distances of 2 km, 5 km, 10 km, and 20 km. Moreover, our inference is robust to alternative clustering at village, district, boundary, and state level as shown in Appendix Table E.5.

Second, we verify that the results are not an artifact of the sample and transformation of the key dependent variable. Appendix Table E.6 shows the results are robust to dropping small and large blocks, indicating the outliers are unlikely to drive our results.[12]

[11] Note that the magnitude of the estimate using both the river-defined and straight-line-like boundaries is greater than the magnitude reported in Panel A of Table 2 indicating the bias is likely to be positive. This suggests that selective sorting may be under-biasing the true effect.

[12] A block is defined as a small block if the total number of firms entering the block during our sample period is less than 0, 5, 10, and 20. A block is defined as a large block if the total number of firms entering the block during our sample period is greater than 1000.

Panel A of appendix Table E.7 shows results are robust to re-estimating our baseline specification using Poisson regression.[13]. Panel B of appendix Table E.7 shows results are robust to using inverse hyperbolic sine (IHS) transformation of the key dependent variable, in alternative samples, instead of $LN(0.001 + x)$ transformation.[14] In appendix Table E.8, we verify that the results are not driven by aggregating observations across time at the village level, by running a non-pooled regression using village-year as the unit of observation. Lastly, we verify that the results are not driven by using the village as the primary unit of analysis. Villages are a predetermined unit of aggregation and could potentially influence our results. For robustness, we create arbitrary grids of cell size - 1×1, 2×2, and 5×5 and aggregate data on firm entry at the cell level. Appendix Table E.9 shows the results are robust to using these arbitrarily defined cells.

Third, we address the concern that the villages located close to the border on either side of the boundary can be different in characteristics, and thereby, the observed impact of having multiple politicians would just be an artifact of such difference in characteristics. Our first defense to this concern is in Table 1 where we show various pre-existing village specific characteristics such as population, education, health, infrastructure, and geography are largely similar across villages along the border. Moreover, we document that firm entry and nightlights vary smoothly across these boundaries before the 2008 delimitation (see appendix Table E.10). Next, we include several covariates in the regression specification 1 to account for several other differences across villages and blocks. These covariates include the 2001 Census population, geographic area of the village, the geographic distance of the village from the district headquarters, and block compactness measured as in Harari (2020).[15] Appendix Table E.11 shows the results are robust to inclusion of these covariates. Another potential concern with our analysis is that the bordering split and unsplit blocks may systematically differ by the quality of the politician. We address this issue by augmenting our baseline specification with constituency fixed effects. Appendix

[13] Chen and Roth (2023) argue that manipulations of the variables such as adding a number before log transformation may affect the empirical results and suggest performing Poisson regression.

[14] The IHS transformation has been used to overcome problems in regression analyses with right-skewed censored dependent variables like ours (Carboni, 2012) and to address the issue of log transformation with zeros (Burbidge, Magee, and Robb, 1988).

[15] Bardhan (2002) and Michalopoulos and Papaioannou (2014) highlight the role of a region's distance to capital as a determinant of growth. Harari (2020) highlights the impact of a city's shape on growth - the more compact the shape of the city, the greater is the growth potential.

Table E.12 reports these results and finds similar results, indicating differences in politician quality are unlikely to drive our results.

Fourth, we verify that the firms that enter split blocks are not more likely to exit. This test addresses the concern that additional firms entering split blocks are of poor quality and hence less likely to survive. We examine the probability of a firm exiting within six months, one year, and two years of its entry as a function of the block being split or unsplit. Appendix Table E.13 presents these results. The estimate of interest is negative and mostly statistically insignificant. This indicates that firms entering in split blocks are not more likely to exit than firms entering in unsplit blocks. Hence, we can rule out the possibility that firms entering unsplit blocks are of systematically lower quality and less likely to survive.

Fifth, we show our main regression is capturing an effect that only appears as we cross the actual boundary between split and unsplit blocks. As a falsification test, we rerun the analysis using arbitrary borders. In particular, we draw fake borders inside each block in our sample such that these borders divide a block into three concentric zones of equal area, as shown in appendix Figure E.2. Appendix Table E.14 presents the results of comparing villages within a block that are only separated by fake borders and have the same number of politicians. The coefficients are insignificant in every case, both economically and statistically. As a placebo test, we randomly define a block as split or unsplit and rerun our analysis using the fake split status assignment. We repeat this exercise and estimate the RD coefficient 10,000 times. Figure E.3 reports these results and finds insignificant results. Hence, the null effects in the falsification and the placebo test indicate our baseline estimation captures an effect that only appears as we cross the actual boundary between split and unsplit block and is unlikely to be spurious or driven by spatial autocorrelation.

4.3 Differences-in-Discontinuity Results

Next, we examine the effect of the exogenous change in the number of politicians governing a block, following the 2008 delimitation by estimating specifications 2 and 3. The specification includes village and boundary $\times$ year fixed effects. The estimate is identi-

fied by comparing villages in the treatment group before and after delimitation with the contiguous villages in the control group, while controlling for boundary-specific shocks to investment opportunities.

Panel A of Table 5 presents estimates from specification 2, focusing on the treatment group of blocks that transitioned from unsplit (single politician) to split (multiple politicians) after the 2008 delimitation. The control group consists of contiguous blocks that remained unsplit throughout. The results show an increase in firm entry by 1.1% - 1.6% when a block transitions from being governed by a single politician to multiple politicians. Panel B of Table 5 reports estimates from specification 3 focusing on the treatment group of blocks that transitioned from split to unsplit after the 2008 delimitation. The control group consists of contiguous blocks that remained split throughout. In contrast to the estimate presented in Panel A, the estimate in Panel B is negative indicating a decline in firm entry by 0.7% - 1.0% when a block switches from multiple politicians to a single politician.

Figure 5, Panels A and B, present the estimates of the dynamic specification associated with equations 2 and 3, respectively. The dynamic plots offer two key takeaways. First, we do not find discontinuity between the treatment and the control group before 2008, indicating an absence of pre-trends. Second, we find evidence of discontinuity between the treatment and the control group after 2008, gradually developing over time.

We note that although the delimitation announcement took place in 2008, the actual changes in assembly constituency boundaries occurred in the subsequent state election. However, employing an event study design using the actual election year after 2008 would violate the no-anticipation assumption. We argue that a part of the impact of the boundary changes is internalized at the time of the announcement, as firms are forward-looking agents. Therefore, an empirical design exploiting the announcement of new electoral boundaries is more appropriate. We verify this by augmenting the previously estimated specification from Table 5 with a Treatment $\times$ Post-Election variable. The results presented in Appendix Table E.15 provide support to our argument.

5 Mechanism

This section aims to investigate the mechanism by which multiple politicians have a positive impact on firm entry. We argue that the presence of multiple politicians enhances state efficiency by imposing checks and balances on each other. These checks and balances increase firm entry by reducing regulatory obstacles and cronyism as well as improving public good provision.

5.1 Checks and Balances

Checks and balances serve as a safeguard against the abuse of power, ensuring that no single entity becomes overly dominant and that the government operates in a manner that ensures accountability and transparency. This section documents that multiple politicians foster firm entry by imposing checks and balances.

5.1.1 Non-Alignment Across Politicians

This section documents that multiple politicians establish checks and balances on each other, leading to favorable conditions for firm entry. We argue that checks and balances are more likely to be relevant force when these politicians are non-aligned. The rationale behind this test is that checks and balances among multiple politicians are most effective when collusion to extract rents becomes difficult either due to high benefits of deviation or high cost of collusion. To this end, we exploit non-alignment in political parties and caste of multiple politicians.

Non-alignment in political parties and the caste of these multiple politicians can reduce collusion by increasing the incentives for deviation or the cost of collusion. For instance, politicians from different parties may find it difficult to find common ground or cooperate due to ideological differences or political animosity (Spenkuch, Teso, and Xu, 2023; Kempf and Tsoutsoura, 2021). Moreover, politicians from different political parties may have stronger incentives to check the dominance of each other. Similarly, differences in caste can reduce collusion due to lower cultural proximity, leading to reduced trust or higher information frictions (Fisman, Paravisini, and Vig, 2017).

We test the checks and balances hypothesis by estimating specification 1 augmented with the interaction term of split blocks and the fraction of non-aligned politicians, boundary × year, and village fixed effects. The unit of observation is village-year. Specifically, the estimation exploits the within-village variation in a split block after 2008; that is, the coefficient of interest is estimated by changes in the fraction of non-aligned politicians within a split block over time while controlling for boundary-specific changes in investment opportunities. Panel A (B) of Table 6 reports results from the estimation exercise in which non-alignment is measured as the difference in political parties (caste) of multiple politicians. The results show split blocks with non-aligned politicians are likely to have greater firm entry than split blocks with all the politicians belonging to the same political party or caste.

5.1.2 Private Returns of Politicians

We further investigate the checks and balances hypothesis by examining the private returns that politicians make through the power of their public office. The intuition of this test is that the checks and balances channel dampens the rent-seeking ability of politicians by ensuring greater accountability and transparency. This tests builds on the prior work of Bhavnani (2012) and Fisman, Schulz, and Vig (2014) who show that a relatively high rate of growth in assets of the winning politician (winner's premium) reflects their rent-seeking behavior. We posit that the presence of checks and balances among multiple politicians reduces winner's premium.

We examine the differences in the winner's premium for politicians across constituencies with a varying share of the population living in split blocks. An essential challenge in estimating the value of the winner's premium is to account for the unobserved skills or resources that politicians possess regardless of their electoral success. Thus, our focus is on closely contested elections where both the winner and runner-up participate in subsequent elections, allowing us to compare the asset growth of candidates who are reasonably similar.

Table 7 presents the findings on the winner's premium in closely contested elections with a margin of victory below 5%. Our main variable of interest is *splitness*, which is examined in both continuous and binary forms. The continuous variable, labeled as

Splitness (Continuous), represents the normalized proportion of the population residing in split blocks within the assembly constituency. The binary variable, named Splitness (Binary), takes a value of one if more than 50% of the constituency population resides in split blocks and zero otherwise. Columns 1 and 2 display results for sub-samples of constituencies with Splitness (Binary) taking a value of zero and one, respectively. We observe a significant 16% increase in the assets of the winner relative to the runner-up in constituencies where Splitness (Binary) is zero. However, we find no statistically significant or positive winner's premium in constituencies where Splitness (Binary) is one. Columns 3 and 4 utilize the complete sample and introduce an interaction between the Winner and the Splitness variable. The negative estimate associated with the interaction term signifies a reduced winner's premium in constituencies with a higher share of split blocks. Overall, our findings indicate lower private returns to public office in constituencies with a significant proportion of split blocks, suggesting the prominence of the checks and balances mechanism.

5.1.3 Multiple Politicians and State Efficiency

This section provides supplementary evidence indicating multiple politicians improve the overall state efficiency, which manifests as better implementation of government programs, greater trust in the state and state machinery, and a higher likelihood of incumbent re-election.

First, we present survey evidence showing multiple politicians improve bureaucratic efficiency through better management. Table 8 presents these results. Using survey data on the opinions of local bureaucrats – Block Development Officers (BDOs) – from the states of Bihar and Kerela, we show that BDOs operating in split blocks report greater management by politicians (MLAs) and their job being more difficult. However, we find no statistically significant differences in resources such as computers and other staff available across split and unsplit blocks.[16] The results suggest greater active management by politicians in split blocks. Since our specification includes assembly constituency fixed

[16]This result is reminiscent of Dasgupta and Kapur (2020). The paper shows that local MLAs aligned with the ruling state party tend to allocate more resources to their respective blocks. However, they also present supplementary results showing that there are no statistically significant differences in resource allocation across blocks based on their level of splitness, according to their preferred specification.

effects, the results suggest the same politician provides active management to the BDO in spilt blocks than the BDO in her constituency's unsplit blocks.

We document higher state efficiency in another government infrastructure program – the *Swachh Bharat Mission* (SBM). SBM is a nationwide program initiated by the Government of India in 2014 to eliminate open defecation and improve solid waste management. Using village-level data on disbursal of public funds under the SBM at the village level, we find a greater disbursal of funds in split blocks than in unsplit blocks (see Appendix Table G.2).

Next, using the survey data from the 2001 and 2011 survey waves of the India Human Development Survey, we document that the greater state efficiency manifests as higher confidence in the local politician (see appendix Table G.3). Additionally, we find households report higher confidence in the local arms of state machinery directly managed by the local politician, such as police and the local governance body referred to as *panchayat*. Moreover, these results are not driven by the popularity of the government at the state level.

Furthermore, we examine the difference in confidence for the same good provided by the private sector and the government – schools and hospitals. The local politician and bureaucrat manage government schools and hospitals, and a direct comparison of these goods with those provided by the private sector can indicate the relative efficiency of the government. Appendix Table G.4 reports these results. This table provides two key takeaways. First, people generally have low confidence in government schools and hospitals to provide education and health care, respectively. However, people in districts with a greater share of split blocks report higher confidence in government schools and hospitals.

Lastly, we find that higher confidence in local politicians and the various public goods provided by them in districts with a higher fraction of split blocks manifests as higher likelihood of incumbent's reelection (see appendix Table G.5). This result implies voters reward local politicians governing alongside other politicians for better management of the state machinery.

5.2 How Do Checks and Balances Increase Firm Entry?

5.2.1 Multiple Politicians Reduce Regulatory Costs

This section examines if the reduction in regulatory costs is a potential driver of the impact of multiple politicians on firm entry. Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002) and Djankov (2009) documents that regulatory costs play a significant role in explaining the disparities in firm entry across different regions. The presence of multiple politicians can influence these costs by enhancing the efficiency of regulatory procedures carried out by local bureaucrats, who are accountable to them.

We establish support for this hypothesis using three tests – exploiting the industry-level heterogeneity in regulatory costs, time taken by private firms to obtain regulatory approvals, and sensitivity to fixed regulatory costs measured by the size of new firms. Overall, the three tests taken together show that lower regulatory costs in areas governed by multiple politicians are a potential driver of higher firm entry in these areas.

First, we exploit the industry-level heterogeneity in regulatory costs by augmenting regression specification 1 with Split × Regulated along with village and boundary × industry fixed effects. Regulated is a binary variable that takes a value of 1 for firms operating in industries subject to high regulatory costs and 0 otherwise.[17] The intuition of this test is that firms facing high regulatory costs are more likely to be affected if regulatory bottlenecks are an important mechanism through which multiple politicians affect firm entry. Panel A of Table 9 reports these results. The results indicate the entry of firms in industries with high regulatory costs is 3.5% greater than the entry of firms operating in industries with low regulatory costs in split blocks relative to unsplit blocks.

Second, we provide supporting evidence that multiple politicians reduce regulatory costs as measured by private firms to receive regulatory approvals for their projects. Using the CapEx database, we compute approval time as the difference between the date when the initial approvals have been received and the date of project announcement. Table 10 reports the results examining the relationship between time for regulatory approvals and the transition of a block from split to unsplit, and vice-versa, following equations 2 and

[17] We define industries with high regulatory costs following the methodology for the classification of regulated industries in the United States based on Pittman (1977) augmented with India specific regulations and factors as discussed in Awasthi, Yayavaram, George, and Sastry (2019). Appendix Table F.1 presents the list of industries classified as industries with high regulatory costs in our sample.

3. Results indicate the time for regulatory approvals for projects decreases (increases) by 80% (57%) when a block switches from being governed by a single politician to multiple politicians (multiple politicians to a single politician).

Third, we exploit the cross-sectional heterogeneity in the size of new firms to examine the effect of multiple politicians on firm entry by size. The intuition of this test is that regulatory costs are more likely to deter the entry of small firms. We group firms into size deciles, with the first (tenth) decile referring to firms in the smallest (largest) size group. We estimate our baseline RD coefficient for each size-decile relative to the 10^{th} decile after including village and boundary $\times$ size decile fixed effects. Figure 6 presents these results. The relative RD estimates monotonically decline with size deciles; that is, the largest effect of the presence of multiple politicians on firm entry is observed among the smallest firms that are likely to find regulatory costs more prohibitive.

5.2.2 Multiple Politicians Reduce Cronyism

The higher entry of new firms in areas governed by multiple politicians could be either driven by greater checks and balances or higher cronyism. Greater checks and balances can lead to increases in efficiency, whereas higher cronyism can act as "*grease in the wheels*" fostering greater entry for crony firms. The lower regulatory costs effect documented in section 5.2.1 is consistent with both the checks and balances and the cronyism channel. We distinguish between the two channels by investigating the effect of the presence of multiple politicians on the entry of crony firms. Lower entry of crony firms would support the checks and balances channel, whereas higher entry of crony firms would support the cronyism channel. Also, investigating the entry of crony firms is important because the quid pro quo between crony firms and politicians can crowd out the entry of other firms (Zingales, 2017).

We exploit the industry-level heterogeneity in seeking rents by engaging in a quid pro quo with the elected politicians. We refer to firms in high rent-seeking industries as crony firms.[18] We augment the regression specification 1 with the interaction term

[18]Specifically, we use the index of industry-level cronyism created by *"Economist"* using the methodology developed by *Transparency International* to classify firms as a high crony and low crony firms. The detailed description of the method to classify industries as crony can be found in the article in The Economist, Planet Plutocrat. Appendix Table F.2 presents the list of high crony industries in the sample.

of crony and split, boundary × industry fixed effects, and village fixed effects. Crony is a binary variable taking a value of 1 if the firm belongs to an industry prone to high crony capitalism. This test compares the entry of high-crony firms with the entry of low-crony firms within the same village and estimates this difference for split blocks relative to unsplit blocks. Panel B of Table 9 reports these results. The estimate of interest is negative and statistically significant across all columns. The results indicate the entry of high-crony firms is 5% lower than the entry of low-crony firms in split blocks relative to unsplit blocks.

The results thus support the premise that greater firm entry is driven by increased efficiency, potentially through improved checks and balances, and not driven by higher corruption greasing the wheels.

5.2.3 Multiple Politicians Improve Provision of Public Goods

The efficiency of state machinery is crucial for firm operations. Businesses rely on state machinery not only to get licenses and permissions but also to improve the overall business climate by providing public goods such as roads and power. Munnell (1992) argues public capital investment in infrastructure can expand the productive capacity of an area by increasing resources and enhancing the productivity of existing resources. We posit that multiple politicians can increase the efficiency of local state machinery, thereby increasing the quantity of public goods and decreasing the time and cost required to provide public goods. We test this hypothesis in the context of public goods that are crucial for a firm's operation and foster the entry of new firms – roads and power supply. Roads and power are two important public capital investments facing severe infrastructure constraints, especially in India.

We investigate the effect of multiple politicians on the efficiency of infrastructure provision by examining a large nationwide road construction program - *Pradhan Mantri Gram Sadak Yojna* (PMGSY).[19] Specifically, we examine the effect of multiple politicians on the cost-effectiveness of road construction. We construct a variable for cost overruns associated with road construction in a village, using estimated and actual cost data for each

[19] We direct the readers to Asher and Novosad (2020) for a detailed discussion on the road construction scheme – PMGSY – and Asher, Lunt, Matsuura, and Novosad (2021) for detailed discussion on the data spanning from 2001 to 2014.

road. Our measure of cost overrun may be driven by three factors. First, local bureaucrats may consistently underestimate the cost of a project, suggesting their incompetency for on-the-ground implementation. Second, local bureaucrats may shirk their responsibilities or siphon off money, contributing to cost overruns.[20] Both of these factors are indicators of the inefficiency of state machinery. Third, idiosyncratic geographic shocks can cause cost overruns in the construction of roads. We control for such shocks in our estimation strategy by including boundary × year fixed effects.

We exploit the 2008 delimitation using the differences-in-discontinuity specification discussed in Section 4.3. Panels A and B of Table 11 report the results from estimating specifications (2) and (3), respectively, with boundary × year and block fixed effects.[21] The estimates in Panel A across all bandwidths are negative and statistically significant. This finding indicates the transition of a block from being governed by a single politician to multiple politicians is associated with a 16% decline in cost overruns on a conservative note. The estimates reported in Panel B are positive and indicate an increase in a cost overrun of 3-5% when a block switches from being governed by multiple politicians to a single politician. Broadly, the results from Table 11 suggest the cost overruns are lower in split blocks indicating the local state machinery is more efficient in the presence of multiple politicians.

We supplement this analysis by examining the provision of another public good - electricity supply for commercial use. We focus on the commercial power supply because it is imperative for firm growth and is controlled by the state. Appendix Table G.1 reports results using high power supply for commercial purposes as the dependent variable.[22] The estimate of interest is positive and statistically significant. The results suggest that a village in a split block is 0.8-1% more likely to receive power for more than 70% of time relative to villages in unsplit blocks.

The results so far support the mechanism whereby checks and balances imposed by multiple politicians on each other increase firm entry by reducing regulatory costs,

[20] Olken (2007) argues that the difference between actual and estimated expenditure reflects the *missing expenditure* and indicates the level of theft by public officials or corruption.
[21] We cannot use village fixed effects as only one road is constructed per village under the PMGSY scheme during our sample period spanning from 2001 until 2014.
[22] We define a village as a high-power-supply village if the village receives commercial power more than 70% of time.

reducing cronyism, and increasing the efficiency and quantity of public good provision most relevant to firms.

6 Conclusion

Political institutions play a vital role in shaping the economy. Hence, understanding what type of political institutions are relatively better at fostering economic growth is of utmost importance. In this paper, we examine a particular feature of political-institutional design – multiple politicians governing an area.

Our empirical investigation is motivated by the theoretical ambiguity surrounding the potential effect of multiple politicians. On one hand, multiple politicians can hurt the local economy due to issues of coordination, free-rider problem, common agency problems, and too many grabbing hands. On the other hand, multiple politicians can improve the local economy by reducing the overall concentration of power, imposing checks and balances on each other, bringing different skills to the table, and the division of labor.

Using a geographic regression discontinuity design across a boundary separating a unit governed by a single politician and multiple politicians, we show that multiple politicians improve the local economy, evidenced by greater firm entry and economic growth. Additionally, we use a differences-in-discontinuity design using delimitation of electoral constituency in 2008 to show firm entry increases when an area transitions from being governed by a single politician to multiple politicians. Furthermore, we find the results are driven by an increase in checks and balances in the presence of multiple politicians, as manifested by a higher impact when they belong to different political parties and castes. We argue that checks and balances imposed by multiple politicians can serve as a safeguard against the abuse of power. This ensures that no single entity becomes overly dominant and that the government operates in a manner that ensures accountability and transparency. Consequently, this can lead to lower regulatory obstacles, less cronyism, and improved provision of public infrastructure, creating an economically favorable environment for the entry of new firms.

A few caveats of our paper are in order. First, our results do not suggest that

multiple politicians solely influence economic growth through the checks and balances channel. There may be other concurrent channels at play, such as the division of labor and the presence of diverse skills. Our study provides evidence for a specific channel, but future research could delve into exploring these alternative channels. Second, our findings do not imply that increasing the number of politicians indefinitely will always lead to a positive effect on economic growth. Our results are bounded by our empirical setting, which allows us to infer the effect of increasing the number of politicians up to a certain point. Beyond this observed range, there might be a possibility that increasing the number of politicians further could yield no additional benefits or even have a negative impact on economic growth.

The results expand our understanding of a specific feature of the political-institutional design that is particularly relevant in understanding the effect of horizontal decentralization, common across several decentralized governance systems. Our results strengthen the faith in the conjecture that imposing checks and balances on agents with authority can result in better governance. Moreover, our results expand our understanding of the costs and benefits of multiple principals and a potential channel through which multiple principals can have a positive impact. The results may be useful in understanding the role of multiple managers that arise in a variety of situations, such as a firm or startup being managed by multiple managers or co-founders, the same entity being regulated by multiple regulatory authorities, doctoral students being advised by multiple chairs, courses being co-taught by multiple instructors, among others.

References

Acemoglu, D., and J. A. Robinson (2015): "The rise and decline of general laws of capitalism," *Journal of Economic Perspectives*, 29(1), 3–28.

Acemoglu, D., J. A. Robinson, and R. Torvik (2013): "Why do voters dismantle checks and balances?," *Review of Economic Studies*, 80(3), 845–875.

Aghion, P., A. Alesina, and F. Trebbi (2004): "Endogenous political institutions," *The Quarterly Journal of Economics*, 119(2), 565–611.

Akhtari, M., D. Moreira, and L. Trucco (2022): "Political turnover, bureaucratic turnover, and the quality of public services," *American Economic Review*, 112(2), 442–493.

Alesina, A., W. Easterly, and J. Matuszeski (2011): "Artificial states," *Journal of the European Economic Association*, 9(2), 246–277.

Alesina, A., and H. Rosenthal (1996): "A theory of divided government," *Econometrica: Journal of the Econometric Society*, pp. 1311–1341.

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

Asher, S., and P. Novosad (2020): "Rural roads and local economic development," *American Economic Review*, 110(3), 797–823.

Awasthi, K., S. Yayavaram, R. George, and T. Sastry (2019): "Classification for regulated industries: A new index," *IIMB Management Review*, 31(3), 309–315.

Bardhan, P. (1997): "Corruption and development: a review of issues," *Journal of economic literature*, 35(3), 1320–1346.

——— (2002): "Decentralization of governance and development," *Journal of Economic Perspectives*, 16(4), 185–205.

Barro, R. J. (1973): "The control of politicians: an economic model," *Public Choice*, pp. 19–42.

Bergemann, D., and J. Välimäki (2003): "Dynamic common agency," *Journal of Economic Theory*, 111(1), 23–48.

Bernheim, B. D., and M. D. Whinston (1986): "Common agency," *Econometrica*, pp. 923–942.

Besley, T., R. Burgess, A. Khan, and G. Xu (2022): "Bureaucracy and development," *Annual Review of Economics*, 14, 397–424.

Bhavnani, R. R. (2012): "Using asset disclosures to study politicians rents: An application to India," in *Annual Bank Conferences on Development Economics, Washington, DC*, vol. 22.

Blanchard, O., and M. Kremer (1997): "Disorganization," *The Quarterly Journal of Economics*, 112(4), 1091–1126.

Burbidge, J. B., L. Magee, and A. L. Robb (1988): "Alternative transformations to handle extreme values of the dependent variable," *Journal of the American Statistical Association*, 83(401), 123–127.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust data-driven inference in the regression-discontinuity design," *The Stata Journal*, 14(4), 909–946.

Calvo, E., and M. V. Murillo (2004): "Who delivers? Partisan clients in the Argentine electoral market," *American journal of political science*, 48(4), 742–757.

Carboni, O. A. (2012): "An empirical investigation of the determinants of R&D cooperation: An application of the inverse hyperbolic sine transformation," *Research in Economics*, 66(2), 131–141.

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

Chen, J., and J. Roth (2023): "Logs with zeros? Some problems and solutions," Discussion paper, Working Paper.

Colonnelli, E., M. Prem, and E. Teso (2020): "Patronage and selection in public sector organizations," *American Economic Review*, 110(10), 3071–3099.

Conley, T. G. (1999): "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92(1), 1–45.

Dasgupta, A., and D. Kapur (2020): "The political economy of bureaucratic overload: Evidence from rural development officials in India," *American Political Science Review*, 114(4), 1316–1334.

Dell, M., and P. Querubin (2018): "Nation building through foreign intervention: Evidence from discontinuities in military strategies," *The Quarterly Journal of Economics*, 133(2), 701–764.

Dixit, A. (1997): "Power of incentives in private versus public organizations," *American Economic Review*, 87(2), 378–382.

Dixit, A., G. M. Grossman, and E. Helpman (1997): "Common agency and coordination: General theory and application to government policy making," *Journal of Political Economy*, 105(4), 752–769.

Djankov, S. (2009): "The regulation of entry: A survey," *The World Bank Research Observer*, 24(2), 183–203.

Djankov, S., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2002): "The regulation of entry," *The Quarterly Journal of Economics*, 117(1), 1–37.

Ferejohn, J. (1986): "Incumbent performance and electoral control," *Public choice*, pp. 5–25.

Fisman, R., D. Paravisini, and V. Vig (2017): "Cultural proximity and loan outcomes," *American Economic Review*, 107(2), 457–92.

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

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

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do institutions cause growth?," *Journal of Economic Growth*, 9, 271–303.

GOI (2004): *Delimitation of Assembly and Parliamentary Constituencies: Guidelines and Methodology*. Delimitation Commission of India, Letter No. 282/DEL/2004, March 17 2004.

Grossman, S. J., and O. D. Hart (1986): "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of Political Economy*, 94(4), 691–719.

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

Harari, M. (2020): "Cities in Bad Shape: Urban Geometry in India," *American Economic Review*, 110(8), 2377–2421.

Hart, O., and J. Moore (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6), 1119–1158.

Hayek, F. A. (1960): *The constitution of liberty*. University of Chicago Press.

Holmstrom, B. R. (1978): *On Incentives and Control in Organizations.* Stanford University.

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

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

Johnson, S., J. McMillan, and C. Woodruff (2002): "Property rights and finance," *American Economic Review*, 92(5), 1335–1356.

Kant, I. (1795): "Perpetual Peace," .

Kempf, E., and M. Tsoutsoura (2021): "Partisan professionals: Evidence from credit rating analysts," *The Journal of Finance*, 76(6), 2805–2856.

Klapper, L., L. Laeven, and R. Rajan (2006): "Entry regulation as a barrier to entrepreneurship," *Journal of Financial Economics*, 82(3), 591–629.

Laussel, D., and M. Le Breton (2001): "Conflict and cooperation: The structure of equilibrium payoffs in common agency," *Journal of Economic Theory*, 100(1), 93–128.

Li, X., Y. Zhou, M. Zhao, and X. Zhao (2020): "A harmonized global nighttime light dataset 1992–2018," *Scientific Data*, 7(1), 1–9.

McMillan, J., and C. Woodruff (2003): "The central role of entrepreneurs in transition economies," *Journal of Economic Perspectives*, 16(3), 153–170.

Mezzetti, C. (1997): "Common agency with horizontally differentiated principals," *The RAND Journal of Economics*, pp. 323–345.

Michalopoulos, S., and E. Papaioannou (2014): "National institutions and subnational development in Africa," *The Quarterly Journal of Economics*, 129(1), 151–213.

Montesquieu, d. S. (1751): "The Spirit of Laws," .

Munnell, A. H. (1992): "Policy watch: Infrastructure investment and economic growth," *Journal of Economic Perspectives*, 6(4), 189–198.

Olken, B. A. (2007): "Monitoring corruption: evidence from a field experiment in Indonesia," *Journal of Political Economy*, 115(2), 200–249.

Olken, B. A., and P. Barron (2009): "The simple economics of extortion: evidence from trucking in Aceh," *Journal of Political Economy*, 117(3), 417–452.

Peltzman, S. (1976): "Toward a more general theory of regulation," *The Journal of Law and Economics*, 19(2), 211–240.

Persson, T., G. Roland, and G. Tabellini (1997): "Separation of powers and political accountability," *The Quarterly Journal of Economics*, 112(4), 1163–1202.

——— (2000): "Comparative politics and public finance," *Journal of Political Economy*, 108(6), 1121–1161.

Peters, M. (2001): "Common agency and the revelation principle," *Econometrica*, 69(5), 1349–1372.

Pittman, R. (1977): "Market structure and campaign contributions," *Public Choice*, 31(1), 37–52.

Polsby, D. D., and R. D. Popper (1991): "The third criterion: Compactness as a procedural safeguard against partisan gerrymandering," *Yale Law & Policy Review*, 9(2), 301–353.

Rasmusen, E., and J. M. Ramseyer (1994): "Cheap bribes and the corruption ban: A coordination game among rational legislators," *Public Choice*, 78(3-4), 305–327.

Shleifer, A., and R. W. Vishny (1993): "Corruption," *The Quarterly Journal of Economics*, 108(3), 599–617.

Spenkuch, J. L., E. Teso, and G. Xu (2023): "Ideology and performance in public organizations," *Econometrica*, 91(4), 1171–1203.

Stigler, G. J. (1971): "The theory of economic regulation," *The Bell Journal of Economics and Management Science*, 2(1), 3–21.

Tsebelis, G. (1999): "Veto players and law production in parliamentary democracies: An empirical analysis," *American Political Science Review*, 93(3), 591–608.

——— (2002): *Veto players: How political institutions work*. Princeton University Press.

Tullock, G. (1965): "The politics of bureaucracy," *Washington: Public Affairs Press.*

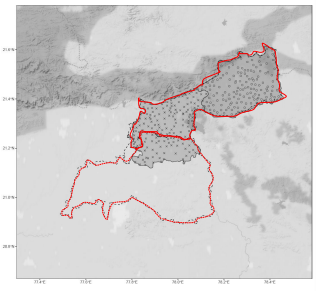
Wilson, J. Q. (2019): *Bureaucracy: What government agencies do and why they do it*. Hachette UK.

Woodruff, C. (2018): "Addressing constraints to small and growing businesses," *International Growth Centre, London*, 20.

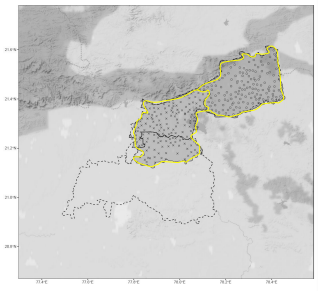
Xu, G. (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*, 108(11), 3170–3198.

Zingales, L. (2017): "Towards a political theory of the firm," *Journal of Economic Perspectives*, 31(3), 113–30.

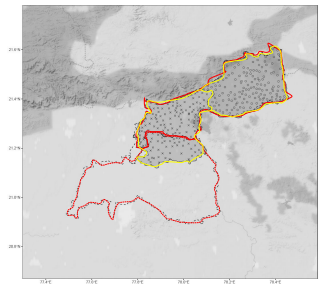
Figure 1: Split and Unsplit Blocks - Overlap of Administrative and Electoral Boundaries



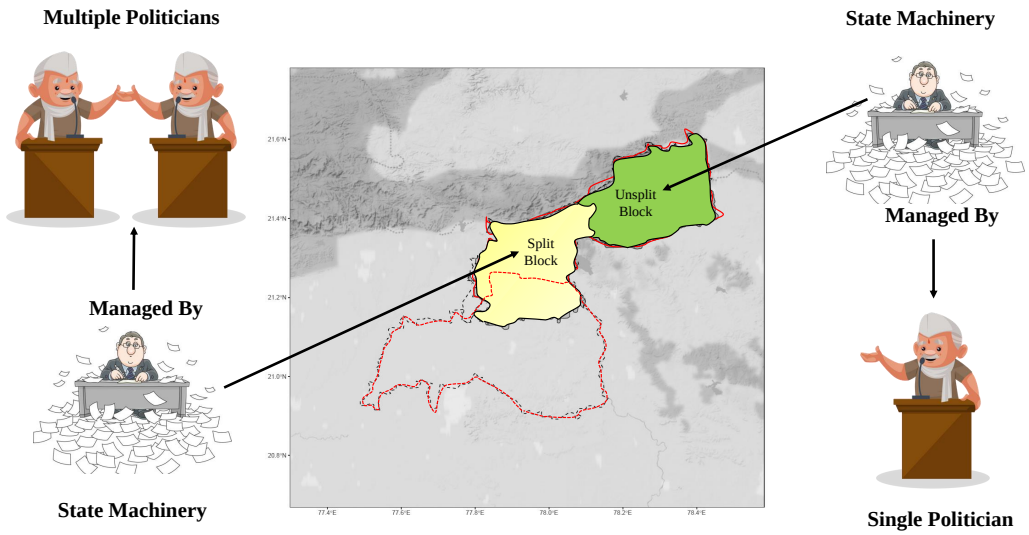
(a) Electoral Boundaries



(b) Administrative Boundaries (Blocks)



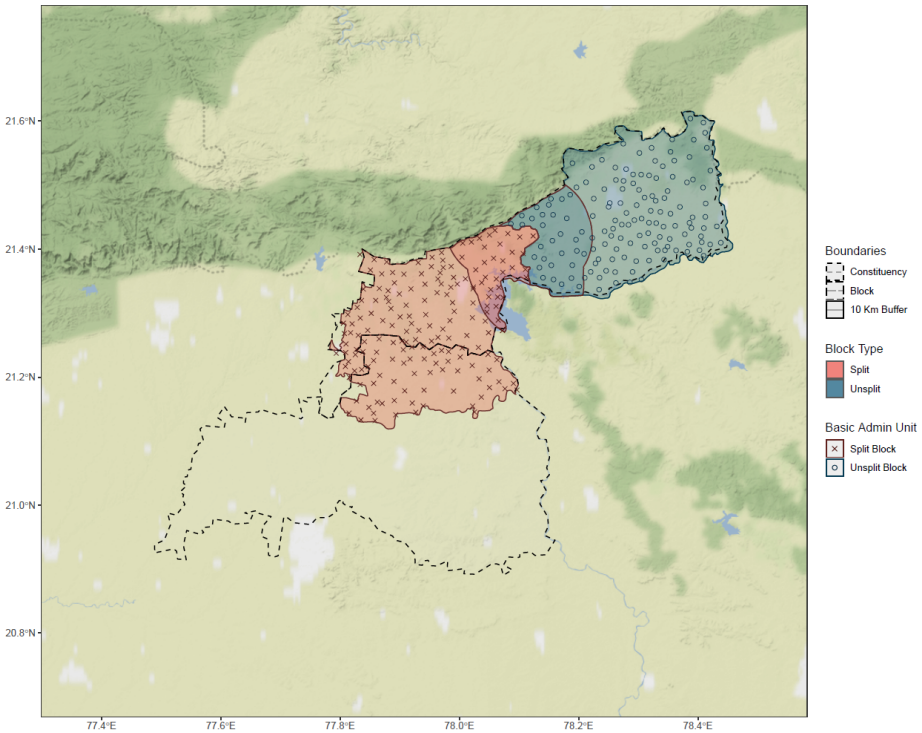
(c) Overlap



(d) Split & Unsplit Blocks

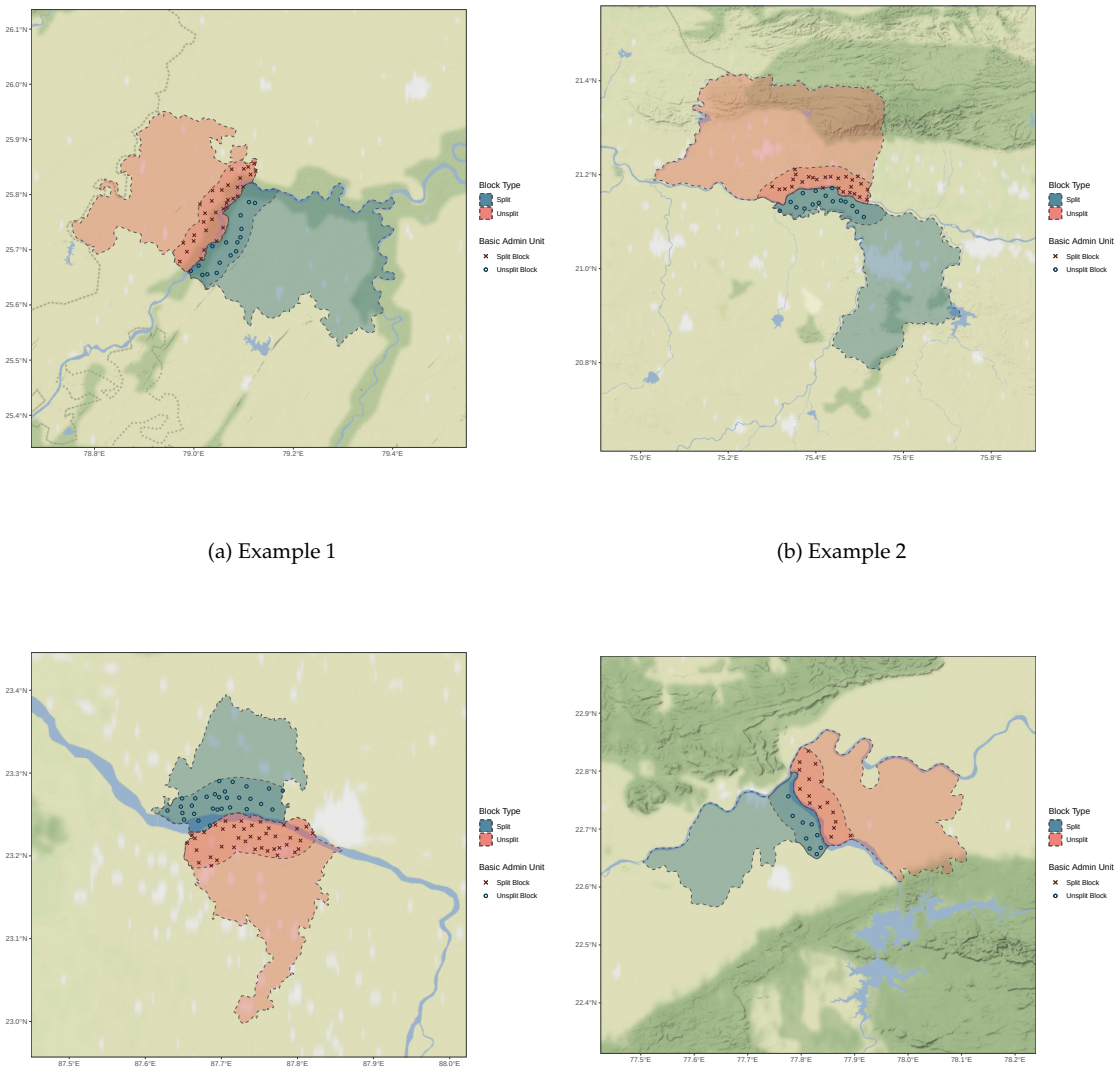
The figure illustrates a case of adjacent split and unsplit blocks in our sample. The split block is the administrative block of Morshi, in the district of Amaravati, Maharashtra. This block is split between the two assembly constituencies of Morshi and Teosa. The neighbouring unsplit block is the administrative block of Warud, in the district of Amaravati, Maharashtra. Figure 1a shows the assembly constituency boundaries of Morshi and Teosa in red. Figure 1b shows the block boundaries of Morshi and Warud in yellow. Figure 1c shows the overlap of assembly constituency boundaries of Morshi and Teosa in red with the block boundaries of Morshi and Warud in yellow. Figure 1d presents a microcosm of our setting showing the formation of split and unsplit blocks due to the overlap of assembly and block boundaries and the state machinery being managed by multiple and single politicians, respectively.

Figure 2: Illustration of Empirical Strategy



This figure illustrates a case of adjacent split and unsplit blocks in our sample. The geographic area shaded in red, is the administrative block of Morshi, in the district of Amaravati, Maharashtra. This block is split between the two assembly constituencies of Morshi and Teosa. The geographic area shaded in blue is the administrative block of Warud, in the district of Amaravati, Maharashtra, and is an unsplit block.

Figure 3: Block Boundaries Delimited by Natural Features



(a) Example 1

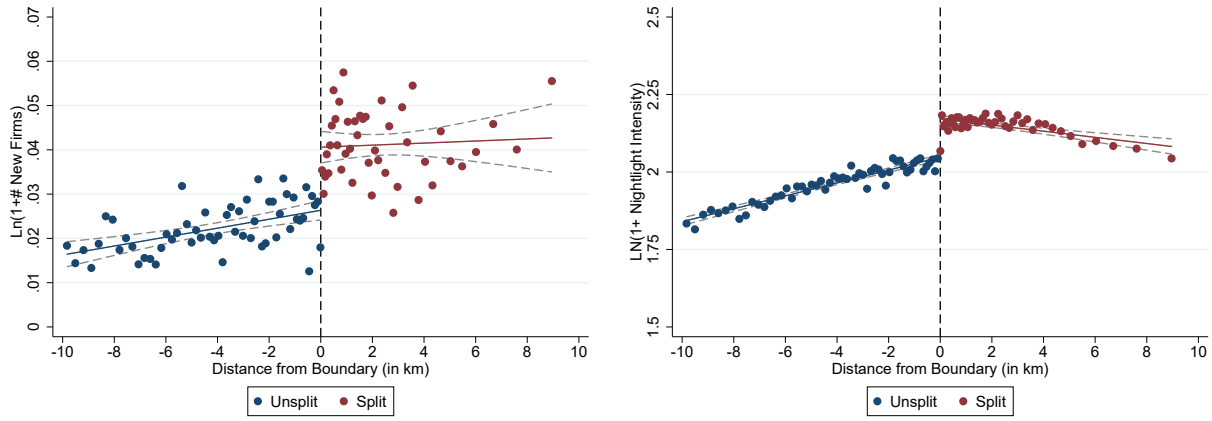
(b) Example 2

(c) Example 3

(d) Example 4

This figure presents four representative examples from the sample of boundaries that are delimited by geographic features, in this case, river segment. Example 1 illustrates River Betwa delimiting the boundary between the blocks, Bamaur and Moth, in the district of Jhansi, Uttar Pradesh. Example 2 illustrates River Tapi delimiting the boundary between the blocks, Jalgaon and Chopda, in the district of Jalgaon, Maharashtra. Example 3 illustrates River Damodar delimiting the boundary between the blocks, Galsi II and Khandaghosh, in the district of Purba Bardhaman, West Bengal;. Example 4 illustrates River Tawa delimiting the boundary between the blocks, Babai and Hoshangabad, in the district of Hoshangabad, Madhya Pradesh.

Figure 4: RD Plots for Firm Entry and Nightlights

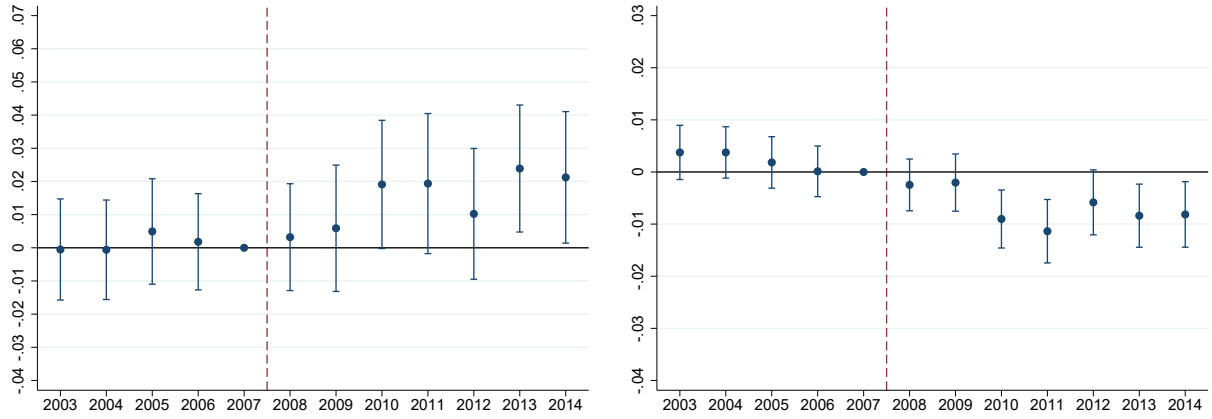


(a) Firm Entry

(b) Nightlights

The figure presents RD plots for our main outcomes and the mean value of each outcome variable at different bins along the running variable (distance to boundary) as well as with a local linear trend estimated separately, along with 95% confidence intervals, on each side of the discontinuity. The unit of observation is a village. The red dots indicate villages in the split blocks and the blue dots indicate villages in the unsplit blocks. Figure 4a presents the RD plots for firm entry. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. Figure 4b presents the RD plots for nightlights. Nightlights is defined as the average nightlight intensity in a village from 2008 to 2016.

Figure 5: Differences-in-Discontinuity Design

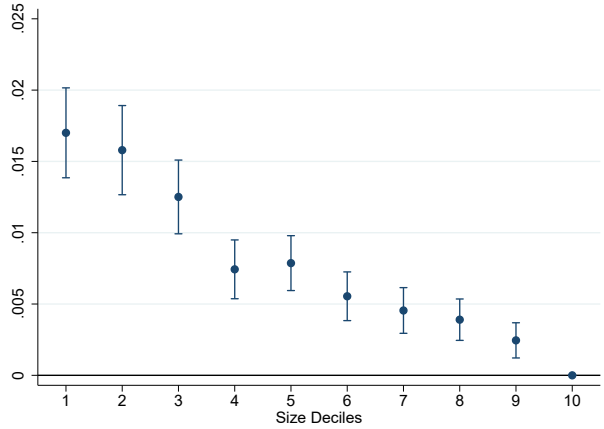


(a) Unsplit → Split

(b) Split → Unsplit

The figure presents the yearly RD estimates between split and unsplit blocks for each year before and after the 2008 delimitation. Panel (a) plots the yearly RD estimates from specification 2 between the treated blocks that switched from being unsplit to split following the 2008 delimitation and the control group that remained unsplit throughout. Panel (b) plots the yearly RD estimates from specification 3 between the treated blocks that switched from being split to unsplit following the 2008 delimitation and the control group that remained split throughout. The unit of observation is a village-year. The dependent variable is the natural logarithm of 0.001 plus the number of new firms in a village during the year. The estimates are plotted with the 95% confidence interval based on clustering the standard errors at the block level. The specification is estimated for the bandwidth of 10 km. All regressions include boundary × year and village fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel.

Figure 6: RD Estimates by Size Deciles



This figure presents the estimates for specification 1 using the natural logarithm of 0.001 plus firm entry as the dependent variable. The estimates are obtained by running the baseline specification 1 augmented to include the interaction terms of size-decile and split binary variable. All estimates are reported relative to the estimate for the firms in 10^{th} decile. We measure firm size based on the book value of assets reported by the firm in the year of entry. We group firms into size deciles with 1^{st} decile referring to firms in the smallest size group, and 10^{th} decile referring to the firms in the largest size group. The unit of observation is a village-size decile. The dependent variable is the natural logarithm of 0.001 plus the number of new firms in a village within a size-decile. The estimates are plotted with the 95% confidence interval based on clustering the standard errors at the block level. The specification is estimated for the bandwidth of 10 km. All regressions include village and boundary $\times$ size decile fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel.

Table 1: Balance on Pre-existing Characteristics

	Full Sample		Split Block (Pre Delim)		Unsplit Block (Pre Delim)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Estimate	SE	Estimate	SE	Estimate	SE
Population						
Ln(0.001 + Population)	0.0153	(0.0251)	0.0278	(0.0301)	0.0833	(0.0840)
Education						
Ln(1 + Schools)	-0.0037	(0.0036)	-0.0027	(0.0044)	0.0178	(0.0139)
Ln(1 + Colleges)	-0.0002	(0.0003)	-0.0003	(0.0004)	-0.0014	(0.0011)
Demographics						
Ln(0.001 + Female)	0.0123	(0.0245)	0.0257	(0.0295)	0.0825	(0.0822)
Ln(0.001 + Lower Caste)	0.0172	(0.0325)	0.0237	(0.0396)	0.1620	(0.1172)
Health						
Ln(1 + Hospital)	-0.0001	(0.0005)	0.0000	(0.0006)	-0.0017	(0.0016)
Ln(1 + Dispensaries)	0.0014	(0.0012)	0.0009	(0.0015)	-0.0005	(0.0027)
Infrastructure						
Ln(1 + Banks)	0.0005	(0.0012)	-0.0007	(0.0013)	0.0038	(0.0070)
Ln(1 + Cinemas)	0.0008	(0.0011)	-0.0012	(0.0011)	0.0047	(0.0068)
Ln(1 + Auditorium)	0.0039	(0.0021)	0.0020	(0.0031)	0.0096	(0.0084)
Geography						
Ln(Area)	-0.0090	(0.0086)	-0.0039	(0.0103)	0.0238	(0.0350)
Ln(Forest Land)	-0.0173	(0.0267)	0.0021	(0.0327)	-0.0263	(0.1101)

This table presents the test for balance on pre-existing characteristics for split and unsplit blocks. Columns (1) and (2) report the RD estimate and the associated standard errors for the full sample. Column (3) and (4) report the RD coefficient on the split variable – defined using the 2008 delimitation – and the standard errors for the sample of blocks that were split before the 2008 delimitation. Column (5) and (6) report the RD coefficient on the split variable – defined using the 2008 delimitation – and the standard errors for the sample of blocks that were unsplit before the 2008 delimitation. The unit of observation is a village that lies within the 10 km bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. The data on the key variables used in the table comes from the 2001 Indian Census. Village-level characteristics are divided into population, education, demographics, health infrastructure, other infrastructure, and geography. Population includes the natural logarithm of the number of people. Education includes the natural logarithm of the number of schools and colleges. Demographics includes the natural logarithm of female, and lower caste population. The latter includes scheduled castes (SC) and scheduled tribes (ST). Health infrastructure includes the natural logarithm of the number of hospitals and dispensaries. Other infrastructure includes the natural logarithm of the number of banks, cinema halls, and auditoriums. Geography includes the natural logarithm of geographic area and area under forest land.

Table 2: RD Estimate: Firm Entry, Nightlights, and Split Block

Panel A: Number of Firms				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split(=1)	0.0331***	0.0326***	0.0312***	0.0309***
	(0.0071)	(0.0060)	(0.0056)	(0.0055)
#Obs	263,307	356,260	416,694	436,980
R^2	0.1537	0.1293	0.1147	0.1062
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Panel B: Nightlights				
Dep Var: LN(0.001+Nightlight)	(1)	(2)	(3)	(4)
Split(=1)	0.0703***	0.0782***	0.0773***	0.0661***
	(0.0085)	(0.0100)	(0.0114)	(0.0127)
#Obs	257,787	348,539	407,616	427,659
R^2	0.6054	0.5736	0.5453	0.5239
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry and nightlight intensity as the dependent variables in Panels A and B, respectively. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm entry is defined as the total number of firms that entered a village from 2008 to 2016. Nightlight is defined as the average nightlight intensity in a village from 2008 to 2016. All regressions include boundary fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: RD Estimate: Employment & Applications for Unemployment Assistance

Panel A: Employment				
Dep Var: LN(0.001+ Employment)	(1)	(2)	(3)	(4)
Split (=1)	0.0595***	0.0597***	0.0599***	0.0587***
	(0.0197)	(0.0180)	(0.0179)	(0.0183)
# Obs	226,010	304,539	355,420	372,431
R^2	0.4420	0.4286	0.4188	0.4107
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Panel B: Applications for Employment Benefits under NREGA				
Dep Var: LN(Days Applied)	(1)	(2)	(3)	(4)
Split (=1)	-0.0622**	-0.0543**	-0.0543**	-0.0561**
	(0.0284)	(0.0256)	(0.0255)	(0.0255)
# Obs	131,620	189,722	228,799	241,962
R^2	0.3454	0.3378	0.3327	0.3325
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus the total number of people employed and the natural logarithm of the total number of days for which unemployment relief is applied for under the National Rural Employment Guarantee Scheme (NREGA) as the dependent variables in Panels A and B, respectively. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. The data on employment comes from the 2013 Economic Survey of India and includes the number of people employed in all non-farm activities. The data on applications for unemployment relief applied under the National Rural Employment Guarantee Scheme (NREGA) comes from the Ministry of Rural Development and spans from 2016 to 2020. All regressions include boundary fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: RD Estimate: Using Exogenous Block Boundaries

Panel A: Boundaries defined by Geographic Features (Rivers & River Basins)				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0965**	0.0913***	0.0819***	0.0777***
	(0.0428)	(0.0343)	(0.0309)	(0.0295)
# Obs	7,137	9,804	11,372	11,792
R^2	0.0823	0.0711	0.0618	0.0562
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Panel B: Boundaries defined by Straight Lines				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split(=1)	0.1763***	0.1360***	0.1106***	0.1076***
	(0.0606)	(0.0485)	(0.0406)	(0.0379)
#Obs	7,018	10,394	13,426	14,328
R^2	0.2012	0.1698	0.1471	0.1331
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for specification 1 using the natural logarithm of 0.001 plus firm entry as the dependent variables. Panel A restricts the sample to boundaries between split and unsplit blocks that are straight lines. We define a boundary as a straight line if the Sinuosity Index (SI) is less than 1.05. Sinuosity Index is defined as the ratio of the length of the curve and the Euclidean distance between the endpoints of the curve. Panel B restricts the sample to boundaries between split and unsplit blocks defined by geographic features – rivers and river basins. We only consider rivers and river basins with an average discharge greater than or equal to 200 m^3/s or cumecs. We describe a boundary to be defined by a river or river basin if the boundary follows a river for at least 80% of its length. Appendix figure 3 provides four representative examples of block boundaries separating a split and unsplit block that are defined by rivers or river basins. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Firm Entry, Change in the Number of Politicians, and 2008 Delimitation

Panel A : Unsplit to Split				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Treat x Post	0.0157**	0.0136**	0.0123**	0.0113**
	(0.0074)	(0.0062)	(0.0055)	(0.0051)
# Obs	687,792	975,312	1,185,816	1,263,456
R^2	0.6489	0.6329	0.6203	0.6129
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes
Panel B : Split to Unsplit				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Treat x Post	-0.0104***	-0.0086***	-0.0074***	-0.0073***
	(0.0023)	(0.0020)	(0.0018)	(0.0018)
# Obs	2,469,408	3,297,648	3,812,364	3,978,144
R^2	0.6888	0.6861	0.6817	0.6769
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (2) and (3) using the natural logarithm of 0.001 plus firm entry as the dependent variable in Panels A and B, respectively. In Panel A the treated group is the set of blocks that switched from being unsplit to split following the 2008 delimitation, whereas the control group comprises of blocks that were always unsplit both before and after the delimitation and border the treated group. In Panel B, the treated group is the set of blocks that switched from being split to unsplit following the 2008 delimitation, whereas the control group comprises blocks that are always split among multiple politicians both before and after the 2008 delimitation and border the treated group. The variable *Post* takes a value of 1 for all years since 2008. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating the treatment and the control groups in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year that lies within the narrow bandwidth of the boundary separating the treatment and the control group. Firm entry is defined as the total number of firms that have entered a village during the year. All regressions include village and boundary × year fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Political Alignment and Firm Entry

Panel A: Political Alignment				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Frac. Non Aligned x Split	0.0053	0.0094**	0.0114***	0.0117***
	(0.0056)	(0.0045)	(0.0041)	(0.0039)
# Obs	1,813,190	2,497,757	2,958,673	3,120,412
R^2	0.7637	0.7581	0.7521	0.7466
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes
Panel B: Caste Alignment				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Frac. Non Aligned x Split	0.0122	0.0157**	0.0148***	0.0142***
	(0.0083)	(0.0070)	(0.0057)	(0.0052)
# Obs	1,387,788	1,917,496	2,284,232	2,422,849
R^2	0.7468	0.7447	0.7402	0.7352
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) augmented for the interaction term of split and the fraction of non-aligned politicians using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms entering a village in that year. The time period spans 2008 to 2016. Panel A of Table 6 reports result from the estimation exercise where non-alignment is measured as difference in political parties. Panel B of Table 6 measures non-alignment in terms of caste. The fraction of non-aligned politicians is a continuous variable taking a value between 0 and 1. It takes a value of 0 when all politicians are affiliated with the same political party or the same caste, and a value of 1 when all politicians are affiliated with different political parties or different castes. It always takes a value of 0 for unsplit blocks. All regressions include boundary × year fixed effect, village fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Winner's Premium and Multiple Politicians

Dep Var: Ln(Final Asset)	(1)	(2)	(3)	(4)
Winner	0.1605**	-0.0800	0.1589*	0.0579
	(0.0809)	(0.0921)	(0.0811)	(0.0607)
Winner X Splitness (Binary)			-0.2348*	
			(0.1226)	
Winner X Splitness (Continous)				-0.1413**
				(0.0615)
Ln(Initial Asset)	0.6112***	0.7224***	0.6580***	0.6583***
	(0.0630)	(0.0764)	(0.0485)	(0.0482)
Constituency FE	Yes	Yes	Yes	Yes
# Obs	330	252	582	582
R^2	0.8195	0.824	0.8233	0.8243
Sample	Splitness (Binary) = 0	Splitness (Binary) = 1	All	All

This table relates the difference in asset accumulated over an election cycle by a winner vis-a-vis runner up to the degree of splitness of an electoral constituency. Splitness (Binary) is defined as a binary variable taking a value of one if more than 50% population in the constituency resides in split block and zero otherwise. We also measure Splitness as a continuous variable: defined as the (normalized) proportion of the assembly constituency population that resides in split blocks. In column (1) and (2) we report the results from the following regression specification performed separately for electoral constituency where splitness is zero and one respectively: $ln(FinalAsset)_{ck} = \beta_1 \times Winner_{ck} + \beta_3 ln(InitialAsset)_{ck} + \beta_c + \epsilon_{ck}$. In columns (3) and (4) we run the following regression specification using all electoral constituencies: $ln(FinalAsset)_{ck} = \beta_1 \times Winner_{ck} + \beta_2 \times Winner_{ck} \times Splitness_c + \beta_3 ln(InitialAsset)_{ck} + \beta_c + \epsilon_{ck}$. In column (3) splitness is defined as a binary variable taking a value of one if more than 50% of the constituency population resides in split block. In column (4) splitness is continuous variable, defined as the normalized proportion of constituency population that resides in split blocks. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effect of Multiple Politicians on Block Development Officers

	(1)	(2)	(3)	(4)	(5)
	Greater Management by Politicians	Job Difficulty	Resources		
			# Computers	# Permanent Staff	# Temporary Staff
Split(=1)	0.6574**	0.8014**	0.0559	0.0047	-0.0495
	(0.3144)	(0.3333)	(0.0711)	(0.0294)	(0.1562)
Assembly FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Model	Ordered Probit	Ordered Probit	Poisson	Poisson	Poisson
# Obs	360	365	354	363	357
Psuedo R^2	0.5697	0.3982	0.2582	0.6799	0.6547

This table presents estimates the differences in the management and the resources for Block Development Officers (BDO's) across split and unsplit blocks. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Column (1) uses the response of the opinion of BDO's on the management by the politicians. The response is noted on a three point scale 1 to 3 where 1 indicates lower management and 3 indicates greater management. Column (2) uses the response of the opinion of BDO's on their job difficulty. The response is noted on a ten point scale 1 to 10 where 1 indicates low difficulty and 10 indicates high difficulty. Column (3)-(5) uses the BDO's response on the availability of resources. Columns (3)-(5) use the number of computers, number of permanent staff, and the number of temporary staff as the dependent variable. Columns (1) and (2) estimate the relationship using an ordered probit model. Columns (3)-(5) estimate the relationship using a Poisson model. All regressions include assembly constituency fixed effects and BDO specific controls. Controls include education fixed effects, gender fixed effects, age, and experience. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Firm Entry by Industry Type

Panel A: Industries with High Regulatory Costs				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Regulated x Split(=1)	0.0349*	0.0346**	0.0359**	0.0343**
	(0.0208)	(0.0172)	(0.0156)	(0.0150)
#Obs	328,104	419,364	465,966	478,656
R^2	0.3623	0.3574	0.3538	0.3524
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Industry FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes
Panel B: Crony Industries				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
High Crony x Split(=1)	-0.0494**	-0.0607***	-0.0599***	-0.0578***
	(0.0210)	(0.0170)	(0.0155)	(0.0150)
#Obs	328,104	419,364	465,966	478,656
R^2	0.3623	0.3574	0.3538	0.3524
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Industry FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

This table presents estimates for specification 1 augmented for the interaction term of split with regulated and high crony industries in Panels A and B, respectively. The dependent variable is the natural logarithm of 0.001 plus firm entry. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is the village-industry that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm entry is defined as the total number of firms in an industry entering a village from 2008 to 2016. All regressions include boundary × industry fixed effect, village fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. We define industries with a high regulatory cost following the methodology for the classification of regulated industries in the United States based on Pittman (1977) augmented with India specific regulations and factors as discussed in Awasthi, Yayavaram, George, and Sastry (2019). Appendix Table F.1 presents the list of industries classified as industries with high regulatory costs in our sample. We use the index of industry-level cronyism created by *Economist* using the methodology developed by *Transparency International* to classify firms as high crony and low crony. The detailed description of the method to classify industries as crony can be found in the article in The Economist titled: Planet Plutocrat. Appendix Table F.2 presents the list of high crony industries in the sample. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Multiple Politicians and Time for Regulatory Approvals

LN(Days to obtain approval)	(1)	(2)
Treat x Post	-0.8084***	0.5706*
	(0.2122)	(0.3461)
Sample	Unsplit to Split	Split to Unsplit
# Obs	2284	9091
R^2	0.2449	0.2745
Bandwidth	10 KM	10 KM
Boundary x Year FE	Yes	Yes
Block FE	Yes	Yes

This table presents estimates for specification (2) and (3) using the natural logarithm of "Days taken to obtain regulatory approvals" as the dependent variable. In column (1) the treated group is the set of blocks that switched from being unsplit to split following the 2008 delimitation, whereas the control group comprises blocks that were always unsplit both before and after the delimitation and border the treated group. In column (2), the treated group is the set of blocks that switched from being split to unsplit following the 2008 delimitation, whereas the control group comprises of blocks that are always split among multiple politicians both before and after the 2008 delimitation and border the treated group. The variable *Post* takes a value of 1 for all years since 2008. We use a bandwidths of 10 km on either side of the boundary, separating the treatment and the control groups. The unit of observation is a project, announced between 2003 and 2016 that lies within the narrow bandwidth of the boundary separating the treatment and the control group.. All regressions include block and boundary × year fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Multiple Politicians and Cost Overruns in Road Construction

Panel A : Unsplit to Split				
Dep Var: $LN(\frac{ActualCost}{EstimatedCost})$	(1)	(2)	(3)	(4)
Treat × Post	-0.2586***	-0.1947**	-0.1658**	-0.1647**
	(0.0980)	(0.0925)	(0.0821)	(0.0770)
# Obs	5,819	8,754	11,124	11,897
R^2	0.7154	0.6898	0.6646	0.6492
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes
Panel B : Split to Unsplit				
Dep Var: $LN(\frac{ActualCost}{EstimatedCost})$	(1)	(2)	(3)	(4)
Treat × Post	0.0313	0.0514*	0.0532**	0.0538**
	(0.0316)	(0.0269)	(0.0230)	(0.0227)
# Obs	22,796	32,363	38,527	40,638
R^2	0.6713	0.6462	0.6319	0.6192
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes

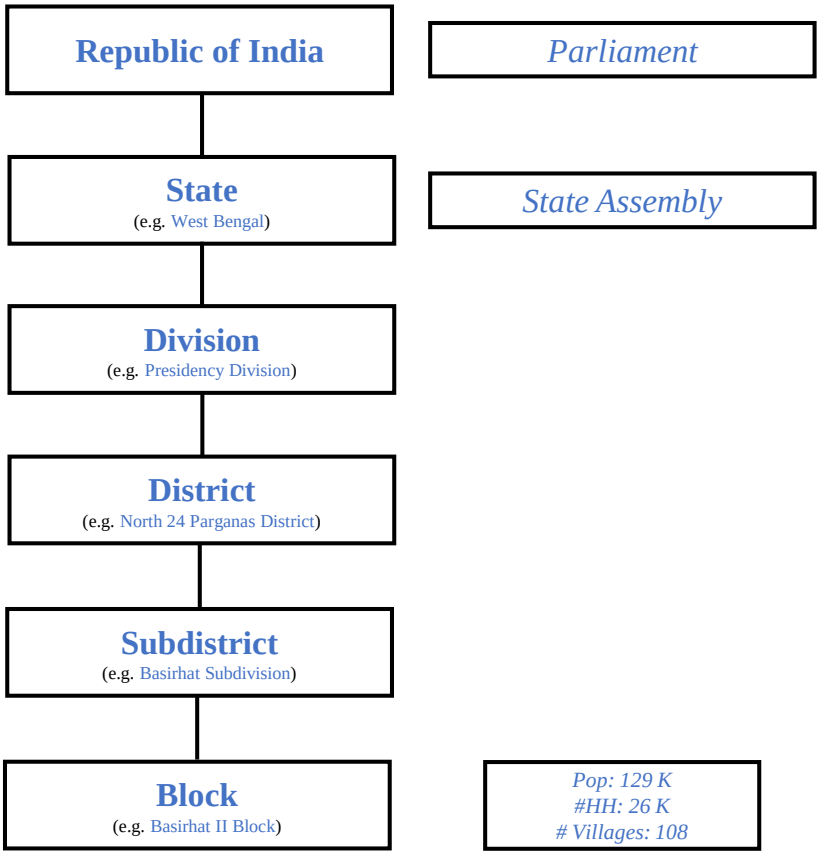
This table presents estimates for specification (2) and (3) using the natural logarithm of actual cost minus the natural logarithm of estimated cost as the dependent variable in panels A and B, respectively. In Panel A the treated group is the set of blocks that switched from being unsplit to split following the 2008 delimitation, whereas the control group comprises of blocks that were always unsplit both before and after the delimitation and border the treated group. In Panel B, the treated group is the set of blocks that switched from being split to unsplit following the 2008 delimitation, whereas the control group comprises of blocks that are always split among multiple politicians both before and after the 2008 delimitation and border the treated group. The variable *Post* takes a value of 1 for all years since 2008. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating the treatment and the control groups in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year spanning from 2001 to 2014 that lies within the narrow bandwidth of the boundary separating the treatment and the control group. A village appears only once in the dataset because only one road could be constructed under the PMGSY per village. All regressions include block and boundary × year fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix for

Political Power-Sharing, Firm Entry, and Economic Growth: Evidence from Multiple Elected Representatives

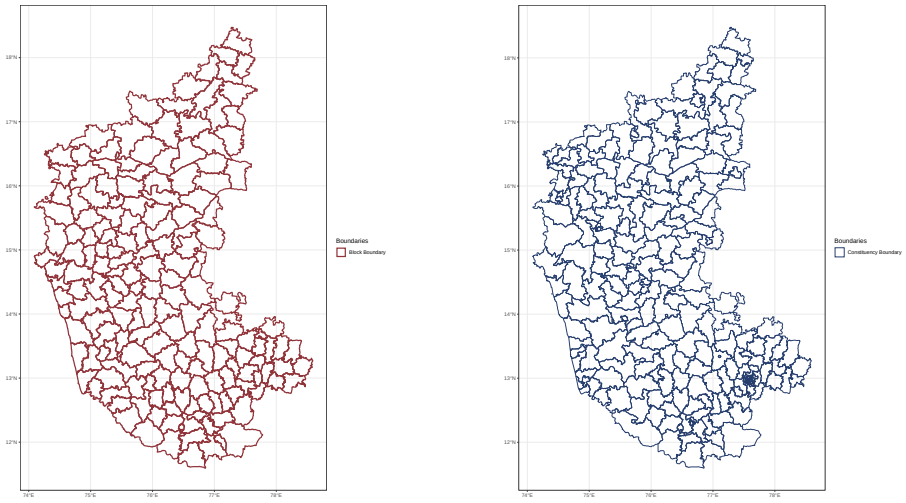
Appendix A Background & Institutional Details

Figure A.1: The Six Tiers of Government in India



The figure illustrates the six tiers of government in India. This hierarchy includes six tiers of government, starting from the federal government of the Republic of India, followed by states, division, district, sub-district, and finally, the local governing bodies. We refer to the local governing bodies in urban and rural areas as blocks which include towns, villages, or clusters of towns and villages. India has a parliamentary system as defined by its constitution, with power distributed between the central government (Parliament) and the states (State Assembly).

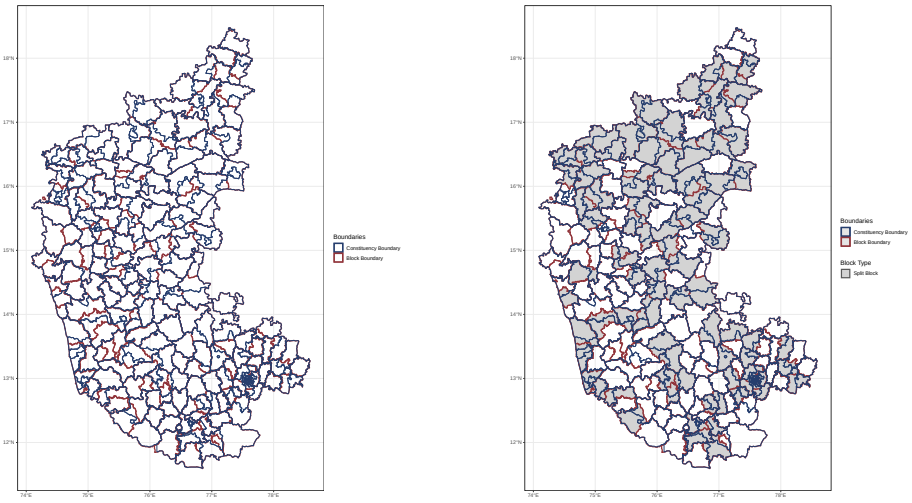
Figure A.2: Administrative and Electoral Boundaries



(a) Administrative Boundaries (Blocks) (b) Electoral Boundaries (Assembly Constituency)

The figure presents the administrative and electoral boundaries, as an example, from the state of Karnataka. The administrative boundaries shown in Figure A.2a in red are based on the basic administrative unit - block. The electoral boundaries shown in Figure A.2b in blue are based on the 2008 delimitation.

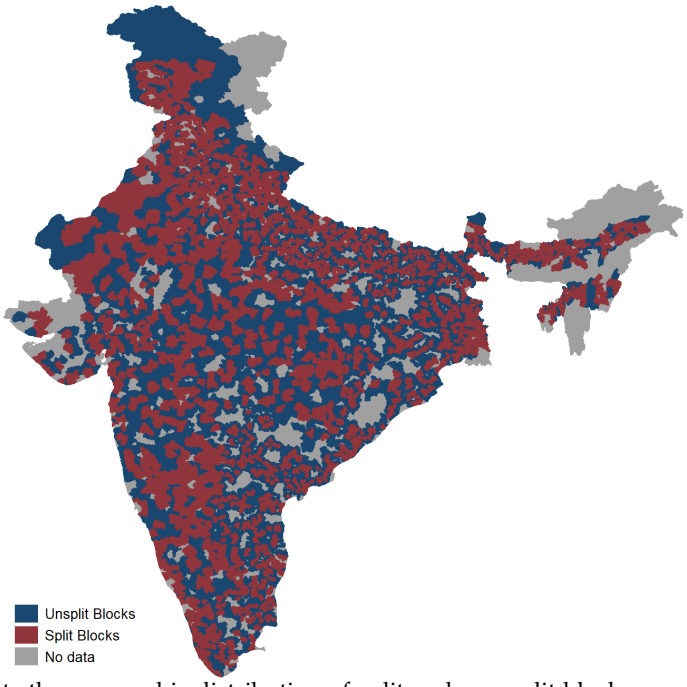
Figure A.3: Overlap of Administrative and Electoral Boundaries



(a) Overlap of Boundaries (b) Split & Unsplit Blocks

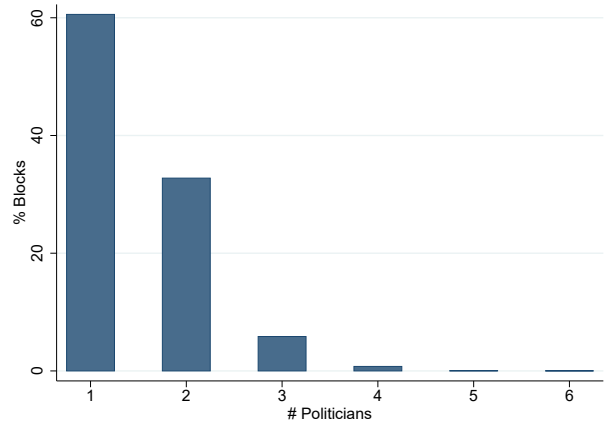
The figure presents the overlap of the administrative and electoral boundaries, as an example, from the state of Karnataka and the formation of split and unsplit blocks. The administrative boundaries in red are based on the basic administrative unit - block. The electoral boundaries shown in blue are based on the 2008 delimitation. Figure A.3a shows the overlap of the electoral and administrative boundaries. Figure A.3b presents a graphic depiction of split blocks in grey and unsplit blocks in white.

Figure A.4: Geography of Blocks Administered by Multiple Politicians



This figure plots the geographic distribution of split and non-split blocks across all blocks in India. A block is defined as a split block if it is split across two or more state legislative assembly constituencies, with at least two constituencies, each accounting for at least 10% of the geographic area of the block based on 2008 assembly constituency boundaries.

Figure A.5: Distribution of Blocks by Number of Governing Politicians



The figure illustrates the distribution of blocks by the number of governing politicians. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation.

Appendix B Extracting Firm Location

This section outlines the methodology used to approximate the locations of companies registered with the Ministry of Corporate Affairs (MCA) in India. We also discuss potential biases associated with the approximation algorithm and provide robustness checks to validate our baseline estimates against these concerns.

B.1 Geolocating Companies Registered with the MCA

We follow a three-step process to geolocate companies registered with the Ministry of Corporate Affairs (MCA):

- **Step 1:** Extract the pincode from the company's registered address as provided in the MCA dataset.[23]
- **Step 2:** Estimate the approximate geographic coordinates of each pincode by using the addresses of corresponding post offices.
- **Step 3:** Determine the approximate geographic location of companies using the pincode extracted in Step 1 and the coordinates derived in Step 2.

Step 1: Extract the pincode from the company's registered address

Section 12 of the Companies Act of 2013 mandates that all businesses, including Limited Liability Partnerships (LLPs), must have a registered office either at the time of incorporation or within 30 days of their registration. This registered office serves as the official address of the company for all communications and legal purposes. The promoters of the company are responsible for designating a specific location as the registered office and must ensure that this address is formally registered with the Ministry of Corporate Affairs (MCA).

We use a regular expression (regex) pattern to extract the pincode from the company's registered address. In India, a valid pincode is always six digits long, does not start with zero, has a first digit ranging from 1 to 9, and the remaining five digits can be any number from 0 to 9. The regex pattern ensures that the extracted pincode adheres to this standard format. Table B.1 presents a few illustrative examples of the address and the extracted pincodes, highlighted in green.

[23] A pincode is a six-digit code in the Indian postal code system used by India Post to simplify the manual sorting and delivery of mail by eliminating confusion over incorrect addresses. It is similar to the concept of zipcode in the United States.

Table B.1: Illustrative Examples of Registered Addresses with Extracted Pincodes

Company Identification Number	Registered Address
U01110KA2016PTC097248	Salughatte Farmers Producer Company Limited, Sira TQ 138, Hunasekatte, Sira, Karnataka 572115
U15323CT2014PTC001592	Jai Hanuman Agroprocessing Pvt. Limited, M.G. Road, beside Makhan Vihar, H.No. 05, Ambikapur, Chhattisgarh 497001
U25114OR2006PTC008532	Utkal Retread Pvt. Ltd., Keonjhar, Plot No. 2577/3298, AT/PO: Raisuar, Odisha 758013

Step 2: Approximate geographic coordinates of a pincode

The country is divided into multiple pin zones, each corresponding to a specific geographic region. A single pincode can be served by more than one post office. To approximate the location of firms within a pincode, we use the location of one of the post offices in that pincode. The following steps outline how we estimate the geographic coordinates for each pincode.

1. Collection of Post Office Data
    - We obtain a dataset of post offices, including their names, addresses, and the pincodes they serve from India Post
    - This dataset provides attributes for 154,797 post offices across the country serving 19,100 unique pincodes
2. Approximating Pincodes to Post Offices
    - We approximate the location of each pincode by first sorting the associated post offices alphabetically by name[24]
    - We then select the first post office in the list to represent the pincode's geographic location, ensuring a consistent and singular geographic point for each pincode
3. Geocoding Post Office Addresses
    - We submit the address of the post office chosen in the previous step to Google Maps API for geocoding
    - The API returns the latitude and longitude of the address, representing the geographic coordinates of the post office

[24]Our estimates are robust to this choice, and we further test the robustness of this selection method in subsection B.3

Illustrative Example

To illustrate the process of approximating pincode coordinates, consider pincode ***225303***, which is served by three post offices: *Basauli B.O*, *Mohammad Pur Khaley S.O*, and *Ram Mandai B.O*, all located within the Taluk of Fatehpur in the Barabanki district of Uttar Pradesh (see Table B.2).

Table B.2: Post Offices serving the Pincode *225303*

Sno	Post Office Name	Pincode	Taluk	District	State
1	Basauli B.O	225303	Fatehpur	Barabanki	Uttar Pradesh
2	Mohammad Pur Khaley S.O	225303	Fatehpur	Barabanki	Uttar Pradesh
3	Ram Mandai B.O	225303	Fatehpur	Barabanki	Uttar Pradesh

We approximate the geographic coordinates of this pincode using the following steps:

1. **Alphabetical Sorting of Post Offices:** The post offices are sorted alphabetically by their names, resulting in the sequence: *Basauli B.O*, *Mohammad Pur Khaley S.O*, and *Ram Mandai B.O*.

2. **Selection of First Post Office:** From the alphabetically sorted list, *Basauli B.O* is the first post office and is selected to represent the geographic location of pincode 225303.

3. **Geocoding the Selected Post Office:** The address of the selected post office, "*Basauli B.O, Fatehpur Taluk, Barabanki District, Uttar Pradesh*", is then geocoded using the Google Maps API. The API returns the latitude and longitude for the location of *Basauli B.O*, which are then used as the approximate geographic coordinates for pincode 225303.

Step 3: Approximate geographic coordinates of companies

Using the pincode extracted from a company's registered address in **Step 1**, we approximate the company's geographic location based on the coordinates of its associated pincode, as determined in **Step 2**. We assume that a company's registered office, as recorded in the Ministry of Corporate Affairs (MCA) dataset, is generally located near the post office serving its pincode.

Data Attrition

The original Ministry of Corporate Affairs (MCA) dataset comprises of 908,267 firms. We are able to retain 900,396 firms after employing the regex pattern matching to extract pincodes form their registered address. The pincode extraction process results in 7,871 firms (0.86%) being dropped due to invalid or missing pincodes in their addresses. Finally, we were able to geolocate 825,980 firms to villages, based on the 2001 census definitions, encompassing 91.7% of the original dataset.

B.2 Accuracy of Custom Geocoding Method

We assess the accuracy of our geocoding process by comparing the coordinates generated by our methodology with those obtained from an external service, the Latlong Maps API. We perform this validation on a sample of 195,000 firms registered across four states – Gujarat, Uttar Pradesh, Karnataka, and Tamil Nadu. We measure the reliability of our method by calculating the distance between the firm geo-coordinates based on our methodology and firm geo-coordinates provided by Latlong Maps API. We express the distance error between the two sets of coordinates in kilometers (km). Table B.3 summarizes the statistical distribution of the distance error.

Table B.3: Statistical Distribution of Distance Errors (in Km)

# Firms	p25	p50	p75	p90	Mean
195,000	0.75	1.58	3.83	11.27	6.42

This table presents the statistical distribution of distance errors in kilometers (km) between the coordinates derived from our geocoding methodology and those obtained from the Latlong Maps API. The columns display the number of firms, along with the 25th percentile (p25), median (p50), 75th percentile (p75), 90th percentile (p90), and mean distance errors, with the mean calculated after winsorizing the right tail at 1%

Table B.3 shows that the median distance error between our geocoding results and those from the Latlong Maps API is 1.58 km. Given that the median block area is 313 square km, corresponding to a median block diameter of roughly 20 km, this error is relatively minor. This suggests that our geocoding method accurately places firms within the correct block.

To evaluate the robustness of our baseline results to our geocoding approach, we re-estimate our baseline specification using data derived from both our custom geocoding method and the Latlong Maps API. The results, shown in Table B.4, indicate that our geocoding method does not materially affect the study's conclusions. In fact, comparing the estimates in Panel A (based on our methodology) and Panel B (based on geo-coordinates from Latlong API) suggests that the measurement error in firm location due to our approximation methodology likely leads to a downward bias in the baseline estimates for this sample.

Table B.4: RD Estimates: Using Different Geocoding Approaches

Panel A: Data constructed using Custom Geocoding Method				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0663***	0.0605***	0.0587***	0.0558***
	(0.0135)	(0.0116)	(0.0112)	(0.0111)
R^2	0.11139	0.09334	0.08321	0.07799
Observations	80,790	105,325	119,053	122,257
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

Panel B: Data constructed using Latlong Maps API				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0873***	0.0955***	0.0866***	0.0737***
	(0.0199)	(0.0176)	(0.0179)	(0.0189)
R^2	0.20503	0.18110	0.16414	0.15450
Observations	80,790	105,325	119,053	122,257
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents RD estimates comparing the effectiveness of two geocoding methods: our custom method and the Latlong Maps API. Panel A uses data constructed via our custom geocoding method, while Panel B uses data from the Latlong Maps API. The dependent variable is the natural logarithm of 0.001 plus the number of new firms. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary separating a split block from an unsplit block in columns 1, 2, 3, and 4, respectively. Each regression includes boundary fixed effects, a local linear specification estimated separately on each side of the boundary, and a triangular kernel. Standard errors are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Robustness Checks on Post Office Selection Within Pincodes

While our initial validation confirms the overall accuracy of our geocoding method, we recognize that the choice of post office within each pincode could introduce bias into our results. This issue is especially relevant when a pincode spans both split and unsplit blocks. Each pincode can be served by multiple post offices. In our geocoding process, we select the first post office listed in alphabetical order to represent the pincode. This selection method may bias our estimates if these chosen post offices are disproportionately located within split blocks. To address this concern and further validate our methodology, we conduct additional analyses on the impact of post office selection within pincodes. The following section outlines our approach and presents the results of our robustness checks.

Distribution of Split and Unsplit Blocks within each pincode

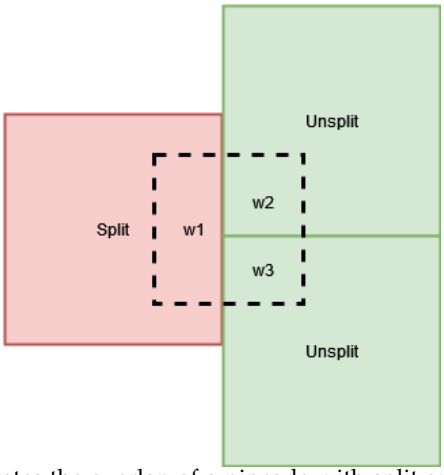
We start by analyzing the spatial distribution of split and unsplit blocks within each pincode. By utilizing shapefiles for pincodes and administrative blocks, we observe their overlap and measure the fraction of the area occupied by split blocks for each pincode using the following expression:

$$\text{Fraction(Split)} = \sum_{i=1}^{N} w_i \times 1(\text{Split})$$

here, N denotes the total number of blocks within the pincode, i is the index for each block, w_i represents the fractional area of block i within the pincode, and 1(Split) is an indicator function that equals 1 if block i is a split block and 0 otherwise. Note that if the value of Fraction(Split) is close to 0, it indicates that the pincode is entirely within unsplit blocks. Conversely, a value close to 1 suggests that the pincode is fully contained within split blocks.

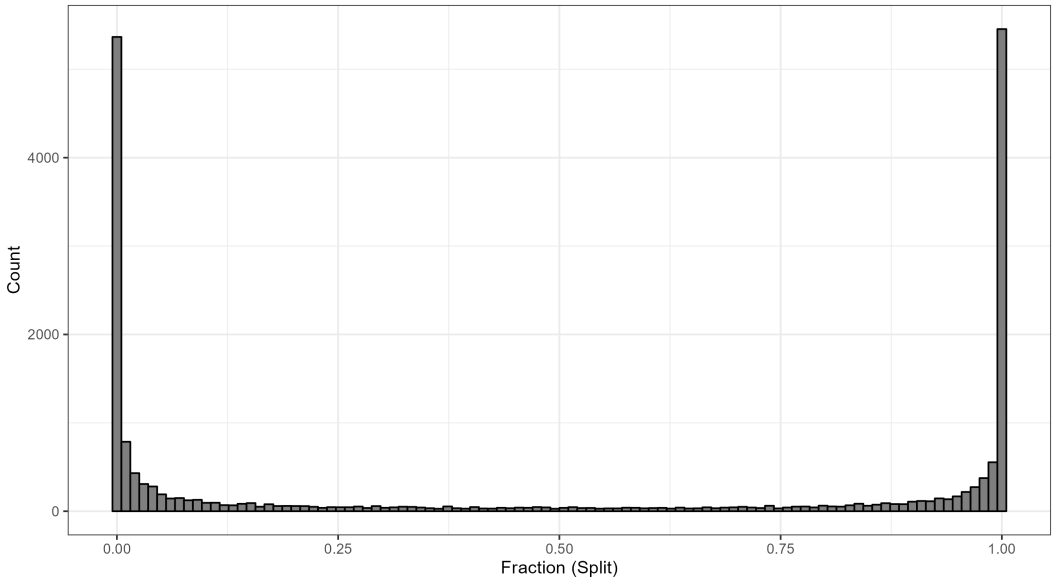
Figure B.1 provides a visual representation of this calculation. The figure shows how a pincode overlaps with both split and unsplit blocks. The shaded squares represent different block types: the red square indicates a split block, while the green squares represent unsplit blocks. The dashed outline marks the pincode boundary. The fractional contributions of each block to the pincode area are labeled as as w_1, w_2, and w_3.

Figure B.1: Visualization of Pincode Overlap Across Split and Unsplit Blocks



This figure illustrates the overlap of a pincode with split and unsplit blocks. The shaded squares denote the different block types, with the red area representing a split block and the green areas indicating unsplit blocks. The dashed outline defines the pincode boundary. The fractional contributions of each block to the area of the pincode are labeled as w_1, w_2, and w_3.

Figure B.2: Histogram of proportion of area occupied by split blocks within each pincode



This figure displays a histogram depicting the distribution of the fraction of split blocks (Fraction(Split)) across all pincodes. Each bar represents the count of pincodes falling within specified ranges of Fraction(Split), illustrating how the area of split blocks is distributed among the pincodes analyzed.

Figure B.2 presents the distribution of the fraction of split blocks across all pincodes. The distribution is predominantly clustered around 0 and 1, indicating that most pincodes are either entirely within split blocks or entirely within unsplit blocks, as reflected by Fraction(Split) values close to 0 or 1. Specifically, 15,559 out of 19,448 pincodes ($\approx 80\%$) have a Fraction(Split) of less than 0.1 or greater than 0.9, indicating that most pincodes are subsumed within split or unsplit blocks.

We estimate our baseline specification using only firms located in pincodes that are clearly either entirely within split blocks or entirely within unsplit blocks. This approach ensures that our geocoding method, which approximates firm locations based on the first post office, does not unfairly favor post offices in split blocks. By focusing on these distinctly categorized pincodes, we can confirm that our baseline results are robust and not unlikely to be influenced by selection bias. Table B.5 presents the results of this analysis across three panels, each with progressively stricter criteria for including pincodes.

Table B.5: Estimates for Firms in pincodes that are subsumed within Split and Unsplit Blocks

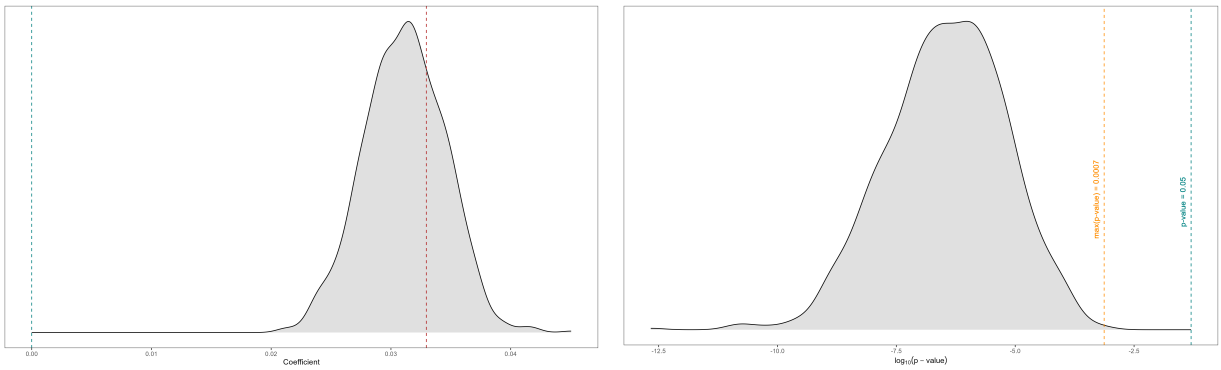
Panel A: Pincodes with Fraction(Split) < 0.1 or Fraction(Split) > 0.1				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0441***	0.0438***	0.0415***	0.0416***
	(0.0106)	(0.0090)	(0.0083)	(0.0080)
R^2	0.20149	0.16425	0.14214	0.12960
Observations	165,922	244,004	300,788	320,556
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Panel B: Pincodes with Fraction(Split) < 0.05 or Fraction(Split) > 0.95				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0499***	0.0489***	0.0466***	0.0465***
	(0.0123)	(0.0104)	(0.0095)	(0.0092)
R^2	0.22580	0.18416	0.15802	0.14298
Observations	136,362	206,603	260,789	280,213
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Panel C: Pincodes with Fraction(Split) < 0.01 or Fraction(Split) > 0.99				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
g10_split	0.1054***	0.1021***	0.0966***	0.0929***
	(0.0240)	(0.0197)	(0.0178)	(0.0169)
R^2	0.28650	0.23506	0.19889	0.17754
Observations	76,044	122,538	165,610	183,343
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for the natural logarithm of 0.001 plus the number of new firms as the dependent variable. Each panel progressively narrows the sample to include villages in pincodes that are more homogeneously within either split or unsplit blocks. Panel A includes villages in pincodes with Fraction(Split) < 0.1 or Fraction(Split) > 0.9, Panel B further restricts to Fraction(Split) < 0.05 or Fraction(Split) > 0.95, and Panel C to Fraction(Split) < 0.01 or Fraction(Split) > 0.99. The models use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary separating split and unsplit blocks in columns 1, 2, 3, and 4, respectively. The unit of observation is a village that lies within the specified bandwidths. A split block is defined based on the overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm entry is calculated as the total number of new firms from 2008 to 2016. All regressions include boundary fixed effects. Standard errors, reported in parentheses, are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To further assess the robustness of our results against potential biases from arbitrary post office selection, we use a bootstrapping approach. We generate 1,000 samples, retain-

ing the original post office choice for the 15,559 pincodes that are entirely within either split or unsplit blocks. For the 3,889 pincodes that overlap, we geocode all post offices and randomly select one post office per pincode. This introduces variability in firm location approximations. We then infer firm coordinates from these selected post offices and map the firms to villages based on these coordinates.

Figure B.3: Bootstrap Estimates



(a) Distribution of coefficient estimates

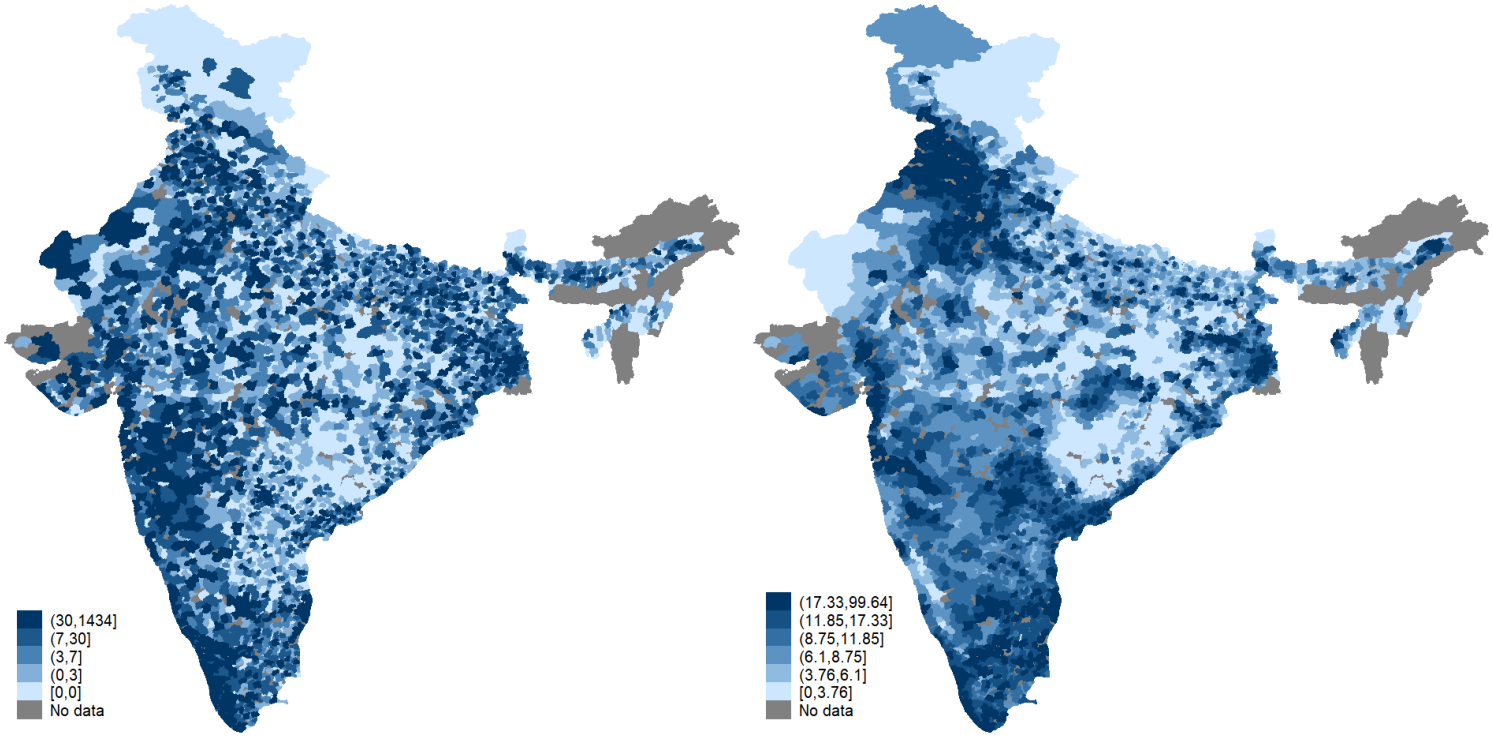
(b) Distribution of log-transformed p-values

The figure presents the results of a bootstrapping analysis to test the robustness of the effect of split blocks on firm entry to potential bias from arbitrary post office selection. Panel A shows the density distribution of coefficient estimates from 1,000 bootstrap samples, with the red dashed line indicating the baseline estimate. Panel B displays the distribution of log-transformed p-values from the bootstrap samples. Vertical lines are included to denote the maximum p-value (0.0007) and the conventional significance threshold (0.05)

Figure B.3 presents the results of the bootstrap analysis. Panel A presents the distribution of coefficient estimates, demonstrating a consistently positive effect of split blocks on firm entry across all samples. The coefficients range from approximately 0.0209 to 0.0450, and none of the 95% confidence intervals cross zero. Panel B illustrates the distribution of log-transformed p-values, with the maximum p-value around 0.0007, which is significantly below the conventional significance threshold of 0.05 (denoted by the dashed line). These findings indicate that the baseline estimates are stable and reliable, suggesting that they are not artifacts of our custom geocoding method.

Appendix C Data

Figure C.1: Geography of Firm Entry and Nightlights

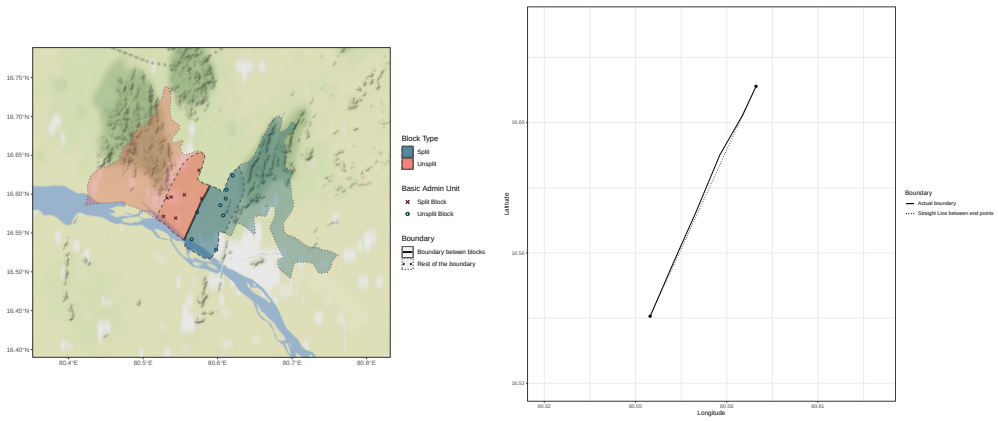


(a) # New Private Firms

(b) Average Nightlights

The figure panel C.1a plots the geographic distribution of the number of new private firms registered across all blocks in India from 2003 to 2016. The figure panel C.1b plots the geographic distribution of average night lights across all blocks in India from 2003 to 2016. The sample of private firms includes the universe of all for-profit firms in India registered from 2003 to 2016. The nightlights are standardized to take a minimum value of zero and a maximum value of 100 across blocks for each year. The standardized nightlights in a block are averaged over the period from 2003 to 2016. Note that we have not verified any boundaries and do not claim authenticity of the same. We do not endorse the geographic boundaries shown here.

Figure C.2: Sinuosity of Straight-Line-Like Block Boundary Segment

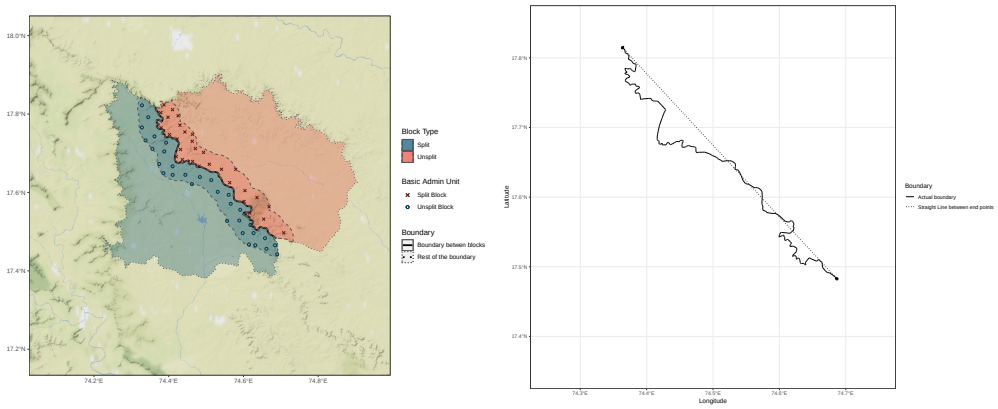


(b) Schematic Diagram

(a) Representative Example

Figure C.2 illustrates the measure of straightness, *Sinuosity*, using a representative example of straight-line-like boundary. Figure C.2a illustrates the boundary between the blocks, Ibrahimpatnam and Vijayawada (Rural), located in Krishna District, Andhra Pradesh. Figure C.2b compares the block boundary and straight line joining the endpoints of the boundary. The length of the boundary segment is 8.69 km, whereas the shortest distance between the endpoints is 8.68 km. The *Sinuosity* of the segment is 1.000965.

Figure C.3: Sinuosity of a Wiggly Block Boundary Segment



(b) Schematic Diagram

(a) Representative Example

Figure C.3 illustrates the measure of straightness, *Sinuosity*, using a representative example of wiggly boundary segment. Figure C.3a illustrates the boundary between the blocks, Khatav and Phaltan located in Satara District, Maharashtra. Figure C.3b compares the block boundary and straight line joining the endpoints of the boundary. The length of the boundary segment is 80.41 km, whereas the shortest distance between the end-points is 50.38 km. The *Sinuosity* of the segment is 1.59.

Table C.1: Number of Firms

Year	# New Firms	Nightlights
2003	10,980	5.98
2004	15,856	7.52
2005	23,051	6.72
2006	39,138	6.20
2007	58,735	7.73
2008	65,299	7.39
2009	55,855	8.10
2010	80,075	8.14
2011	88,833	8.10
2012	94,213	8.96
2013	83,100	9.01
2014	62,858	14.64
2015	72,375	15.47
2016	75,612	15.26
Total/Average	825,980	9.22

This table presents the number of new firms and the average nightlights index for each year from 2003 and until 2016.

Appendix D Number of Multiple Politicians & Firm Entry

Table D.1: Number of Politicians in Split Blocks and Firm Entry

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Politician (=2)	0.0261***	0.0256***	0.0236***	0.0229***
	(0.0072)	(0.0060)	(0.0056)	(0.0054)
Politician (=3)	0.0524***	0.0472***	0.0487***	0.0502***
	(0.0180)	(0.0163)	(0.0173)	(0.0185)
Politician (>=4)	0.3661***	0.3956***	0.4322***	0.4526***
	(0.1099)	(0.1163)	(0.1095)	(0.1088)
#Obs	263,307	356,260	416,694	436,980
R^2	0.1541	0.1297	0.1151	0.1066
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Politician (=3) - Politician (=2)	0.0263	0.0216	0.0251	0.0273
Politician (=4) - Politician (=2)	0.340	0.370	0.4086	0.4297
F-Stat	5.80	5.89	7.98	8.85
Prob > F	0.0030	0.0028	0.0003	0.0001

This table presents estimates for specification (1) augmented for the number of politicians in split blocks using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. We further segregate split blocks based on the number of politicians governing the block. Politician (=2) is a binary variable taking a value of 1 for split blocks with exactly two politicians. Politician (=3) is a binary variable taking a value of 1 for split blocks with exactly politicians. Politician (>=4) is a binary variable taking a value of 1 for split blocks with four or more politicians. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix E Robustness Tests

Table E.1: RD Estimate on Firm Entry with Polynomial of Degree 2

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0282***	0.0322***	0.0297***	0.0281***
	(0.0101)	(0.0080)	(0.0070)	(0.0064)
# Obs	263,307	356,260	416,694	436,980
R^2	0.1537	0.1293	0.1147	0.1062
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Polynomial Degree	2	2	2	2

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local quadratic specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.2: RD Estimate with Uniform Kernel

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split(=1)	0.0334***	0.0294***	0.0308***	0.0307***
	(0.0073)	(0.0062)	(0.0057)	(0.0056)
# Obs	257,787	348,539	407,616	427,659
R^2	0.6054	0.5736	0.5453	0.5239
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Kernel	Uniform	Uniform	Uniform	Uniform

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a uniform kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.3: RD Estimates with Spatial Polynomial

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split(=1)	0.2456***	0.2867***	0.3184***	0.3314***
	(0.0821)	(0.0742)	(0.0707)	(0.0706)
# Obs	263,307	356,260	416,694	436,980
R^2	0.1537	0.1293	0.1146	0.1061
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Spatial Polynomial	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification with latitude and longitude estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.4: RD Estimates with Conley (1999) Standard Errors

Bandwidth ↓	(1)	(2)	(3)	(4)
5 KM	0.0331***	0.0331***	0.0331***	0.0331***
	(0.0077)	(0.0077)	(0.0078)	(0.0079)
10 KM	0.0326***	0.0326***	0.0326***	0.0326***
	(0.0063)	(0.0063)	(0.0064)	(0.0067)
20 KM	0.0312***	0.0312***	0.0312***	0.0312***
	(0.0056)	(0.0056)	(0.0057)	(0.0060)
50 KM	0.0309***	0.0309***	0.0309***	0.0309***
	(0.0053)	(0.0053)	(0.0054)	(0.0058)
Conley Cutoff →	2 KM	5KM	10 KM	20 KM

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are Conley (1999) standard errors, estimated for cutoff values of 2, 5, 10, and 20 km in columns (1), (2), (3), and (4), respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.5: Standard Errors with Alternative Clustering

Bandwidth ↓	(1)	(2)	(3)	(4)
5 KM	0.0331***	0.0331***	0.0331***	0.0331***
	(0.0078)	(0.0081)	(0.0082)	(0.0111)
10 KM	0.0326***	0.0326***	0.0326***	0.0326***
	(0.0064)	(0.0072)	(0.0068)	(0.0089)
20 KM	0.0312***	0.0312***	0.0312***	0.0312***
	(0.0056)	(0.0068)	(0.0062)	(0.0086)
50KM	0.0309***	0.0309***	0.0309***	0.0309***
	(0.0053)	(0.0066)	(0.0060)	(0.0089)
Clustering →	Village	District	Boundary	State

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are based on alternative clusters, with clustering at the village, district, boundary, and state level in columns (1), (2), (3), and (4), respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.6: RD Estimate with Alternative Sample

Bandwidth ↓	(1)	(2)	(3)	(4)	(5)	(6)
5 KM	0.0331***	0.0462***	0.1113***	0.1606***	0.1824***	0.0270***
	(0.0071)	(0.0108)	(0.0220)	(0.0322)	(0.0464)	(0.0068)
10 KM	0.0326***	0.0443***	0.1107***	0.1575***	0.1825***	0.0245***
	(0.0060)	(0.0092)	(0.0191)	(0.0283)	(0.0411)	(0.0057)
20 KM	0.0312***	0.0421***	0.1031***	0.1464***	0.1727***	0.0229***
	(0.0056)	(0.0086)	(0.0181)	(0.0265)	(0.0389)	(0.0052)
50 KM	0.0309***	0.0411***	0.1009***	0.1428***	0.1726***	0.0226***
	(0.0055)	(0.0084)	(0.0177)	(0.0258)	(0.0381)	(0.0050)
Sample	Full Sample	Firms > 0	Firms > 5	Firms > 10	Firms > 20	Firms <= 1000

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable in panel A, and the inverse hyperbolic sine (IHS) transformation of the number of firms in panel B. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Column (1) includes the baseline sample. Column (2) drops all blocks with no firm entry. Columns (3), (4), and (5) drops blocks where less than five, ten, and twenty firms, respectively, entered during the sample period. Column (6) drops all large blocks with number of new firms greater than 1000. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.7: RD Estimate with Alternative Transformation

Panel A : Poisson Pseudo Maximum Likelihood						
Bandwidth ↓	(1)	(2)	(3)	(4)	(5)	(6)
5 KM	0.2117***	0.1667***	0.2545***	0.2653***	0.2487***	0.1731***
	(0.0515)	(0.0525)	(0.0649)	(0.0728)	(0.0835)	(0.0523)
10 KM	0.1856***	0.1332***	0.1912***	0.2015***	0.1606**	0.1411***
	(0.0430)	(0.0443)	(0.0551)	(0.0624)	(0.0713)	(0.0435)
20 KM	0.2003***	0.1444***	0.2046***	0.1938***	0.1589**	0.1555***
	(0.0385)	(0.0401)	(0.0505)	(0.0568)	(0.0659)	(0.0389)
50 KM	0.2011***	0.1424***	0.2016***	0.1978***	0.1741***	0.1535***
	(0.0369)	(0.0387)	(0.0492)	(0.0560)	(0.0654)	(0.0370)
Sample	Full Sample	Firms > 0	Firms > 5	Firms > 10	Firms > 20	Firms <= 1000
Panel B : Outcome - IHS(#New Firms)						
Bandwidth ↓	(1)	(2)	(3)	(4)	(5)	(6)
5 KM	0.0126***	0.0196***	0.0440***	0.0628***	0.0757***	0.0102***
	(0.0022)	(0.0035)	(0.0078)	(0.0119)	(0.0178)	(0.0021)
10 KM	0.0128***	0.0196***	0.0432***	0.0616***	0.0753***	0.0096***
	(0.0020)	(0.0031)	(0.0069)	(0.0107)	(0.0160)	(0.0017)
20 KM	0.0121***	0.0186***	0.0402***	0.0571***	0.0704***	0.0088***
	(0.0019)	(0.0030)	(0.0067)	(0.0103)	(0.0155)	(0.0016)
50 KM	0.0120***	0.0183***	0.0396***	0.0562***	0.0709***	0.0087***
	(0.0019)	(0.0030)	(0.0067)	(0.0102)	(0.0155)	(0.0016)
Sample	Full Sample	Firms > 0	Firms > 5	Firms > 10	Firms > 20	Firms <= 1000

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows 1, 2, 3, and 4, respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Column (1) includes the baseline sample. Column (2) drops all blocks with no firm entry. Columns (3), (4) and (5) drops blocks where less than 5, 10, and 20 firms, respectively, entered during the sample period. Column (6) drops all large blocks with more than 1000 firms. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.8: RD Estimates with Non-Pooled Data

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split(=1)	0.0215***	0.0214***	0.0199***	0.0198***
	(0.0036)	(0.0032)	(0.0031)	(0.0031)
# Obs	1,813,190	2,497,757	2,958,673	3,120,412
R^2	0.1219	0.0997	0.0873	0.0805
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Spatial Polynomial	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year that lies within the narrow bandwidth of the boundary separating a split and an unsplit block from 2008 to 2016. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village during the year. All regressions include boundary fixed effect, a local linear specification with latitude and longitude estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.9: RD Estimates with Alternative Grids

Bandwidth ↓	Grid Size		
	(1) 1×1	(2) 2×2	(3) 5×5
5 KM	0.0150***	0.0480***	0.2244***
	(0.0028)	(0.0094)	(0.0480)
10 KM	0.0163***	0.0556***	0.2045***
	(0.0025)	(0.0084)	(0.0348)
20 KM	0.0184***	0.0636***	0.2266***
	(0.0025)	(0.0082)	(0.0292)
50 KM	0.0226***	0.0783***	0.2693***
	(0.0022)	(0.0072)	(0.0248)

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows 1, 2, 3, and 4, respectively. The unit of observation is an arbitary grid of size $x \times x$ that lies within the narrow bandwidth of the boundary separating a split and an unsplit block from 2008 to 2016. Columns 1, 2 and 3 use a grid size of 1×1, 2×2 and 5×5, respectively. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered in the cell during the period. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.10: RD Estimates before the 2008 Delimitation

Dep Var ↓	(1)	(2)	(3)	(4)
LN(0.001+# New Firms)	0.0142	0.0114	0.0064	0.0046
	(0.01310)	(0.01170)	(0.01120)	(0.01090)
LN(0.001+Nightlights)	0.0059	0.0442	-0.0408	-0.0763
	(0.13170)	(0.15160)	(0.16380)	(0.16900)
Bandwidth →	5 KM	10 KM	20 KM	50 KM

This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry and nightlights as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year that lies within the narrow bandwidth of the boundary separating a split and an unsplit block from 2003 until 2007. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village during the year. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.11: RD Estimate with Additional Controls – Dep Var: LN(0.001+# New Firms)

Bandwidth ↓	(1)	(2)	(3)	(4)	(5)	(6)
5 KM	0.0331***	0.0291***	0.0292***	0.0474***	0.0317***	0.0416***
	(0.0071)	(0.0069)	(0.0069)	(0.0097)	(0.0070)	(0.0093)
10 KM	0.0326***	0.0277***	0.0283***	0.0464***	0.0308***	0.0382***
	(0.0060)	(0.0058)	(0.0058)	(0.0084)	(0.0060)	(0.0080)
20 KM	0.0312***	0.0264***	0.0273***	0.0424***	0.0293***	0.0337***
	(0.0056)	(0.0054)	(0.0054)	(0.0083)	(0.0056)	(0.0078)
50 KM	0.0309***	0.0263***	0.0269***	0.0401***	0.0289***	0.0325***
	(0.0055)	(0.0052)	(0.0052)	(0.0082)	(0.0054)	(0.0078)
Controls	None	Population	Area	Distance to HQ	Compactness	All

This table presents estimates for specification (1), augmented for controls, using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in rows 1, 2, 3, and 4, respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Column (1) includes no controls. Columns (2), (3), and (4) include a linear and quadratic term for the 2001 population, geographic area, and the distance to district headquarters for a village. Column (5) controls for compactness of the block measured as in Harari (2020). Column (6) includes all controls included in columns (2)-(5). Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.12: RD Estimates Controlling for Politician Quality

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Split (=1)	0.0239***	0.0164***	0.0130**	0.0121**
	(0.0073)	(0.0062)	(0.0057)	(0.0056)
#Obs	263,156	356,101	416,528	436,814
R^2	0.1965	0.1670	0.1497	0.1394
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes

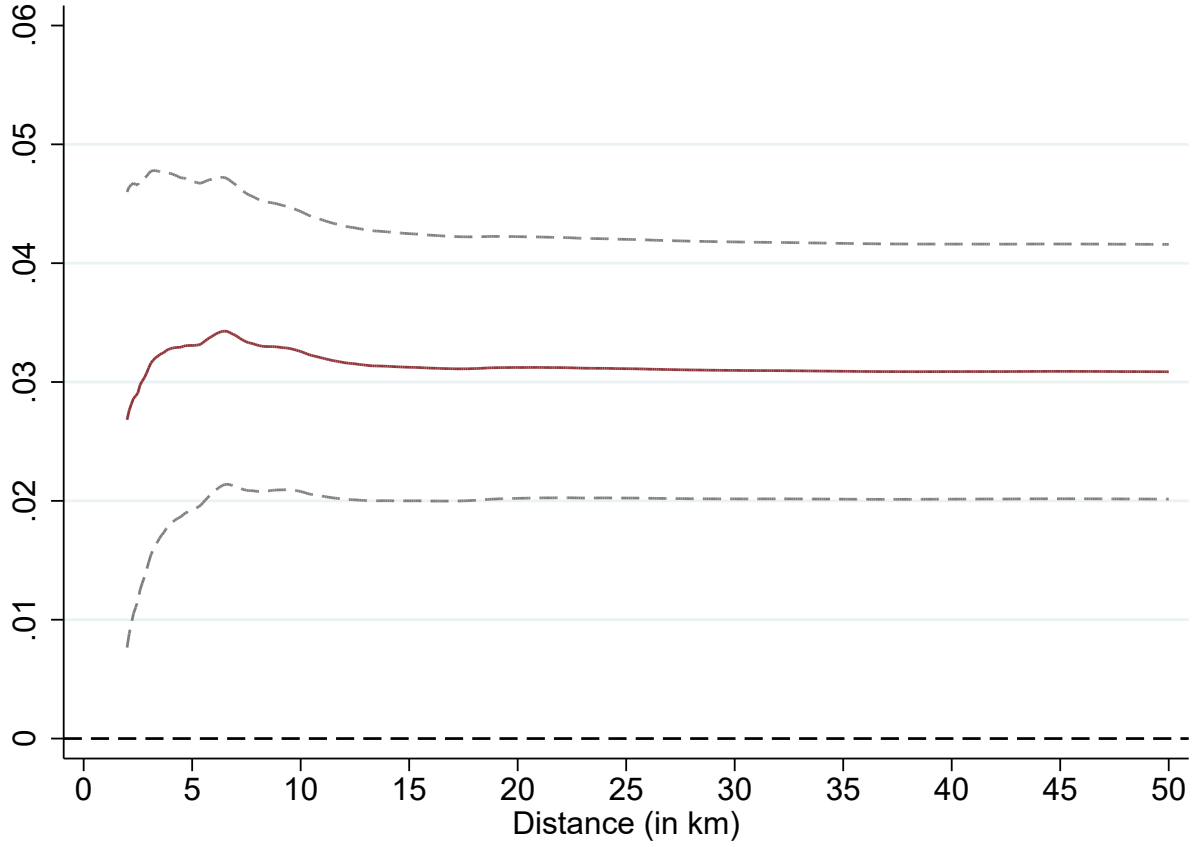
This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect and constituency fixed effect.a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E.13: Probability of Firm Exit

Dep Var: Firm exits within x time	(1)	(2)	(3)	(4)
6 months	-0.0008	-0.0004	-0.0003	-0.0002
	(0.0007)	(0.0004)	(0.0003)	(0.0002)
1 year	-0.0017	-0.0003	-0.0003	-0.0005
	(0.0014)	(0.0008)	(0.0007)	(0.0006)
2 year	-0.0044	-0.0008	-0.0010	-0.0022
	(0.0028)	(0.0019)	(0.0017)	(0.0016)
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary X Cohort Year FE	Yes	Yes	Yes	Yes

This table presents results from a firm-level regression, estimating the relationship between the probability of firm exit and split variable. It uses an RD specification similar to equation (1), but the data is at the firm level. For each firm that started between 2008 and 2016, we estimate the probability of it closing within six months, one year, and two years of its start as a function of the split status of the block. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a firm that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. All regressions include boundary × cohort fixed effects. All firms that entered during the same year are considered part of the same cohort. Each row of the table comes from a different regression. Rows 1, 2, and 3 estimate the probability of firm exit within six months, one year, and two years of firm entry, respectively. a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure E.1: RD Estimates for various Bandwidth choices

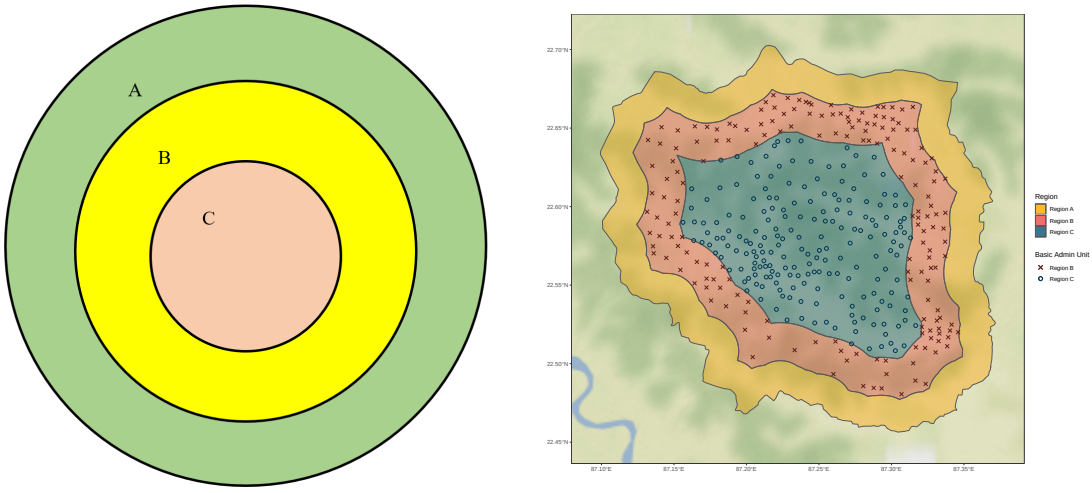


This table presents RD estimates and 95% confidence interval for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We estimate specification (1) for several bandwidths between 2 km and 50 km in increments of 0.1 km. The Y-axis reports the RD estimate, and the X-axis reports the bandwidth used to estimate the RD coefficient. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. The solid red line reports the RD estimate, and the dashed grey lines indicate the 95% confidence bands calculated based on standard errors clustered at the block level.

E.1 Falsification and Placebo Tests

We first describe the methodology to construct the falsification sample. We divide each block into three regions A, B and C, as shown in Figure E.2. These regions are equal in area and account for one-third of the entire block area. Our falsification exercise compares villages in region B with villages in region C within a narrow bandwidth of the boundary separating the two regions. The intuition of this test is that crossing from region B to region C does not change the number of politicians governing the area. Hence, if our baseline results capture the effect of the change in the number of politicians and not spatial correlation, we should observe a null effect when moving from region B to region C.

Figure E.2: Falsification Test: Sample Construction



(a) Schematic Diagram

(b) Representative Example

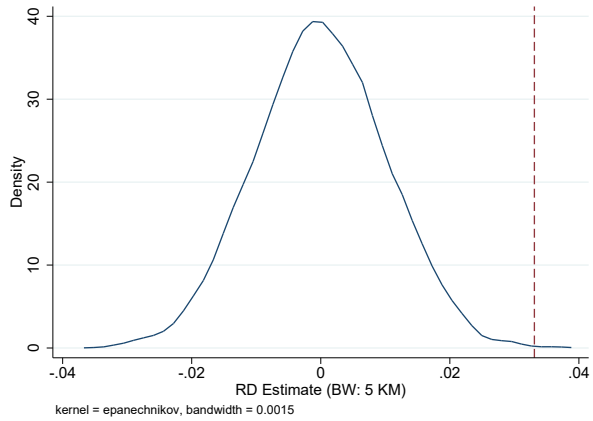
Figure E.2a presents the schematic diagram of a block that is split into three regions A, B and C. These regions are equal in area and account for one-third of the entire block area. Our falsification exercise compares villages in region B with villages in region C within a narrow bandwidth of the boundary separating the two regions. Figure E.2b presents the schematic diagram using a representative example using the block of Salbani in the West Medinipur district in the state of West Bengal.

Table E.14: Falsification Test: Results

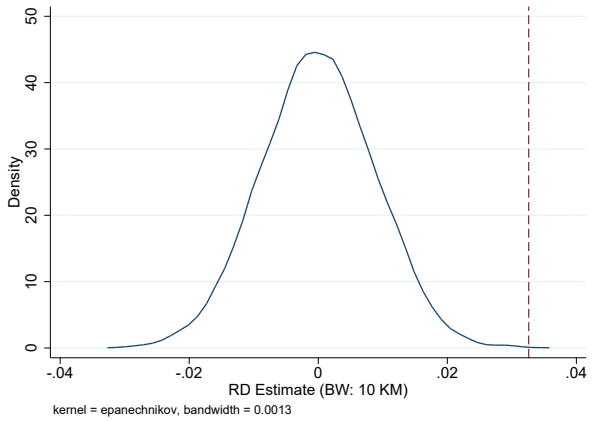
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Inside Region C (=1)	-0.0007	-0.0037	-0.0058	-0.0047
	(0.0088)	(0.0067)	(0.0058)	(0.0055)
# Obs	200,843	278,097	320,484	332,161
R^2	0.1587	0.1306	0.1190	0.1157
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) for the falsification sample using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating region B and region C within a block. A block is divided into three regions, A, B, and C, which are equal in area as shown in Figure E.2. The falsification test compares villages in region B and region C. Firm Entry is defined as the total number of firms that have entered a village from 2008 to 2016. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

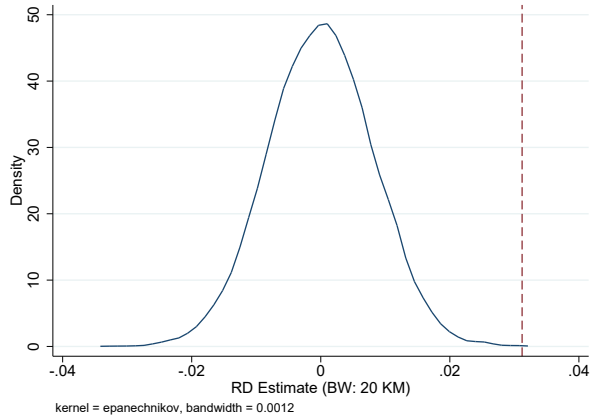
Figure E.3: Placebo Test



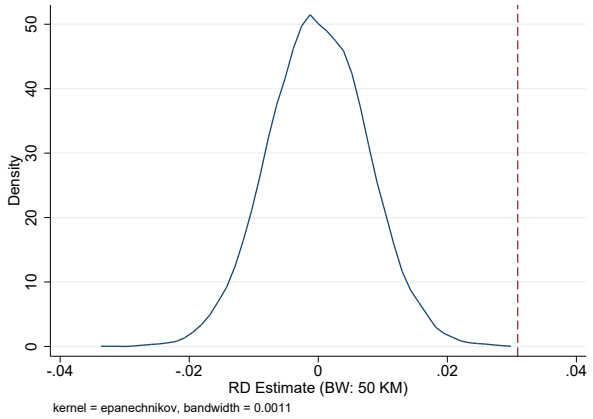
(a) Bandwidth: 5 KM



(b) Bandwidth: 10 KM



(c) Bandwidth: 20 KM



(d) Bandwidth: 50 KM

The figure plots the kernel density of the RD estimates for specification (1) for the randomly assigned split status obtained from the 10,000 Monte-Carlo simulations. The dependent variable is the natural logarithm of 0.001 plus firm entry. We randomly assign each block to be a split block or an unsplit block. Hence, a placebo split binary variable is generated for each block. Every block has a probability of 0.411 to be assigned as a split block. This is based on the fraction of total split blocks in the sample. We repeat this process 10,000 times and estimate the RD specification (1) with the placebo split variable for different bandwidths of 5 (Figure E.3a), 10 (Figure E.3b), 20 (Figure E.3c), and 50 km (Figure E.3d) on either side of the boundary. The dashed red line shows the point estimate from the corresponding bandwidth as in Panel A of Table 2.

Table E.15: 2008 Delimitation and Elections After 2008: Firm Entry and Change in Number of Politicians

Panel A : Unsplit to Split				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Treat x Post Delimitation	0.0103*	0.0093*	0.0083*	0.0079*
	(0.0061)	(0.0051)	(0.0047)	(0.0044)
Treat x Post Election	0.0140**	0.0098*	0.0083*	0.0075
	(0.0062)	(0.0054)	(0.0050)	(0.0048)
# Obs	687,792	975,312	1,185,816	1,263,456
R^2	0.6489	0.6329	0.6204	0.6129
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes
Panel B : Split to Unsplit				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Treat x Post Delimitation	-0.0092***	-0.0103***	-0.0108***	-0.0111***
	(0.0017)	(0.0015)	(0.0015)	(0.0015)
Treat x Post Election	-0.0029	-0.0023	-0.0026*	-0.0024*
	(0.0018)	(0.0016)	(0.0015)	(0.0014)
# Obs	2,469,408	3,297,648	3,812,364	3,978,144
R^2	0.6888	0.6861	0.6817	0.6769
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Village FE	Yes	Yes	Yes	Yes

This table presents estimates for supplemented specification (2) and (3) using the natural logarithm of 0.001 plus firm entry as the dependent variable in Panels A and B, respectively. The table supplements these specifications to include the interaction term of Treat and Post Election. In Panel A the treated group is the set of blocks that switched from being unsplit to split following the 2008 delimitation, whereas the control group comprises of blocks that were always unsplit both before and after the delimitation and border the treated group. In Panel B, the treated group is the set of blocks that switched from being split to unsplit following the 2008 delimitation, whereas the control group comprises blocks that are always split among multiple politicians both before and after the 2008 delimitation and border the treated group. The variable *PostxDelimitation* takes a value of 1 for all years since the announcement of delimitation in 2008. The *PostxElection* variable assumes a value of 1 for all years succeeding the initial state legislative election held after the announcement of delimitation. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating the treatment and the control groups in columns (1), (2), (3), and (4), respectively. The unit of observation is a village-year that lies within the narrow bandwidth of the boundary separating the treatment and the control group. Firm entry is defined as the total number of firms that have entered a village during the year. All regressions include village and boundary × year fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix F Industry Classification

Table F.1: List of Industries with High Regulatory Costs in the Sample

NIC Code	Industry Name	Broad Industry
1	Crop and animal production, hunting and related service activities	Agriculture, forestry and fishing
2	Forestry and Logging	Agriculture, forestry and fishing
5	Mining of coal and lignite	Mining and quarrying
19	Manufacture of coke and refined petroleum products	Manufacturing
35	Electricity, gas, steam and air conditioning supply	Electricity, gas, steam and air conditioning supply
36	Water collection, treatment and supply	Water supply; sewerage, waste management and remediation activities
41	Construction of buildings	Construction
45	Wholesale and retail trade and repair of motor vehicles and motorcycles	Wholesale and retail trade; repair of motor vehicles and motorcycles
50	Water transport	Transportation and storage
51	Air transport	Transportation and storage
52	Warehousing and support activities for transportation	Transportation and storage
61	Telecommunications	Information and communication
64	Financial service activities, except insurance and pension funding	Financial and insurance activities
65	Insurance, reinsurance and pension funding, except compulsory social security	Financial and insurance activities
66	Other financial activities	Financial and insurance activities
70	Activities of head offices; management consultancy activities	Professional, scientific and technical activities
72	Scientific research and development	Professional, scientific and technical activities
73	Advertising and market research	Professional, scientific and technical activities
74	Other professional, scientific and technical activities	Professional, scientific and technical activities

This table presents the list of industries with high regulatory costs in our sample. We define industries with a high regulatory cost following the methodology for the classification of regulated industries in the United States based on Pittman (1977) augmented with India specific regulations and factors as discussed in Awasthi, Yayavaram, George, and Sastry (2019).

Table F.2: List of Industries with High Rent Seeking or Crony-Capitalism

NIC C ode	Industry Name	Broad Industry
1	Crop and animal production, hunting and related service activities	Agriculture, forestry and fishing
2	Forestry and Logging	Agriculture, forestry and fishing
5	Mining of coal and lignite	Mining and quarrying
19	Manufacture of coke and refined petroleum products	Manufacturing
20	Manufacture of chemicals and chemical products	Manufacturing
36	Water collection, treatment and supply	Water supply; sewerage, waste management and remediation activities
41	Construction of buildings	Construction
52	Warehousing and support activities for transportation	Transportation and storage
64	Financial service activities, except insurance and pension funding	Financial and insurance activities
65	Insurance, reinsurance and pension funding, except compulsory social security	Financial and insurance activities
66	Other financial activities	Financial and insurance activities
93	Sports activities and amusement and recreation activities	Arts, entertainment and recreation

This table presents the list of industries with high rent-seeking or high degree of crony-capitalism in our sample. We use the index of industry-level cronyism created by *Economist* using the methodology developed by *Transparency International* to classify firms as high crony and low crony firms. The detailed description of the method to classify industries as crony can be found in The Economist: Planet Plutocrat.

Appendix G Mechanism

Table G.1: Multiple Politicians and Power Supply

Dep Var: High Power Supply	(1)	(2)	(3)	(4)
Split (=1)	0.0083*	0.0102**	0.0117**	0.0109**
	(0.0048)	(0.0046)	(0.0046)	(0.0047)
# Obs	105,264	143,320	168,952	177,819
R^2	0.7026	0.6944	0.6854	0.6791
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using high power supply as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. A village is defined as a high power supply village if it receives commercial electricity supply for more than 70% of time, both during winter and summer seasons, and is defined as low power supply village otherwise. The data on village level commercial power supply comes from the 2011 Indian Census. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table G.2: Multiple Politicians and Swachh Bharat Mission

Dep Var: LN(1 + Funding Received)	(1)	(2)	(3)	(4)
Split (=1)	0.0931*	0.1066**	0.1165**	0.1163**
	(0.0514)	(0.0485)	(0.0480)	(0.0486)
# Obs	220,665	298,546	349,219	366,325
R^2	0.2959	0.2912	0.2887	0.2884
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

This table presents estimates for specification (1) using the natural logarithm of funding disbursed under the Swachh Bharat Mission (SBM) as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. The unit of observation is a village that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. The data on village level funding disbursed under SBM is the sum of all funding received from 2015 until 2018. All regressions include boundary fixed effect, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table G.3: Confidence in Local Politician and State Machinery

Dep Var: Confidence Level	(1)	(2)	(3)	(4)
	Local	State Machinery		State
	Politician	Police	Panchayats	Government
Fraction of Split Blocks	0.2841**	0.1659**	0.2112***	-0.0109
	(0.1389)	(0.0786)	(0.0812)	(0.1424)
Sample Mean	1.4257	1.9617	2.1107	2.0751
Standard Deviation	0.5249	0.3397	0.3075	0.4709
# Obs	686	686	686	686
R^2	0.6277	0.7125	0.5592	0.5192
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

This table presents estimates from the regression of confidence in local politicians, state machinery, and the state government on the fraction of split blocks in the district. The unit of observation is district-year. The data on confidence comes from the two survey waves of the India Human Development Survey (IHDS) conducted in 2001 and 2011. Each respondent is asked, "How much confidence do you have in X?", where "X" denotes local politicians, police, panchayat, and the government at the state level. A respondent can say - (1) No Confidence coded as 1, (2) Intermediate level of confidence coded as 2, or (3) High level of confidence coded as 3. We collapse the respondent level data at the village level using survey weights to create a new variable that measures the district-level average confidence in "X." For each district, we create a variable - fraction of split blocks. It is calculated as the ratio of the number of split blocks to the total number of blocks in the district. We define a block as split in 2001 and 2011 based on 1977 and 2008 delimitation. Columns (1), (2), (3), and (4) use confidence in the local politician, police, panchayat, and the state government as the dependent variable. We also present the average and the standard deviation of the dependent variables. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table G.4: Difference in Confidence among Government and Private Schools and Hospitals

Dep Var: Confidence Level	(1) Schools	(2) Hospitals
Govt × Fraction of Split Blocks	0.2061**	0.2089**
	(0.0956)	(0.1022)
Govt	-0.2456**	-0.3635**
	(0.1178)	(0.1456)
Sample Mean	2.5856	2.5535
Standard Deviation	0.2646	0.2937
District FE	Yes	Yes
# Obs	688	688
R^2	0.6986	0.6950

This table presents estimates from the regression of confidence in schools and hospitals on the interaction term of schools and hospitals provided by the government and the fraction of split blocks in the district. The unit of observation is district-provider; that is, for each district, we have two observations, one indicating the confidence in the government good and another in the private sector good. The data on confidence comes from the 2011 survey wave of the India Human Development Survey (IHDS). Each respondent is asked - "How much confidence do you have in X?", where "X" denotes government and private schools and hospitals. A respondent can say - (1) No Confidence coded as 1, (2) Intermediate level of confidence coded as 2, or (3) High level of confidence coded as 3. We collapse the respondent level data at the village-provider level using survey weights to create a new variable that measures the district-provider-level average confidence in "X." Provider is either the private sector or the government. For each district, we create a variable - fraction of split blocks. It is calculated as the ratio of the number of split blocks to the total number of blocks in the district. We define a block as a split block based on the 2008 delimitation. Columns (1) and (2) use confidence in schools and hospitals as the dependent variable. We also present the average and the standard deviation of the dependent variables. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table G.5: Likelihood of Incumbent's Re-election

Dep Var: Incumbent Re-elected	(1)	(2)	(3)
Fraction of Split Blocks	0.0545***	0.0518***	0.0517***
	(0.0183)	(0.0184)	(0.0184)
State × Year FE	Yes	Yes	Yes
Reservation FE		Yes	Yes
By-Election FE			Yes
# Obs	4,667	4,667	4,667
R^2	0.0798	0.0809	0.0812

This table presents estimates from the regression of the likelihood of the incumbent's reelection and the fraction of split blocks in the assembly constituency (AC). The unit of observation is AC-year from 2008 until 2018. The data on election outcomes comes from the Election Commission of India. For each AC, we create a variable - fraction of split blocks. It is calculated as the ratio of the number of split blocks to the total number of blocks in the AC. We define a block as a split block based on the 2008 delimitation. Column (1) includes state × year fixed effect. Column (2) adds reservation fixed effect, which controls for whether the AC is reserved for lower castes. Column (3) adds a by-election fixed effect and controls for the election being outside the normal election cycle. Standard errors reported in parentheses are clustered at the AC level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.