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

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

Abstract

Unchecked political power impedes economic growth, and constraints on such power are inherently limited. We study whether increasing the number of politicians sharing authority over administration, a feature of political institutional design intended to constrain power, fosters economic growth. We implement discontinuity-based research designs, leveraging the haphazard overlap of electoral and administrative boundaries in India, which generates quasi-random variation in the number of politicians governing adjacent areas. We assemble administrative records to measure economic outcomes at high spatial resolution and find that areas governed by multiple politicians exhibit higher firm entry, higher employment, and greater economic activity. We find suggestive evidence that checks and balances operate. Economic benefits concentrate in areas governed by partisanly non-aligned politicians, who face stronger incentives to monitor one another. Politicians facing higher intensity of mutual oversight accumulate lower private returns from public office. We further find that firm entry improves along margins typically affected by unchecked political power. In particular, patronage and cronyism decline, regulatory hold-up eases, and delivery of essential public infrastructure improves.

*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 on Development Economics, 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 Imperial College London. email: sdutta@imperial.ac.uk

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

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

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

1 Introduction

"Ambition must be made to counteract ambition. . . . It may be a reflection on human nature, that such devices should be necessary to control the abuses of government. . . . If men were angels, no government would be necessary."

— James Madison, *Federalist No. 51* (1788)

Political institutions such as constitutions are inherently incomplete contracts. The authority these institutions delegate leaves discretion in the hands of those in power, creating scope for abuse that generates substantial growth costs. A long-standing concern in the design of political institutions has therefore been the extent of unchecked power delegated to those in authority and the mechanisms through which that power is disciplined (Montesquieu, 1751; Kant, 1795; Hayek, 1960). In this paper, we examine one widely used dimension of institutional design through which political power is constrained: the degree of horizontal power-sharing among elected representatives.[1] More specifically, we study whether varying the number of politicians who hold effective authority over the same administrative apparatus affects economic outcomes.

The relationship between the number of 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.[2] 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.[3] 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] Institutional arrangements where multiple politicians oversee the same administrative apparatus are common but understudied. Examples include the two-senator representation of each state in the U.S. Senate and Switzerland's seven-member Federal Council.

[2] 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.

[3] 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 estimating the effect of being governed by multiple politicians on firm entry and local economic growth, using a large-scale natural experiment in India. We take advantage of the nationwide haphazard overlap of electoral and administrative boundaries, which generates quasi-random variation in whether local administrative units fall under the jurisdiction of one or multiple elected representatives. Our central finding is that being governed by multiple politicians increases firm entry and economic growth. We find this positive effect to be stronger when politicians have greater incentives to impose mutual checks and balances on each other. We further show that such constraints foster firm entry in precisely those domains where unchecked political authority typically impedes it, reducing preferential access for connected firms, limiting cronyistic arrangements, and reducing regulatory hold-up in the issuance of permits.

The ideal research design for answering this question would require variation in the number of politicians exercising authority over otherwise similar administrative units, in a way that is orthogonal to local economic conditions. We exploit a unique setting in India that approximates such a scenario, arising from the haphazard overlap between electoral and administrative boundaries. *Administrative Blocks* are the basic governance units through which local policy implementation occurs. *Assembly Constituencies* are the electoral units through which representation in the state legislature is organized. Electoral boundaries are drawn by an independent, non-partisan Delimitation Commission with the objective of equalizing populations across all assembly constituencies within a state. When defining electoral boundaries, this commission does not take into account administrative boundaries at the block level. As a result, electoral boundaries may intersect with administrative blocks, leading some blocks to be divided across multiple constituencies (*Split Blocks*), while others are entirely contained within a single assembly constituency (*Unsplit Blocks*). This incongruence between electoral and administrative boundaries generates quasi-random variation in the number of politicians with authority over a given administrative block.

At the block level, the key administrative authority is the Block Development Officer (BDO), a bureaucrat who serves as the "eye of the needle" through which most public projects and administrative decisions must pass (Dasgupta and Kapur, 2020). BDOs exercise statutory authority over a range of local approvals, including permits that are pre-

requisites for economic activity and public service provision. Their interpretive authority in applying these regulatory requirements results in variation in administrative outcomes across blocks. State-level politicians (Members of the Legislative Assembly, or MLAs) are elected to represent their assembly constituencies. They face electoral incentives to deliver for constituents and opportunities to extract private benefits from public office (Fisman, Schulz, and Vig, 2014). However, MLAs are legislators and lack direct executive authority over policy implementation, so achieving these objectives requires bureaucratic implementation. MLAs therefore work through BDOs, using their influence over transfers and career advancement to induce bureaucratic compliance. This politician–bureaucrat nexus determines how policy is implemented at the block level.

The MLA–BDO interactions occur through both informal channels and regular block-level meetings, which provide a structured setting where MLAs can observe BDO performance and, when multiple MLAs are present, one another's behavior. When a single MLA's constituency encompasses an entire block, that MLA is the sole political principal overseeing the BDO, plausibly making it easier for the two to coordinate on rent-sharing arrangements with limited scrutiny, external exposure, or electoral discipline. By contrast, when a block spans multiple constituencies, several MLAs oversee the same BDO, making their actions mutually observable to political rivals with electoral incentives to monitor and expose self-serving behavior. These differences in mutual oversight create the possibility of systematically different patterns of rent extraction and policy implementation across blocks.

This institutional design allows us to use discontinuity based designs to test whether governance by multiple elected representatives causally affects local economic outcomes. The boundaries separating split blocks from unsplit blocks generate a sharp discontinuity in the number of elected representatives exercising authority over local administration, while other fundamental determinants of economic activity are assumed to vary smoothly across the boundary. Our primary design is a cross-sectional geographic regression discontinuity that compares villages just inside split blocks to those just outside in unsplit blocks, within a narrow bandwidth around the boundary. We augment this with a second strategy that relies on the timing of the 2008 delimitation, which redrew constituency boundaries nationwide and altered the number of elected representatives exercising authority within

blocks over time. This allows us to implement a differences-in-discontinuity design, comparing changes in outcomes over time for locations on either side of split–unsplit boundaries. The design allows us to how local economic activity responds when areas transition between single and multiple representative authority over local administration, while controlling for time-invariant local characteristics.

We assemble rich administrative data on firm entry from the Ministry of Corporate Affairs (MCA), Government of India, which maintains the statutory registry for the near-universe of the formal private corporate sector. Firm entry is our primary outcome variable and is well-suited for our analysis for three reasons. First, new firm creation is a key determinant of productivity growth and local employment (McMillan and Woodruff, 2003; Woodruff, 2018), and also serves as an indicator of whether the local regulatory environment facilitates or impedes productive economic activity (Djankov, La Porta, Lopez-de Silanes, and Shleifer, 2002; Djankov, 2009). Second, establishing a firm requires permits and approvals that fall under BDO discretion and MLA influence, making firm entry a direct reflection of how political oversight shapes the local business environment. Third, the MCA registry records precise registration addresses for each firm, which we geocode at high spatial resolution, enabling us to implement our discontinuity-based designs.

Our baseline analysis employs a geographic regression discontinuity (RD) design, comparing villages just inside split blocks to those just outside in unsplit blocks within a narrow bandwidth around the boundary. We find that villages in split blocks exhibit 3% higher firm entry and 7% higher nightlight intensity relative to villages in unsplit blocks. The effect increases monotonically with the number of representatives exercising authority over the block, consistent with a dose-response relationship.

We augment the cross-sectional estimates with a differences-in-discontinuity design that leverages the timing of the 2008 delimitation. This design examines how firm entry responds when areas transition from a single representative having authority to multiple representatives having authority over local administration. More specifically, we compare blocks that switched from unsplit to split following the delimitation against blocks that always remained unsplit, and separately compare blocks that switched from split to unsplit against blocks that remained split. We find that the transition from being governed by a single to multiple politicians is associated with a 1.1–1.6% increase in firm entry, while

the reverse transition is associated with a 0.7–1.0% reduction. This symmetry further strengthens our causal interpretation.

These effects extend to labor market outcomes. Villages in split blocks exhibit 5% higher employment rates, consistent with the increased firm entry documented above. We further examine applications to the National Rural Employment Guarantee Scheme (NREGA), India's large-scale public workfare program, which provides a direct measure of demand for public employment assistance. We find that split blocks exhibit lower NREGA application rates. This pattern is directionally consistent with Gulzar and Pasquale (2017), who document lower NREGA disbursement in areas governed by multiple politicians and interpret this as evidence of poor program implementation. However, the observed lower disbursement is an equilibrium outcome. The data on NREGA applications at our disposal allows us to isolate demand for such assistance. The lower application rates are consistent with our findings of increased firm entry and higher private sector employment. This pattern suggests that workers in split blocks have better outside options in the private labor market, reducing their demand for public employment assistance.

The causal interpretation of the RD estimates rests on two identifying assumptions. First, the underlying determinants of economic activity are assumed to vary smoothly at block boundaries. We conduct balance tests using data from the 2001 Census to support this assumption. Villages located just within split blocks closely resemble those in adjacent unsplit blocks across multiple dimensions, including population, public infrastructure, financial access, and geography. Second, we assume boundary placement was free of selective sorting, i.e., that the delimitation process did not systematically assign more productive areas to split or unsplit blocks. Iyer and Reddy (2013) documents that the 2008 delimitation was largely free from gerrymandering, supporting this assumption. Nevertheless, we restrict the sample to boundaries where this assumption is more plausibly satisfied: a) those following straight-line patterns, and b) those defined by natural features such as rivers and watersheds. We obtain similar estimates across these subsamples. Together, these tests support the assumptions underlying our causal estimates.

Next, we investigate the mechanism through which multiple politicians foster firm entry. We propose that multiple politicians holding effective authority over the same administrative apparatus impose mutual checks and balances that constrain rent extrac-

tion. The institutional structure facilitates such oversight: in split blocks, multiple MLAs participate jointly in block-level Panchayat Samiti meetings alongside the same BDO. This arrangement places politicians in a shared information environment where they can observe each other's interactions with bureaucrats. Politicians from rival parties face electoral incentives to monitor and expose self-serving behavior. For the BDO, the presence of multiple political principals with divergent interests makes exclusive rent-sharing arrangements difficult to sustain. We first establish that checks and balances operate by leveraging institutional features that induce variation in the intensity of mutual oversight. We then provide further evidence by examining domains where unchecked political discretion typically impedes firm entry.

We first leverage variation in the partisan composition of politicians governing split blocks. Politicians from different parties face stronger incentives to monitor one another, as they benefit electorally from exposing rivals' misconduct. Ideological differences and political animosity also create barriers to collusive arrangements. Checks and balances should therefore operate with greater intensity when split blocks are governed by non-aligned politicians. We test this by examining whether firm entry varies with partisan alignment among politicians sharing authority over the same block. We find that split blocks governed by non-aligned politicians experience significantly greater firm entry than those governed by aligned politicians. These patterns are suggestive of checks and balances in play.

We further examine whether politicians subject to greater oversight accumulate less private wealth. Politicians can extract rents through their discretionary authority over administrative processes, and such rent extraction often manifests in personal wealth accumulation. Electoral law in India requires candidates to file sworn asset affidavits, allowing us to measure changes in private wealth over a politician's term in office. Following Fisman, Schulz, and Vig (2014), we proxy for rent extraction through the winner's premium, the differential asset growth of election winners relative to runners-up. A politician's constituency contains multiple blocks. Haphazard boundary overlap implies that constituencies vary in the share of split blocks they contain, and hence in the degree of mutual oversight politicians face. We find that the winner's premium is lower in constituencies with a higher share of split blocks. Such constrained rent extraction

in constituencies with greater mutual oversight is suggestive of checks and balances in operation.

Having established that checks and balances operate, we examine how they foster firm entry. We focus on three domains where unchecked political discretion typically impedes firm entry: preferential treatment that favors connected agents, regulatory hold-up that delays permits, and public goods provision that suffers from rent extraction. First, we examine preferential treatment. Politicians wielding unconstrained authority can facilitate firm entry based on personal connections, potentially favoring co-ethnic entrepreneurs. We identify such potential patronage connections through shared caste (jati) identity, drawing on surname-caste associations from the People of India Project (Singh, 1996) and applying recent advances in language to process names. We observe that connected firms experience lower entry in split blocks. Additionally, we examine firm entry in industries prone to political favoritism and find that such industries experience relatively lower entry in split blocks. These patterns suggest that checks and balances constrain preferential treatment, expanding opportunities for firms outside patronage networks.

Second, we examine regulatory hold-up. Political discretion in the issuance of regulatory permits creates scope for rent extraction, as politicians could impose costly delays on firms seeking approvals (Djankov, 2009). We find that projects in split blocks receive regulatory approvals substantially faster than those in unsplit blocks. We further exploit heterogeneity across industries and firm sizes. Industries subject to extensive regulatory requirements face higher entry barriers where discretionary delays are prevalent. We find that such regulated industries experience disproportionately higher entry in split blocks. Similarly, regulatory compliance often involves fixed costs that impose greater burdens on smaller firms. We find that smaller firms experience disproportionately higher entry in split blocks, with effects declining monotonically with firm size. These patterns suggest that checks and balances reduce costs arising from discretionary regulatory implementation.

Third, we examine public goods provision. Public infrastructure serves as a productive input for private firms, yet the discretionary nature of implementation creates scope for rent extraction, often manifesting as cost overruns and resource misallocation.

We study road construction under the Pradhan Mantri Gram Sadak Yojana (PMGSY), a nationwide program where implementation is overseen by local authorities and political intervention is well documented (Lehne, Shapiro, and Vanden Eynde, 2018). We find that cost overruns decline substantially when blocks transition from single to multiple politician governance. We complement this with evidence on electricity: villages in split blocks are more likely to receive reliable commercial power supply. These patterns are consistent with improved efficiency in public goods provision when mutual oversight constrains rent extraction during implementation.

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 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 docu-

[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.

ment 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 Institutional Details

2.1 Administrative and Political Structure in India

India's governance system operates through a multi-layered administrative hierarchy that extends from the federal government down to the village level (see appendix Figure A.1 for a visual representation of this administrative structure). National and state governments formulate policies, while implementation occurs through a three-tier local governance system: the district, the block, and the village.[6]

Policy implementation at the local level occurs through blocks, which serve as key administrative units in India's governance system. Block Development Officers (BDOs)

[6]Districts, headed by District Collectors, coordinate programs across blocks, while villages serve as final delivery points for policy implementation.

administer these blocks and have statutory powers for policy implementation, serving as the "eye of the needle" through which government programs must pass (Dasgupta and Kapur, 2020). Parallel to this administrative structure is India's political system, where Members of Legislative Assembly (MLAs) represent electoral constituencies and serve as legislators in the State Legislative Assembly.

We focus on the interplay between block-level bureaucrats and politicians. Blocks occupy a unique position in the administrative hierarchy as the largest units that can either fall within a single constituency or span multiple constituencies. This creates our key variation: some blocks fall under the jurisdiction of one MLA while others fall under the jurisdiction of multiple MLAs. BDOs are accountable to all MLAs whose constituencies encompass their blocks. The quasi-random variation in political oversight forms the foundation of our research design.

In the rest of this section, we describe India's administrative and political boundaries, the formal powers held by bureaucrats and politicians at different levels, and the interaction between politicians and bureaucrats that shapes governance outcomes.

2.2 Overlap of Electoral and Administrative Boundaries

Local governance and policy implementation are territorially organized at the block level, while political representation operates through constituency boundaries. Political boundaries are determined by the Delimitation Commission, which redraws electoral boundaries to equalize population across all constituencies within a state. The administrative boundaries of a block are not considered during the design of constituencies. This independent boundary-drawing process results in a haphazard overlap between the constituency and block 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 a constituency comes under the jurisdiction of the MLA.[7] As a result of this overlap, blocks are either under the jurisdiction of multiple MLAs (split blocks) or a single MLA (unsplit blocks), as shown in appendix

[7]Note that although a block can be under the jurisdiction of multiple MLAs from different constituencies, individuals can only vote for candidates in one constituency, depending on the village or town they reside in.

Figure A.3. The local bureaucrats are accountable to all the MLAs with jurisdiction over the block, regardless of whether a block is wholly or partially included in the 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 Amravati, Maharashtra. The haphazard overlap of constituency and block boundaries generates the variation in the number of MLAs with jurisdiction over 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. The variation in whether a block is split or unsplit is the primary source of heterogeneity we exploit in this paper.

2.3 Administrative Authority at the Block Level

Blocks represent a key implementation tier in India's governance system, where government programs and policies are implemented at the local level. Block Development Officers (BDOs), appointed by the state government, serve as the primary administrative authority at this level with discretionary power in interpreting and implementing government directives.[8] BDOs derive their authority from state legislation that establishes their position as senior gazetted officers responsible for implementing federal and state development programs.

BDOs exercise effective control over licensing and other implementation decisions through their supervisory authority over village-level administration. While many procedures, including initial permit applications, are nominally processed at the village level, BDOs hold decisive influence through budget approval powers, audit authority, and inspection directives that village institutions must comply with. Through mandatory budget approvals, BDOs can control the resources available for processing applications, while their audit powers enable them to "disallow" local decisions and impose personal financial liability on village officials who do not align with BDO preferences.[9] This super-

[8] In Karnataka, Section 155(1) mandates that the state government appoint a Group-A Officer of the State Civil Services to serve as the Executive Officer of the Taluk Panchayat (block). Similar provisions exist across states, with variations in specific rank requirements and appointment procedures; see Appendix A.1 for a detailed discussion of BDO powers.
[9] See Appendix A.1.2 for a comprehensive description of BDO oversight powers over Grama Panchayats.

visory framework ensures that BDOs function as the relevant administrative authority for decisions that ultimately affect local economic activity.

In addition to program implementation, BDOs control permits and approvals that are prerequisites for business creation, including trade licenses, factory establishment permissions, and building construction permits.[10] These permits serve as prerequisites for bank accounts, state-level clearances, and other business requirements, creating cascading dependencies through which BDO decisions determine the timing and feasibility of firm entry.

2.4 Political Representatives and Administrative Oversight

India operates a parliamentary democracy with elected representatives at multiple levels of government. At the state level, Members of Legislative Assembly (MLAs) are elected from single-member constituencies to serve five-year terms in their respective State Legislative Assemblies. MLAs derive their authority from India's constitutional framework, which grants them legislative power to make laws and oversee government policy. Executive authority for policy implementation is vested in the bureaucracy.[11] MLAs must therefore work through bureaucrats to achieve local governance outcomes in their constituencies.

MLAs face electoral incentives to demonstrate effective constituency service and development outcomes. In India's competitive democracy, voters evaluate MLAs based on visible improvements in infrastructure, public services, and economic opportunities, with electoral punishment for those who fail to deliver (Bussell, 2019). Simultaneously, political office provides MLAs with opportunities for private benefit extraction through their position and control over bureaucratic decision-making (Fisman, Schulz, and Vig, 2014). Both electoral success and private benefit extraction require effective bureaucratic implementation.

Despite constitutional separation of powers, the system provides MLAs with mechanisms to influence bureaucratic behavior. MLAs exercise oversight through formal leg-

[10] See Appendix A.2 for a detailed description of the subset of local business permits that fall under block-level administration and BDO oversight.

[11] MLAs do not possess direct executive authority and cannot issue administrative orders to bureaucrats or directly implement policies. This follows the Westminster parliamentary system where the legislature and executive are constitutionally separate branches.

islative institutions and can affect bureaucrat career progression through the transfer and posting system. Block Development Officers and other state-level bureaucrats are subject to periodic transfers, and MLAs provide influential feedback on bureaucrat performance to state government officials who make transfer decisions (Iyer and Mani, 2012). This system enables MLAs to influence bureaucratic actions, as non-compliance with MLA preferences can result in unfavorable career consequences.

2.5 Governance Structures and Barriers to Firm Entry

Firm entry in India requires extensive regulatory approvals across multiple tiers of governance.[12] At the block level, BDOs exercise statutory authority over the issuance of local permits essential for business establishment, including trade licensing, factory establishment permissions, and building construction permits. The BDO's interpretive authority over the issuance of regulatory permits creates differential costs for firms, with identical business applications facing different processing timelines and requirements despite uniform policy frameworks.

State-level politicians (MLAs) have strong incentives to shape which firms are able to enter and operate in their constituencies, since influence over the allocation of business permits generates opportunities for private rent extraction while in office (Sukhtankar and Vaishnav, 2015). However, MLAs operate as legislators and lack direct executive authority over policy implementation. [13] Consequently, MLAs must work through the local bureaucrats who hold statutory authority over permit issuance.[14] By credibly promising job security and favorable treatment to compliant bureaucrats, MLAs can influence the enforcement and interpretation of regulatory requirements. This politician–bureaucrat nexus determines the effective barriers and costs of firm entry at the block level.

The MLA–BDO interactions occur through both informal channels and formal institutional venues. The Panchayat Samiti, an intermediate tier of local governance at the

[12] The World Bank's Ease of Doing Business Report (Bank, 2009) provides comprehensive documentation of the regulatory procedures required for business establishment across Indian cities. These costs are substantial: building permits and utility connections in Mumbai alone cost 2,717.8% of GDP per capita and require approximately 200 days. We refer readers to the report for the complete list of procedures.

[13] In practice, MLAs devote little time to legislative activities: only 3% of surveyed MLAs report assembly work as their primary focus, and they spend most of their time interacting with local bureaucrats and addressing constituents (Chopra, 1996).

[14] This aligns with the argument put forth by Chopra (1996) that MLAs often act as 'fixers' who work with bureaucrats to allocate government-produced goods.

block level, serves as the primary formal venue for these interactions. MLAs whose constituencies overlap with a block sit as ex officio members of that block's Panchayat Samiti, while the BDO serves as the committee's executive officer. Panchayat Samiti meetings thus create a structured setting in which MLAs regularly observe BDO performance and, when multiple MLAs are present, can also observe one another's behavior.

When a single MLA's constituency encompasses an entire block, that MLA serves as the sole political representative in the Panchayat Samiti. In this setting, the MLA and BDO can coordinate over rent-sharing arrangement, without facing scrutiny from rival political principals. The absence of competing political principals within the block, combined with the fact that much of this behavior operates through low-visibility bureaucratic channels, limits both external exposure and electoral discipline. As a result, patterns of rent extraction can persist even when they impose substantial aggregate costs.

When a block spans multiple constituencies, multiple MLAs sit as ex officio members of the same Panchayat Samiti. This institutional arrangement creates conditions under which each MLA's actions and demands become observable to rival politicians through regular committee meetings. MLAs with diverging political interests face electoral incentives to monitor each other and publicize behavior that appears self-serving. For the BDO, the presence of several elected principals with potentially divergent interests makes it harder to sustain an exclusive rent-sharing arrangement with any single MLA. The heightened monitoring and coordination difficulties might make MLAs reluctant to pursue extractive arrangements that risk exposure by political rivals. Consequently, blocks with multiple MLAs are more likely to experience reduced rent extraction and lower barriers to firm entry than blocks represented by a single MLA

3 Data

We assemble several administrative datasets that measure economic and political activity at fine spatial resolution across India. This section provides a brief overview of the most important datasets and describes their use in subsequent analyses. We provide further details in Appendix B.3.

3.1 Administrative and Political Boundaries

3.1.1 Identifying Split Blocks

We obtain geospatial data on administrative blocks from MLInfomap and on electoral constituency boundaries from DataMeet for both the 1977 and 2008 delimitations. Administrative blocks represent the unit of local governance in India, with each block overseen by a Block Development Officer (BDO) who manages regulatory approvals and public program implementation. Assembly constituencies define the electoral boundaries for state legislative elections, with each constituency electing one Member of Legislative Assembly (MLA).

We overlay these shapefiles to identify which blocks fall under the jurisdiction of multiple politicians. Following the institutional structure that a block wholly or partially located within an assembly constituency falls under that constituency's elected representative, we define a block as *Split* if it is part of more than one assembly constituency. To account for discrepancies in boundary alignment across data sources, we require each constituency partition to occupy at least 10% of the block's geographic area when classifying a block as *Split*.

Appendix Figures A.2 and A.3 illustrate this procedure for the state of Karnataka. Appendix Figure A.2a shows boundaries for the administrative block, while Appendix Figure A.2b shows boundaries for electoral constituency within Karnataka. Appendix Figure A.3a displays the overlay of these two boundary systems, and Appendix Figure A.3b identifies the resulting split and unsplit blocks. Figure 2 illustrates a case of adjacent split and unsplit blocks in our sample.

Split blocks constitute 39.43% of our sample and are not geographically concentrated (Appendix Figure A.4). These blocks account for 50% of the 2001 population and 40% of total geographic area. Among *Split* blocks, 83.08% are governed by two politicians, 14.83% by three politicians, and 2.09% by more than three politicians (Appendix Figure A.5).

We use villages as the primary unit of observation for our analysis. We obtain village boundaries from the 2001 Census provided by Meiyappan, Roy, Soliman, Li, Mondal, Wang, and Jain (2018). We map villages to their corresponding blocks and assembly

constituencies by overlaying the shapefiles. Using village centroids, we calculate each village's distance to the nearest boundary separating a *Split* from an *Unsplit* block, which serves as the running variable in our regression discontinuity design. The village-level resolution allows us to estimate local polynomial within narrow bandwidths around block boundaries, comparing outcomes in villages that are geographically proximate but experience differential political oversight.

3.1.2 Exogenous Block Boundaries

We further refine our identification strategy by restricting analysis to boundaries with characteristics that preclude strategic placement based on economic conditions. We employ two complementary approaches: (1) identifying boundaries defined by geographic features - in this case, river segments, and (2) identifying straight-line-like boundaries. River boundaries constitute natural barriers that predate modern economic development and are unlikely to reflect strategic placement to separate economically distinct areas. Straight-line boundaries reflect historical surveying conventions that imposed geometric regularity without regard to local topographic or economic features. These sample restrictions provide robustness checks on our baseline estimates.

First, we identify boundaries determined by river courses using high-resolution geospatial data on river segments from HydroATLAS (Linke, Lehner, Ouellet Dallaire, Ariwi, Grill, Anand, Beames, Burchard-Levine, Maxwell, Moidu, Tan, and Thieme, 2019). We classify a boundary segment as determined by river segments if it aligns with a river for at least 80% of its length and the river has average discharge greater than or equal to 200 m^3/s (cumecs). The discharge threshold ensures the river is large enough to serve as a natural territorial delimiter. This definition yields 3.02% of boundaries and 5.74% of blocks from our sample. Figure 3 presents representative examples of block boundaries determined by river courses.

Second, we identify straight-line boundaries. We calculate the Sinuosity Index (SI) for each boundary segment—the ratio of actual boundary length to the Euclidean distance between its endpoints. Boundaries with SI values less than 1.05 exhibit near-straight characteristics. We include boundaries with SI less than 1.05, yielding 6.24% of boundaries

and 10.68% of blocks. Appendix Figures B.5 and B.6 illustrate examples of straight-line and irregular boundary segments.

3.2 Universe of Firm Registration

We utilize data from the Ministry of Corporate Affairs (MCA), Government of India, which administers the statutory registry for companies incorporated under the Companies Act.[15] Our dataset covers approximately 850,000 firms established between 2003 and 2016, providing a near-census of the formal private corporate sector in India.[16] This registry serves as the foundation for India's national accounts estimates of the private corporate sector.

The Companies Act requires all registered companies to file incorporation documents and ongoing returns with the Registrar of Companies, including registration dates, operational addresses, financial statements, and information on directors and key managerial personnel. The MCA records include each firm's precise operational address and date of incorporation, enabling us to construct a spatio-temporal dataset of firm entry across India. The data also contain detailed information on company directors, including the names of these directors, which we leverage to infer the social identities of key personnel in charge of operating these firms.

3.2.1 Geolocating Firm Entry

The MCA registry provides information on the address of the registered office for each firm, which includes street addresses and six-digit postal codes (pincodes). This allows us to locate these firms in space and measure private sector activity at highly granular spatial resolution, an improvement over existing datasets on private sector activity in India, which typically provide aggregates at the state or district level.

We start constructing village-level measures of firm entry by first geolocating firm addresses into spatial coordinates. [17] We approximate the location of a firm by that of

[15] The MCA electronic filing system maintains comprehensive records for all registered companies, including private limited companies, public limited companies, and one-person companies.

[16] The dataset excludes entities not subject to the Companies Act, such as sole proprietorships, ordinary partnerships, and cooperatives, which constitute a substantial share of India's non-corporate enterprises.

[17] The MCA records provide addresses for each firm's registered office but do not include latitude and longitude coordinates.

the post office corresponding to its pincode. Geocoordinates for post offices are obtained using the Google Maps API. Appendix B.1 provides detailed discussion of the geolocation procedure, potential measurement errors associated with pincode-based assignment, and validation exercises indicating these errors are unlikely to significantly affect our results.

With these firm-level coordinates, we then map firms to villages by performing spatial joins with village boundary shapefiles. This procedure enables us to aggregate firm counts at the village level, creating a spatio-temporal dataset of firm entry by location and year from 2003 to 2016. Appendix Figure B.4a and Appendix Table B.6 present the spatial and temporal distribution of firm entry in our sample.

3.2.2 Firm Directors and Social Identity

The Companies Act requires firms to file information on their directors and key managerial personnel with the Registrar of Companies, which includes the names of directors. Surnames in India often carry identity markers that signal caste affiliation, as well as broader religious identities.[18] We leverage this feature of the data to infer the social identities of firm directors and construct measures of patronage connections between firms and local politicians.

We rely on recent advances in AI/ML methods and combine rich ethnographic documentation as well as large-scale datasets on self-reported identity affiliations to link names to social identities. This enables us to classify firms based on whether their directors share social identity markers with the local politician in power. We further discuss this in Section 5.2.1. Appendix B.2 provides a detailed discussion of the methodology followed to infer social identity.

3.3 Auxiliary Data

3.3.1 Data on Economic Activity

NREGS Application Data: The National Rural Employment Guarantee Scheme (NREGS) is the world's largest employment guarantee program, under which the government serves

[18] In India, *jati* — endogamous subcastes that constitute narrowly defined groups — represents a salient identity around which social and political networks form (Munshi, 2019).

as employer of last resort. We collect data on individual-level applications filed for unemployment relief covering the period 2016 to 2020. The number of days for which applicants seek work through NREGS provides a measure of demand for unemployment assistance. To our knowledge, we are the first to collect and use application-level NREGS data.

Harmonized Nightlights. We use global nighttime light data from Li, Zhou, Zhao, and Zhao (2020), which harmonizes data from the Defense Meteorological Satellite Program (DMSP) and the Visible Infrared Imaging Radiometer Suite (VIIRS).[19] We construct village-level measures of average nightlight intensity from 2003 to 2016, which serve as a proxy for economic activity (Donaldson and Storeygard, 2016).

CMIE CapEx Data. We obtain data on investment projects from the Centre for Monitoring Indian Economy (CMIE) Capital Expenditure database, covering projects announced from 2003 to 2016 that involve capital expenditure of Rs. 10 million or more. For each project, the database records announcement dates, regulatory approval dates, and implementation milestones. We use the time between project announcement and initial regulatory approval as a measure of bureaucratic processing delays.

Economic Census (2013). We use data from the 2013 Economic Census of India (Central Statistics Organisation, Ministry of Statistics and Programme Implementation, Government of India, 2013) , which provides data on non-farm employment at the village level.[20] These data were accessed through the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG) (Asher, Lunt, Matsuura, and Novosad, 2021).

3.3.2 Data on Legislative Elections

Electoral Data: We use data on politicians and election outcomes from the Trivedi Center for Political Data (Jensenius and Verniers, 2017a), accessed through SHRUG (Asher, Lunt, Matsuura, and Novosad, 2021). This dataset provides election results, party affiliations,

[19] DMSP provides coverage until 2013, while VIIRS coverage starts from 2012. The transition between these satellite systems occurs within our study period (2003–2016), necessitating harmonization to ensure consistent measurement of nightlight intensity over time.

[20] The 2013 Economic Census is the first comprehensive economic census conducted after the 2008 delimitation.

and names of candidates for state legislative assembly elections. We use this information to construct measures of political alignment, defined as the degree to which multiple politicians governing split blocks belong to different political parties. We also use candidate names to infer their social identities and measure connections to entrepreneurs entering their jurisdictions.

Candidate Affidavits: We obtain data on candidate characteristics from sworn affidavits filed with the Election Commission of India, which were digitized by the Association for Democratic Reforms (ADR).[21] These affidavits include self-reported assets and liabilities, allowing us to measure changes in private wealth over a politician's term in office by comparing declarations across elections.

3.3.3 Data on Public Good Provision

PMGSY Road Construction: We obtain administrative data from the Pradhan Mantri Gram Sadak Yojana (PMGSY) rural road construction program, accessed through SHRUG (Asher and Novosad, 2020; Asher, Lunt, Matsuura, and Novosad, 2021). The data cover road projects from 2001 to 2014 and include estimated costs, actual costs, and completion timing. We use cost and time overruns as measures of inefficiency in road construction, testing whether split blocks experience more efficient public goods provision.

Population Census: We use village-level data from the 2001 and 2011 Population Census of India (Office of the Registrar General and Census Commissioner, India, 2011), accessed through SHRUG (Asher, Lunt, Matsuura, and Novosad, 2021). The Population Census provides village-level data on demographics and public infrastructure.

[21] Electoral law in India requires candidates to file sworn asset affidavits (Form 26) at the time of nomination (Section 33A, Representation of the People Act, 1951, inserted following the Supreme Court's 2002 ruling in *Union of India v. Association for Democratic Reforms*). The ADR dataset is available under the Open Database License at https://github.com/Vonter/india-election-affidavits.

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.[22] 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 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.

[22] 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 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 × 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 C.14 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 B.5 and B.6 present a representative example 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.[23]

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 C.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 C.2. Appendix Table C.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 C.1 shows the results are robust to choosing any bandwidth between 2 km and 50 km. Appendix Table C.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 C.5.

Second, we verify that the results are not an artifact of the sample and transformation of the key dependent variable. Appendix Table C.6 shows the results are robust to dropping small and large blocks, indicating the outliers are unlikely to drive our results.[24] Panel A of appendix Table C.7 shows results are robust to re-estimating our baseline specification using Poisson regression.[25]. Panel B of appendix Table C.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.[26] In appendix

[23] 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.

[24] 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.

[25] 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.

[26] The IHS transformation has been used to overcome problems in regression analyses with right-skewed censored

Table C.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 C.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 C.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).[27] Appendix Table C.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 Table C.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

dependent variables like ours (Carboni, 2012) and to address the issue of log transformation with zeros (Burbidge, Magee, and Robb, 1988).

[27] 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.

unsplit. Appendix Table C.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 C.2. Appendix Table C.16 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 C.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 identified 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 C.17 provide support to our argument.

5 Mechanism

This section investigates the mechanism through which multiple politicians foster firm entry. Firm entry involves engaging with local governance apparatus for regulatory compliance and resource access. Politicians hold discretion over how such policies are interpreted and implemented, creating opportunities for rent extraction, affecting firm entry. When multiple politicians share governance over the same area, they can observe each other's actions and impose mutual checks and balances, limiting individual rent

extraction. We first establish that checks and balances are operative, by leveraging institutional features that induce variation in mutual oversight intensity. We show that firm entry and politician rent extraction respond to this variation. We then provide further evidence that checks and balances potentially enhance firm entry by examining domains where unchecked political discretion deters firm creation.

5.1 Checks and Balances

The institutional structure facilitates checks and balances in split blocks, where multiple electoral constituencies overlap across the same administrative block. The MLAs governing these constituencies operate alongside the same BDO and are required to participate jointly in block-level Panchayat Samiti meetings. Such an arrangement places politicians in a common information environment where each can observe the others' interactions with bureaucrats, creating mutual observability of administrative decisions.

We begin by providing suggestive evidence that checks and balances operate by leveraging institutional features that generate variation in the intensity of checks and balances. First, we rely on partisan composition of politicians governing split blocks. Politicians from different parties have stronger incentives to monitor one another, as they benefit electorally from exposing rivals' wrongdoings and face greater barriers to collusive arrangements. Consistent with this, we find that split blocks governed by non-aligned politicians — those from different parties — experience significantly higher firm entry compared to those governed by aligned politicians from the same party. Second, we leverage variation in the share of split blocks across constituencies. Politicians governing constituencies with a higher share of split blocks face greater oversight, as more of their administrative decisions occur in blocks that overlap with other constituencies where rival politicians can observe their actions. Consistent with this, we find that politicians governing such constituencies exhibit lower gains private wealth.

5.1.1 Non-Alignment Across Politicians

We first examine variation in the intensity of checks and balances generated by partisan composition of politicians governing split blocks. Politicians from different parties have

stronger incentives to monitor each other, as they gain electoral advantage by exposing rivals' corrupt behavior. Moreover, ideological differences and political animosity between parties create barriers to collusive rent extraction, making coordinated corruption more difficult to sustain (Spenkuch, Teso, and Xu, 2023; Kempf and Tsoutsoura, 2021). Conversely, politicians from the same party can more easily coordinate rent-seeking arrangements because they share electoral interests and face fewer coordination barriers. Partisan composition thus generates variation in the intensity of checks and balances across split blocks. We examine whether this variation corresponds with patterns of firm entry.

The variation in political alignment operates both across split blocks at any given time and within split blocks over time, as party composition of governing MLAs changes around elections. We estimate specification 1 augmented with the interaction between split blocks and the fraction of non-aligned politicians. The specification includes boundary × year and village fixed effects, relying on within-village variation after 2008. The coefficient of interest is estimated from changes in the fraction of non-aligned politicians over time while controlling for boundary-specific investment opportunities. Table 6 reports estimates showing that split blocks with non-aligned politicians experience significantly greater firm entry than those with aligned politicians. These estimates are consistent with checks and balances operating to increase firm entry

5.1.2 Private Returns of Politicians

We provide further evidence that checks and balances operate by examining whether politicians subject to stronger oversight accumulate less private wealth. Politicians can extract rents through their discretionary authority over administrative processes, and such rent extraction often manifests in personal wealth accumulation. Electoral law in India requires candidates to file sworn asset affidavits at the time of nomination, allowing us to measure changes in private wealth over a politician's term in office. Following Bhavnani (2012) and Fisman, Schulz, and Vig (2014), we proxy for rent extraction through the winner's premium, defined as the differential asset growth of election winners relative to runners-up over a term. By comparing similar candidates who differ only in electoral success, this measure isolates wealth gains from holding office.

A politician's constituency contains multiple blocks. Haphazard overlap of boundaries implies that constituencies vary in the extent to which they share split blocks with the adjacent constituencies. Politicians governing constituencies with more split blocks face greater oversight from politicians in the overlapping constituencies. We test whether the winner's premium varies across constituencies with varying intensity of such mutual oversight. A key challenge in estimating the winner's premium is accounting for unobserved skills or resources that politicians possess independent of electoral success. We address this by restricting our sample to elections with a margin of victory below 5 percent. This allows us to compare politicians with plausibly similar characteristics who differ primarily in whether they won office.

Table 7 presents estimates examining how rent-seeking (winner's premium) is affected by the intensity of checks and balances. We construct *Splitness*, which captures the intensity of checks and balances that a constituency-level politician faces. Splitness is measured two ways: (i) a continuous measure—the normalized share of the constituency's population living in split blocks—and (ii) a binary indicator equal to one if this share exceeds 50% and zero otherwise. Columns 1 and 2 report estimates separately for constituencies with differing intensity of checks and balances. In constituencies with low intensity of checks and balances (Splitness = 0), we observe a 16% increase in the assets of the winner relative to the runner-up. We find no statistically significant winner's premium in constituencies with high intensity of checks and balances (Splitness = 1). Columns 3 and 4 utilize the complete sample and include a specification with an interaction between the Winner and the Splitness variable to test whether the winner's premium varies with the intensity of checks and balances. The negative estimate associated with the interaction term indicates that the winner's premium decreases in constituencies with a higher share of split blocks. These estimates suggest that greater intensity of mutual oversight reduces politicians' private wealth accumulation, consistent with checks and balances operating to constrain rent extraction.

5.2 How Do Checks and Balances Increase Firm Entry?

We provide further evidence for checks and balances in operation by examining domains where unchecked political discretion impedes firm entry. First, we document reduced preferential treatment in market access, showing increased entry of firms outside politicians' co-ethnic networks and lower entry in industries prone to cronyism. Second, we provide evidence of reduced regulatory barriers, finding shorter approval delays and higher entry in industries with greater regulatory intensity where unchecked power typically creates hold-up. Third, we demonstrate improved efficiency in public good provision, which serves as a productive input for private firms while representing a domain where political rent extraction is prevalent.

5.2.1 Reduced Preferential Treatment

Firm entry typically requires obtaining regulatory permits across different levels of governance, including at the block level where local politicians direct implementation through bureaucrats. When political institutions provide insufficient constraints on such authority, politicians gain discretionary power that often leads to preferential allocation based on personal connections. Such preferential treatment can manifest through patronage networks that favor co-ethnic individuals and cronyistic arrangements that benefit politically valuable industries. Checks and balances among politicians might therefore foster firm entry by constraining preferential allocation arising from such discretionary authority.

India's political landscape is characterized by ethnic networks that influence resource allocation (Chandra, 2007), while well-documented quid pro quo arrangements exist between politicians and rent-seeking industries (Kapur and Vaishnav, 2013; Crabtree, 2018). Against this backdrop, we examine firm entry across contexts in which preferential treatment would create differential advantages to connected agents. First, we study relative entry of firms with directors sharing caste identity with local politicians, where co-ethnic connections might facilitate preferential access. Second, we exploit heterogeneity across industries prone to cronyism, where rent-seeking sectors typically receive political favoritism.

Lower Entry of Patron Firms Political patronage often operates through shared identities that facilitate preferential access to resources. In India, *jati* — endogamous subcastes that constitute narrowly defined groups — represents a salient identity around which social and political networks form (Munshi, 2019). We proxy for patronage by identifying firms led by entrepreneurs that share the same jati with local politicians, classifying such firms as potentially connected through patronage.

Surnames in India often carry identity markers that signal jati affiliation. We use the names of firm directors and politicians to deduce whether they share common identity markers. We rely on the People of India Project (Singh, 1996), a comprehensive anthropological study that systematically documented 2,205 jatis with their associated surnames at the state level. The salience of jati as an identity is particularly pronounced for the Hindu community compared to other religious groups, and surname-caste associations are more systematic in North Indian states compared to South Indian states.[28] This limits our sample to blocks in North Indian states governed by Hindu politicians, where we can reliably establish caste connections through surname analysis. Processing Indian names presents substantial complexity, hence we leverage recent advances in language processing tools to identify potential patronage connections between politicians and firm directors.[29] A detailed description of our methodology is provided in Appendix B.2.

We test whether patterns of firm entry are consistent with reduced patronage by estimating specification 1 augmented with the interaction term Split × Connected and include boundary × connected and village fixed effects. Connected is an indicator variable equal to one for firms with directors sharing caste identity with the local politician and zero otherwise. Estimates in Table 8 show that patronage-connected firms experience 0.9 to 1.3 percentage points lower entry relative to non-connected firms in split blocks compared to unsplit blocks. This pattern suggests that the relative advantages of patronage-connected firms diminish in split blocks, consistent with checks and balances constraining preferential treatment when multiple politicians potentially provide mutual

[28] South Indian naming conventions reduce the informativeness of surnames for caste inference. In many southern states, people commonly use patronymics and initials (e.g., father's given name as an initial) rather than fixed, hereditary surnames; caste suffixes that were historically used have been legally discouraged and socially abandoned; and names may include neutral given names or village identifiers. By contrast, North Indian names more often retain inherited family surnames that correlate with jati, making surname-based inference more reliable.

[29] Challenges include varying surname conventions across states, non-standardized naming formats, transliteration differences creating multiple spelling variants, and complex name structures.

oversight.

Lower Entry in Industries Prone to Cronyism Cronyistic arrangements, where industries offer rents in exchange for preferential treatment, are symptomatic of political favoritism. Industries vary in their vulnerability to cronyism, with sectors characterized by high rents, extensive state interaction, and opaque regulations being structurally more susceptible. We rely on such heterogeneity to examine relative entry patterns across industries that vary in their propensity for cronyism. We capture this by estimating specification 1 augmented with interactions between Split and Crony. We also include village and boundary × industry fixed effects. Crony is an indicator for industries vulnerable to cronyism.[30] Estimates in Table 9 show that the differential entry between industries vulnerable to cronyism and those less vulnerable is 6 percentage points lower in split blocks compared to unsplit blocks. This pattern suggests that multiple politicians foster firm entry by constraining (rather than enhancing) cronyistic arrangements, consistent with checks and balances operating to limit political favoritism towards connected sectors.

5.2.2 Reduced Hold-Up in Regulatory Implementation

Political discretion in the issuance of regulatory permits enables hold-up opportunities by inducing artificial delays. Such regulatory implementation is a documented source of high entry costs in developing economies (Djankov, La Porta, Lopez-de Silanes, and Shleifer, 2002; Djankov, 2009; Campos, Estrin, and Proto, 2010). Checks and balances among politicians might therefore foster firm entry by reducing costs arising from such discretionary implementation of regulations.

We examine firm entry across contexts in which regulatory discretion would create distinct cost burdens. First, we study approvals for regulatory permits, where discretionary delays impose direct costs. Second, we exploit heterogeneity across industries in regulatory intensity, where industries requiring more approvals face higher entry barriers. Third, we examine size-based heterogeneity, where smaller firms' entry is disproportionately affected by fixed regulatory costs because of capital constraints.

[30] We use the industry-level cronyism index created by *The Economist* using *Transparency International* methodology. See Planet Plutocrat. Appendix Table B.9 lists industries with higher vulnerability to cronyism in our sample.

Faster Regulatory Approvals: Approval times for regulatory permits can reflect costs arising from discretionary implementation. We examine whether the presence of multiple politicians reduces such discretionary delays by measuring the time required for private firms to receive regulatory approvals for their projects. Using the CapEx database, we compute approval time as the difference between project announcement and initial approval dates. Table 10 reports estimates examining blocks that transition between single and multiple politician governance compared to blocks maintaining constant governance arrangements, following equations 2 and 3. We find that approval times decrease by 80% when blocks switch from single to multiple politician governance, and increase by 57% in the reverse transition. These patterns suggest that checks and balances may reduce discretionary delays in regulatory implementation.

Higher Entry in Regulated Industries Regulatory complexity varies across industries. Those with more complex regulatory requirements are potentially more vulnerable to administrative delays that create barriers to entry. We rely on this heterogeneity to examine relative entry patterns across industries with different regulatory intensities. To capture this, we estimate specification 1 augmented with interactions between Split and Regulated. We also include village and boundary $\times$ industry fixed effects. Regulated is an indicator for industries subject to extensive regulatory requirements.[31] Estimates in Table 11 show that the differential entry between highly regulated and less regulated industries is 3.5 percentage points higher in split blocks relative to unsplit blocks. This pattern suggests that checks and balances may promote entry of firms in industries where regulatory hold-up is more likely to create entry barriers.

Higher Entry for Small Firms Regulatory compliance often involves a fixed cost component that impose disproportionate barriers on smaller firms due to their capital constraints. When regulatory costs decline through enhanced checks and balances, smaller firms ought to experience relatively higher firm entry. We group firms into size deciles, with the first (tenth) decile referring to firms in the smallest (largest) size group. We augment regression specification 1 with interaction terms between Split and each size decile. We also

[31] We define high regulatory cost industries broadly following Pittman (1977) methodology augmented with India-specific factors from Awasthi, Yayavaram, George, and Sastry (2019). Appendix Table B.8 lists classified industries.

include village and boundary × size decile fixed effects. Figure 6 shows that differential entry is higher for smaller firms, with the effect decreasing monotonically with firm size. This pattern suggests that checks and balances operate to reduce regulatory barriers with smaller firms experiencing disproportionately larger improvements.

5.2.3 Improved Efficiency in Public Goods Provision

Public goods serve as productive inputs for private firms (Munnell, 1992). Provision of public goods in developing countries is frequently characterized by inefficient delivery, cost overruns, and resource misallocation due to rent extraction by politicians and bureaucrats (Olken, 2007). We examine two domains — provision of roads and electricity — that represent key inputs to production while also being sectors prone to political rent extraction (Lehne, Shapiro, and Vanden Eynde, 2018; Mahadevan, 2024). We provide evidence that patterns of public goods provision are consistent with improved efficiency when multiple politicians govern the same jurisdiction, suggesting that mutual checks and balances may enhance the delivery of such essential productive inputs.

The *Pradhan Mantri Gram Sadak Yojna* (PMGSY) is a large nationwide road construction program. While provisioned by the Central Government, implementation is overseen by local authorities, creating opportunities for rent extraction during the construction process. The program has faced significant corruption allegations, with the Comptroller and Auditor General reporting construction delays, cost overruns relative to initial bids, and award of contracts to dubious contractors.[32] We examine whether the presence of multiple politicians in a jurisdiction is associated with patterns consistent with improved efficiency in road construction within the program. We measure cost overruns — our proxy for implementation efficiency — as the ratio of actual implementation costs to estimated costs at the time of tender.

We rely on the 2008 Delimitation and use a differences-in-discontinuity design discussed previously in Section 4.3. Panels A and B of Table 12 report estimates from specifications (2) and (3), respectively, augmented with boundary × year and block fixed

[32]Standing Committee on Rural Development and Panchayati Raj, "Pradhan Mantri Gram Sadak Yojana (PMGSY)," Sixteenth Lok Sabha, March 20, 2017.

effects.[33] Estimates in Panel A across all bandwidths are negative and statistically significant. These estimates indicate that 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. Estimates reported in Panel B are positive and indicate an increase in cost overruns of 3-5% when a block switches from being governed by multiple politicians to a single politician. Taken together, estimates from Table 12 suggest that governance by multiple politicians is associated with improved efficiency in public goods provision, patterns consistent with enhanced mutual oversight constraining rent extraction during implementation processes.

We supplement this analysis by examining electricity provision for commercial use. Commercial power supply serves as another essential productive input that is state-controlled, creating scope for rent extraction through inefficient delivery. Appendix Table D.1 reports estimates using reliable commercial power supply as the dependent variable.[34] The estimates are positive and statistically significant, indicating that villages in split blocks are 0.8 - 1% more likely to receive reliable commercial power relative to villages in unsplit blocks. These complementary findings across road construction and electricity provision indicate that public goods delivery patterns are consistent with improved efficiency when multiple politicians govern the same jurisdiction. The evidence suggests that mutual oversight constrains rent extraction across different infrastructure domains that directly affect firm entry.

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

[33] 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.
[34] We define a village as having reliable power supply if it receives commercial electricity more than 70% of the time.

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.

Bank, W. (2009): *Doing Business in India 2009*. World Bank, Washington, DC.

Bardhan, P. (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.

Bussell, J. (2019): *Clients and constituents: Political responsiveness in patronage democracies*. Oxford University Press.

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.

Campos, N. F., S. Estrin, and E. Proto (2010): "Corruption as a barrier to entry: Theory and evidence," Discussion paper, IZA Discussion Papers.

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.

Central Statistics Organisation, Ministry of Statistics and Programme Implementation, Government of India (2013): "Economic Census of India," Accessed via SHRUG v2.1 (Pakora), https://www.devdatalab.org/shrug_download/.

Chandra, K. (2007): *Why ethnic parties succeed: Patronage and ethnic head counts in India*. Cambridge University Press.

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

Chopra, V. K. (1996): *Marginal Players in Marginal Assemblies: Indian M.L.A.* Sangam Books Ltd, London.

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.

Crabtree, J. (2018): *The Billionaire Raj: A Journey Through India's New Gilded Age*. Tim Duggan Books.

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.

Donaldson, D., and A. Storeygard (2016): "The view from above: Applications of satellite data in economics," *Journal of Economic Perspectives*, 30(4), 171–198.

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

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.

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.

Huh, M., B. Cheung, T. Wang, and P. Isola (2024): "The Platonic Representation Hypothesis," *arXiv preprint arXiv:2405.07987*.

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.

Jensenius, F. R., and G. Verniers (2017a): "Studying Indian Politics with Large-Scale Data: Indian Election Data 1961–Today," *Studies in Indian Politics*, 5(2), 269–275.

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

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

Kapur, D., and M. Vaishnav (2013): "Quid pro quo: Builders, politicians, and election finance in India," *Center for Global Development Working Paper*, (276).

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.

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

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

Linke, S., B. Lehner, C. Ouellet Dallaire, J. Ariwi, G. Grill, M. Anand, P. Beames, V. Burchard-Levine, S. Maxwell, H. Moidu, F. Tan, and M. Thieme (2019): "Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution," *Scientific Data*, 6, 283.

List, J. A., A. M. Shaikh, and A. Vayalinkal (2023): "Multiple testing with covariate adjustment in experimental economics," *Journal of Applied Econometrics*, 38(6), 920–939.

List, J. A., A. M. Shaikh, and Y. Xu (2019): "Multiple hypothesis testing in experimental economics," *Experimental Economics*, 22(4), 773–793.

Mahadevan, M. (2024): "The price of power: Costs of political corruption in Indian electricity," *American Economic Review*, 114(10), 3314–3344.

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

Meiyappan, P., P. S. Roy, A. Soliman, T. Li, P. Mondal, S. Wang, and A. K. Jain (2018): "Documentation for the India Village-Level Geospatial Socio-Economic Data Set: 1991, 2001," .

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.

Munshi, K. (2019): "Caste and the Indian economy," *Journal of Economic Literature*, 57(4), 781–834.

Office of the Registrar General and Census Commissioner, India (2011): "Census of India," Accessed via SHRUG https://www.devdatalab.org/shrug_download/, v2.1.pakora.

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.

Singh, K. (1996): *Communities, Segments, Synonyms, Surnames and Titles, People of India Volume 8*. Oxford University Press, digitized by Cassan, Keniston and Kleineberg, A Division of Laborers: Identity and Efficiency in India, NBER Working Paper 28462, 2024.

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.

Sukhtankar, S., and M. Vaishnav (2015): "Corruption in India: Bridging Research Evidence and Policy Options," *India Policy Forum*, 11, 193–261.

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.

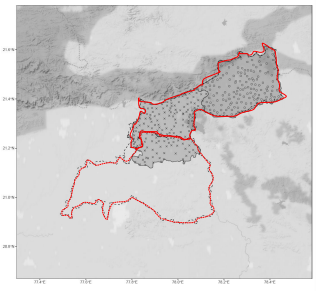
Tullock, G. (1965): "The politics of bureaucracy," *Washington: Public Affairs Press.*

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

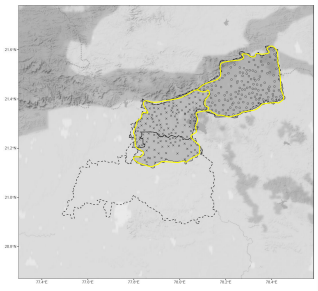
Xu, G. (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*, 108(11), 3170–3198.

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

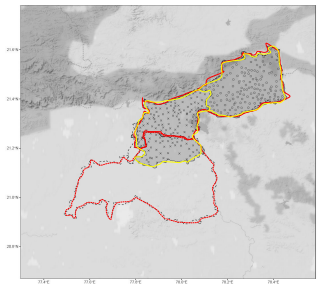
Figure 1: Split and Unsplit Blocks - Overlap of Administrative and Electoral Boundaries



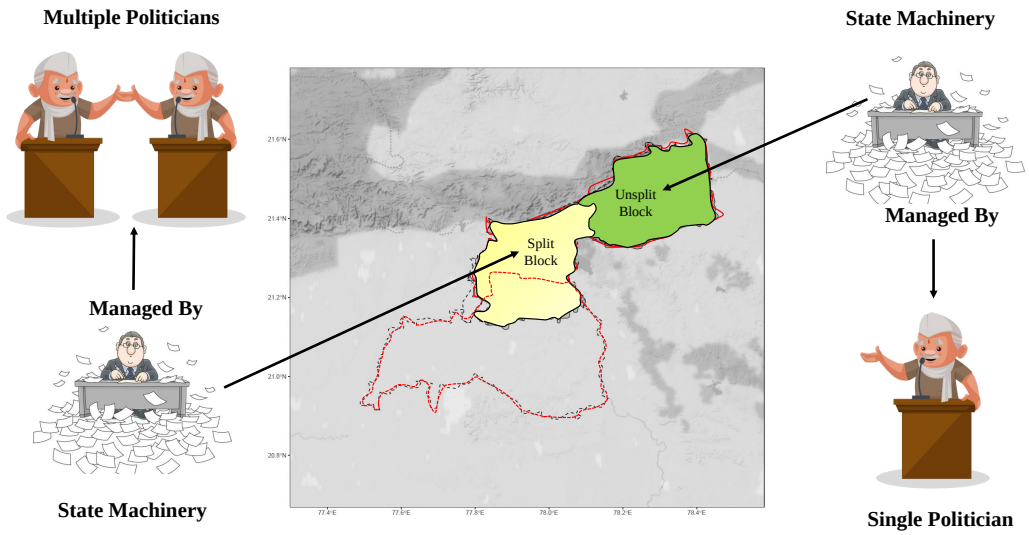
(a) Electoral Boundaries



(b) Administrative Boundaries (Blocks)



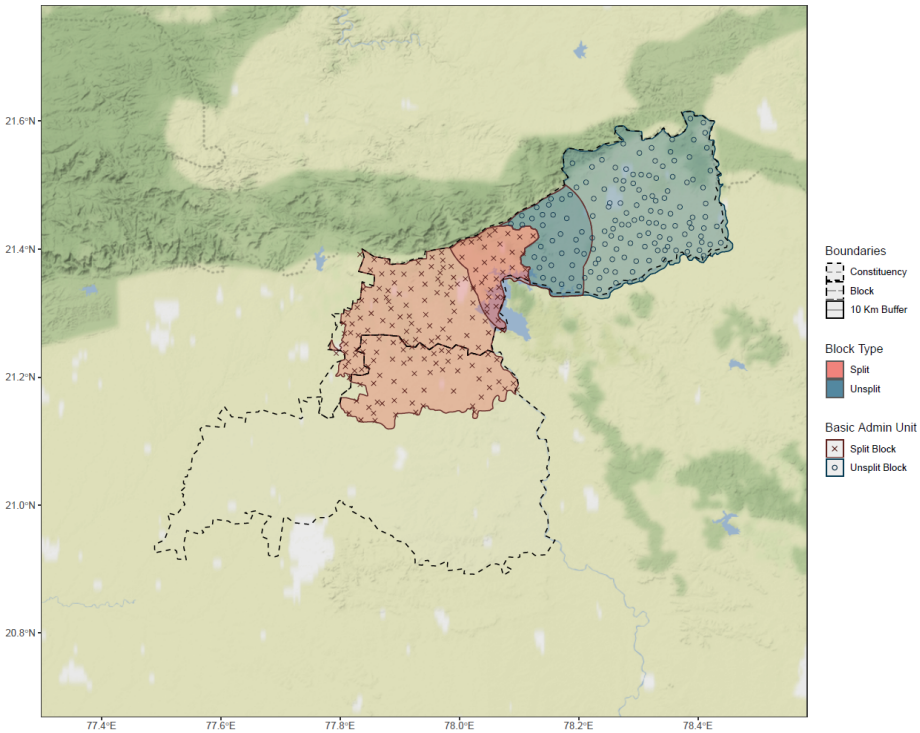
(c) Overlap



(d) Split & Unsplit Blocks

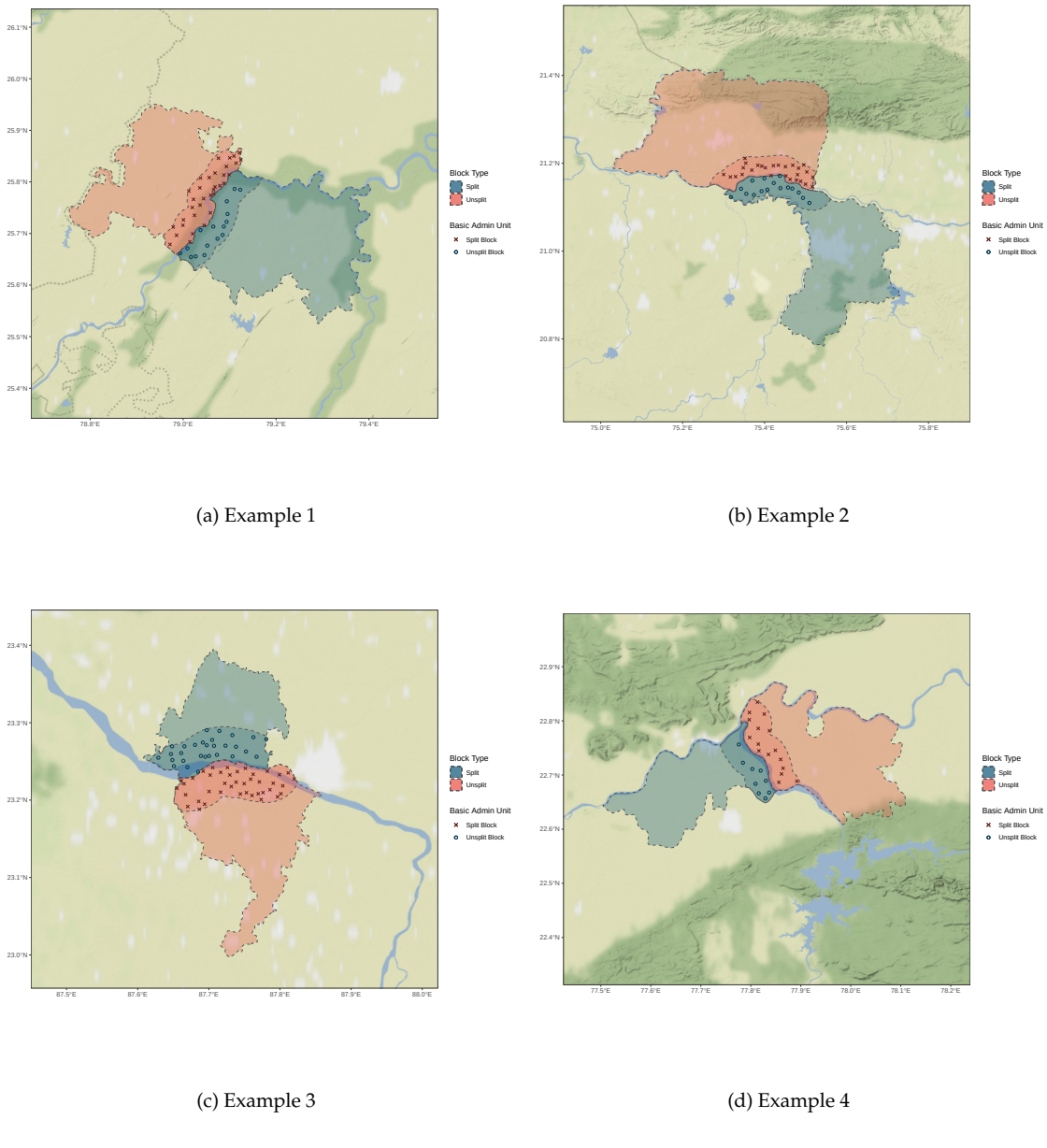
The figure illustrates a case of adjacent split and unsplit blocks in our sample. The split block is the administrative block of Morshi, in the district of Amaravati, Maharashtra. This block is split between the two assembly constituencies of Morshi and Teosa. The neighbouring unsplit block is the administrative 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

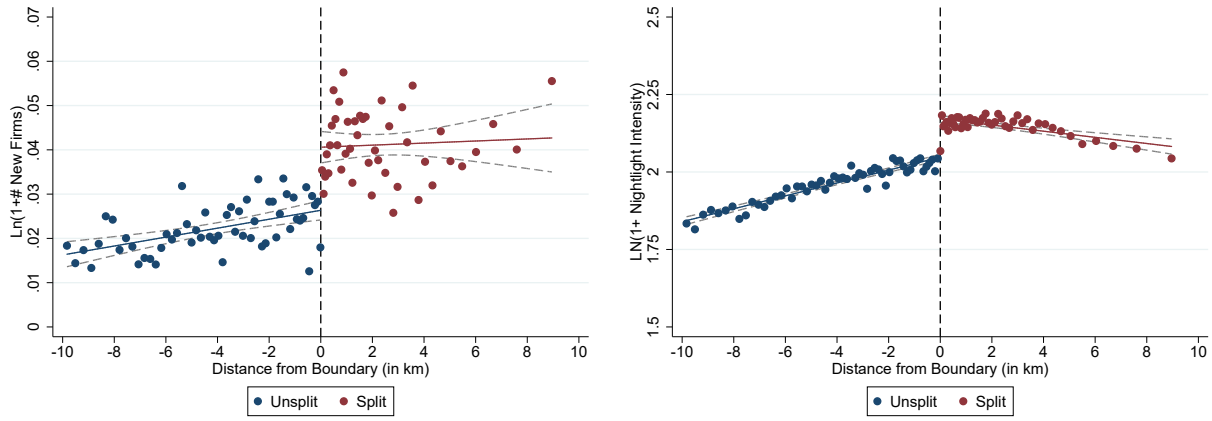
(b) Example 2

(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

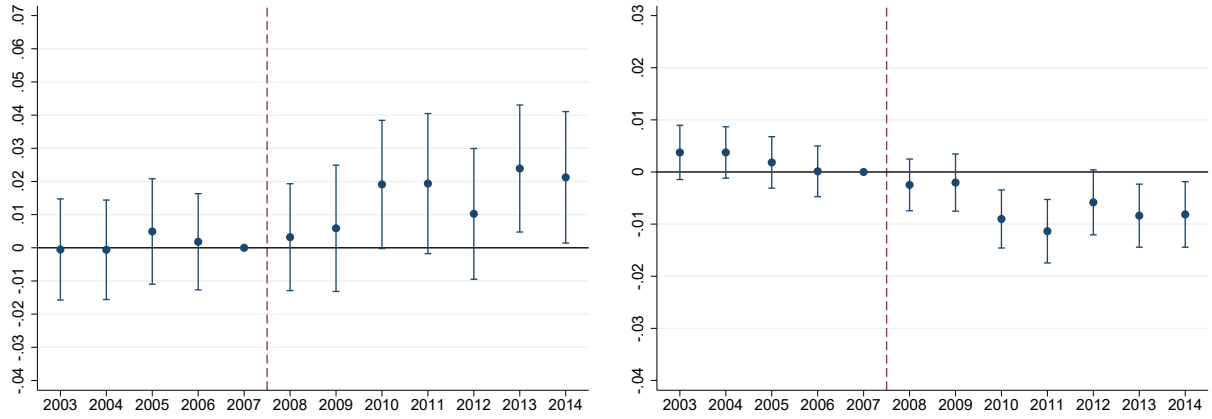


(a) Firm Entry

(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

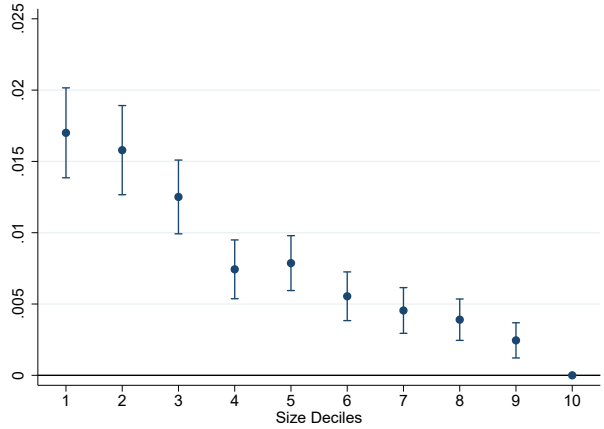


(a) Unsplit → Split

(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

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
Within R^2	0.0001	0.0002	0.0002	0.0002
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 with the interaction term between split blocks and the fraction of non-aligned politicians, using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use bandwidths of 5, 10, 20, and 50 km on either side of the boundary—separating split blocks from unsplit blocks—in columns (1) through (4), respectively. The unit of observation is a village-year within the narrow bandwidth around the boundary. A block is classified as split based on the random overlap of block boundaries with electoral boundaries following the 2008 delimitation. Firm entry is defined as the total number of new firms entering a village in a given year, for the period from 2008 to 2016. Results are reported for an estimation where non-alignment is measured solely based on political party affiliation. The fraction of non-aligned politicians is a continuous variable that takes a value between 0 and 1: it equals 0 when all politicians belong to the same party and 1 when each is affiliated with a different party. In unsplit blocks, this measure is always 0. All regressions include boundary × year fixed effects, 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
Within R^2	0.0001	0.0002	0.0002	0.0002
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. In columns (3) and (4) we run the regression specification using all electoral constituencies. 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. Standard errors reported in parentheses are robust. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Firm Entry by Political Patronage

Dep Var: LN(0.001+# Firms)	(1)	(2)	(3)	(4)
Split × Connected	-0.0131**	-0.0086*	-0.0116***	-0.0108***
	(0.0058)	(0.0049)	(0.0042)	(0.0039)
#Obs	350,706	487,848	584,472	619,866
R^2	0.6846	0.6798	0.6741	0.6736
Within R^2	0.0001	0.0002	0.0002	0.0002
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Connected 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 connected. The dependent variable is the natural logarithm of 0.001 plus firm entry. Firm entry is defined as the total number of firms of each type entering a village in the first political cycle after the delimitation. A block is defined as a split block based on the haphazard overlap of block boundaries with electoral boundaries as per the 2008 delimitation. Firms are classified into two types: (1) Connected firms if they have at least one director sharing the same caste identity as the local politician in power, determined through ethnographic mapping using the Persons of India database, and (2) Non-connected firms if they do not share caste identity with the local politician. Connected is an indicator variable that equals one for connected firms and zero for non-connected firms. The interaction term captures the differential effect of block splitting on connected versus non-connected 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. The unit of observation is the village-firm type that lies within the narrow bandwidth of the boundary separating a split and an unsplit block. The data is structured in long format where each village appears twice: once for connected firms and once for non-connected firms. All regressions include boundary × connected fixed effects, 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 9: Firm Entry across Industries Vulnerable to Cronyism

Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
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
Within R^2	0.0001	0.0002	0.0002	0.0002
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 with the interaction term between Split and Crony. 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 split and unsplit blocks in columns (1), (2), (3), and (4), respectively. The unit of observation is the village-industry within the narrow bandwidth of the boundary. A block is defined as split based on the haphazard overlap of block boundaries with electoral boundaries following 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 effects, village fixed effects, a local linear specification estimated separately on each side of the boundary, and use a triangular kernel. Crony is an indicator for industries vulnerable to cronyism, based on the industry-level cronyism index created by *The Economist* using *Transparency International* methodology. The detailed description can be found in Planet Plutocrat. 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
Within R^2	0.0001	0.0002
Bandwidth	10 KM	10 KM
Boundary x Year FE	Yes	Yes
Block FE	Yes	Yes

This table presents estimates for specification (1) and (2) 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: Firm Entry across Industries with Varying Regulatory Intensity

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
Within R^2	0.0001	0.0002	0.0002	0.0002
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 industries. 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 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 et al. (2019). Standard errors reported in parentheses are clustered at the block level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: 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
Within R^2	0.0001	0.0002	0.0002	0.0002
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
Within R^2	0.0001	0.0002	0.0002	0.0002
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary × Year FE	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes

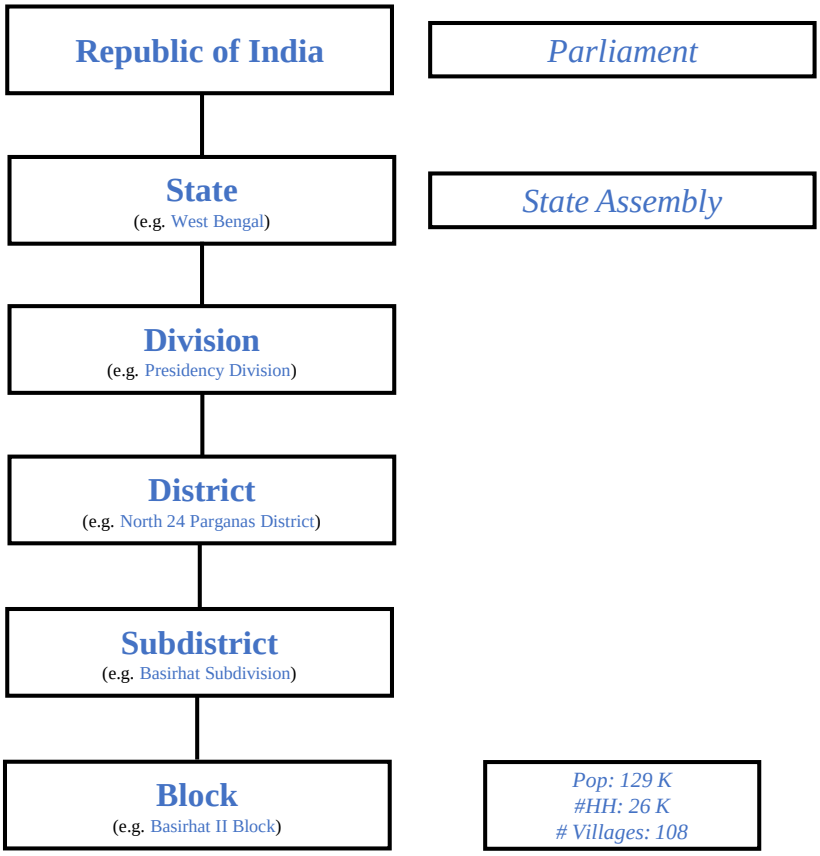
This table presents estimates for specification (1) and (2) using the natural logarithm of actual cost minus the natural logarithm of estimated cost as the dependent variable in panels A and B, respectively. In Panel A the treated group is the set of blocks that switched from being unsplit to split following 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).

A.1 Roles and Responsibilities of the Block Development Officer

This appendix examines the statutory powers of Block Development Officers (BDOs) using Karnataka's governance structure as a case study.[35] Karnataka implements a three-tier local self-governance system through the Karnataka Panchayat Raj Act, 1993: Zilla Panchayat (district level), Taluk Panchayat (block), and Grama Panchayat (village level). The State Appointed Executive Officer (BDO) of the Taluk Panchayat serves as the administrative head with powers that establish a hierarchical relationship affecting governance processes at the village level. The discussion that follows details the BDO's legal mandate, specific oversight mechanisms for Grama Panchayats, and how these powers enable administrative control over local economic and governance activities.

A.1.1 The Block (Taluka) and its BDO: Definitions, Mandate, and General Powers

The Block (Taluk Panchayat) as an Administrative Unit The Taluk Panchayat (block) serves as the intermediate tier in Karnataka's rural local self-government system, functioning as a pivotal unit for area-based planning, implementation of development programs, and administrative coordination for Grama Panchayats (villages) within its boundaries. Its mandate includes consolidating village development plans, executing inter-village or block-wide works, and providing administrative and technical support to village-level bodies. The block thus plays a crucial role in translating district-level policies into actionable programs at the block level.

- **Legal Basis:** The establishment and jurisdiction of the Taluk Panchayat are detailed in Section 119 of the Act. Its functions are enumerated in Section 145 and further elaborated in Schedule II of the Act.

The BDO: Appointment and General Statutory Authority within the Block The BDO serves as the chief executive of the Taluk Panchayat (block), appointed by the state government as a career civil servant. The BDO's general statutory authority encompasses the management of the block's day-to-day administration, its financial resources, and the implementation of various developmental schemes and public works undertaken by the Taluk Panchayat.

- **Legal Basis for Appointment & Status:** Section 155(1) of the Act mandates that "The Government shall appoint a Group-A Officer of the State Civil Services or equal to the rank of the Assistant Commissioner to be the Executive Officer of the Taluk Panchayat."
- **Legal Basis for General Powers:** Section 156(1) of the Act outlines the primary functions of the BDO (Executive Officer), which include:
  - "exercise all the powers specifically imposed or conferred upon him by or under this Act, or under any other law for the time being in force" (Section 156(1)(a));

[35] For terminological consistency, we refer to the Executive Officer of the Taluk Panchayat as the Block Development Officer (BDO) and the Taluk Panchayat as the block throughout this paper.

- "control the officers and officials of, or holding office under the Taluk Panchayat subject to the general superintendence and control of the Adhyaksha and such rules as may be prescribed" (Section 156(1)(b));
- "supervise and control the execution of all works of the Taluk Panchayat" (Section 156(1)(c));
- "take necessary measures for the speedy execution of all works and developmental schemes of the Taluk Panchayat" (Section 156(1)(d)). The term "necessary measures" inherently implies a degree of discretion in operational decisions.
- "have custody of all papers and documents connected with the proceedings of the meetings of the Taluk Panchayat and of its Committees" (Section 156(1)(e));
- "draw and disburse monies out of the Taluk Panchayat fund" (Section 156(1)(f)).

A.1.2 The BDO's Authority and Influence over Grama Panchayats

The Taluk Panchayat's (block) operational mandate includes oversight of all Grama Panchayats (village institutions) within its jurisdiction, making this a core BDO responsibility. As the Taluk Panchayat's executive arm, the BDO connects block and village governance through statutory powers that provide discretionary authority over the functioning and resource management of Grama Panchayats. These powers enable the BDO to shape governance outcomes at the village level, with mechanisms to ensure village compliance with block-level policy preferences. The statutory provisions conferring these jurisdictional powers are examined below.

Financial and Budgetary Control The BDO exercises control over Grama Panchayat finances through budget approval requirements and audit authority, creating dependence of village institutions on block-level authorization.

- **Legal Basis (Budget):** Section 241(2-5) mandates Grama Panchayat budget submission to the Taluk Panchayat, which holds modification rights to "secure compliance." If a village fails to pass its budget, the Taluk Panchayat "shall approve it with or without modification," with the BDO certifying the final document. This creates a direct approval requirement whereby village-level spending plans cannot proceed without BDO certification.
- **Legal Basis (Audit):** Section 246 grants the BDO post-expenditure control, including powers to "direct that the defects or irregularities... shall be removed or remedied" (Sec 246(7)) and to "disallow any item of expenditure" and "surcharge the amount thereof on the person making or authorising the illegal payment" (Sec 246(8)). These provisions establish personal financial liability for Grama Panchayat officials who make expenditures later rejected by the BDO—creating incentives for compliance with BDO preferences.

Mandatory Compliance and BDO Inspection Directives The BDO holds authority to conduct inspections of Grama Panchayat operations and issue binding directives that village institutions must implement.

- **Legal Basis:** Section 232(a) explicitly states the BDO may "inspect the offices or premises or works taken up by the... Grama Panchayat" and that "the... Grama Panchayat concerned shall comply with the instructions issued after such inspections." This mandatory compliance requirement transforms routine inspections into a directive mechanism through which the BDO can direct operational changes at the village level without requiring further authorization.

Non-Compliance Response and Takeover Authority The Act establishes intervention mechanisms for addressing deficiencies in Grama Panchayat performance, with the BDO serving as the implementing authority for Taluk Panchayat directives.

- **Legal Basis:** Section 235(1) authorizes the Taluk Panchayat to determine when a Grama Panchayat has "made default in performing any duty" and to "fix a period for the performance of that duty." Upon continued non-compliance, the Taluk Panchayat may "appoint a person to perform it" and "direct that the expense... shall be forthwith paid by the Grama Panchayat." As the executive officer responsible for executing Taluk Panchayat resolutions under Section 156(1), the BDO operationalizes these interventions, effectively serving as the administrative channel through which higher-level directives supplant village-level authority.

Information Access and Document Requisition Powers The Act grants the BDO powers to requisition, seize, and investigate Grama Panchayat records and finances.

- **Legal Basis:** Section 157(1-3) empowers the BDO to requisition "moneys or... accounts, records or other property" from Grama Panchayats, with authority to "issue a search warrant and exercise all such powers... as may lawfully be exercised by a Magistrate." Section 157(3) creates a legal obligation for any person with knowledge of concealed village resources to report this to the BDO. These powers extend the BDO's authority beyond administrative oversight into investigative domains.
- **Legal Basis (Routine Information):** Section 232(b) further authorizes the BDO to "call for any return, statement, account or report" from Grama Panchayats, establishing documentation requirements that facilitate monitoring.

Personnel Control Through Appellate Authority The BDO's authority extends to personnel matters, establishing administrative judgment over disciplinary actions against Grama Panchayat employees.

- **Legal Basis:** Section 113(4) establishes that appeals against disciplinary orders issued by the Panchayat Development Officer of a Grama Panchayat "shall lie... to the Executive Officer [BDO]." This appellate authority creates a supervisory relationship between the BDO and village-level staff, affecting the authority of elected Grama Panchayat officials in personnel matters.

A.2 Business Permits and Licensing Authority by the Local Administration

This appendix examines the business permit and licensing framework at the local level, using Karnataka's governance structure to illustrate how local regulatory authority operates in practice. We focus on the comprehensive licensing powers formally vested in Grama Panchayats under the Karnataka Gram Swaraj and Panchayat Raj Act, 1993, which serve as foundational requirements for business establishment and operation. The Act grants Grama Panchayats substantial discretionary authority in permit decisions, including power to impose conditions specified in such permission" and to refuse to grant the permission... for reasons to be recorded in writing."[36] This discretionary space in legal interpretation allows local authorities to influence the cost and feasibility of business operations. Grama Panchayats also possess enforcement powers to penalize non-compliance, issue violation notices, and determine penalty levels.[37] These local permits frequently serve as prerequisites for state-level clearances, creating cascading dependencies where denial or delay at the village level can obstruct multiple higher-level approvals required for business operation. While these decisions occur under Grama Panchayat authority, they operate within a broader administrative framework where Block Development Officers exercise supervisory oversight through multiple mechanisms, including appellate authority, budget approval, and compliance monitoring.[38] The statutory provisions detailed below demonstrate the foundational role of local permits in business formation and the mechanisms through which administrative oversight can influence firm entry outcomes.

Mandatory Business Permits Issued by Grama Panchayats

Trade Licenses (Section 69)

- **Legal Mandate**: Section 69 mandates that "No place within the jurisdiction of a Grama Panchayat shall be used as a shop whether permanently or temporarily... except under a licence granted or renewed by the Grama Panchayat."
- **Coverage**: All commercial establishments, retail shops, service centers, and trading activities within GP jurisdiction.
- **Downstream Requirements**: Bank Accounts, KSPCB Hazardous Waste Authorization, Municipal Solid Waste Authorization.

Factory and Industrial Establishment Permissions (Section 66)

[36] Sections 66, 67, 68, 69 for conditions specification; Section 70(2) for refusal authority, Karnataka Panchayat Raj Act, 1993.

[37] Sections 64(3), 70(2F), and 292 provide enforcement authority and discretionary compliance standards, Karnataka Panchayat Raj Act, 1993.

[38] As detailed in the previous appendix, BDOs possess inspection powers (Section 232), audit authority (Section 246), and serve as final appellate authority for permit decisions (Sections 64(5) and 70(3), Karnataka Panchayat Raj Act, 1993).

- **Legal Mandate**: Section 66 requires that "No person shall, without the permission of the Grama Panchayat... construct or establish any factory, workshop or workplace... to employ steam power, water power or other mechanical power or electrical power, or... install... any machinery or manufacturing plant driven by any power."
- **Coverage**: All mechanized production facilities, requiring permission before factory construction or machinery installation.
- **Downstream Requirements**: Factory Plan Approval, Factory License.

Building Construction Permits (Section 64)

- **Legal Mandate**: Section 64(1) states "no person shall erect any building or alter or add to any existing building... or reconstruct any building without the written permission of the Grama Panchayat."
- **Coverage**: All new construction, alterations, additions, and reconstructions within GP limits, including all business premises, factories, and commercial buildings.
- **Downstream Requirements**: Occupancy Certificate, Commencement Certificate.

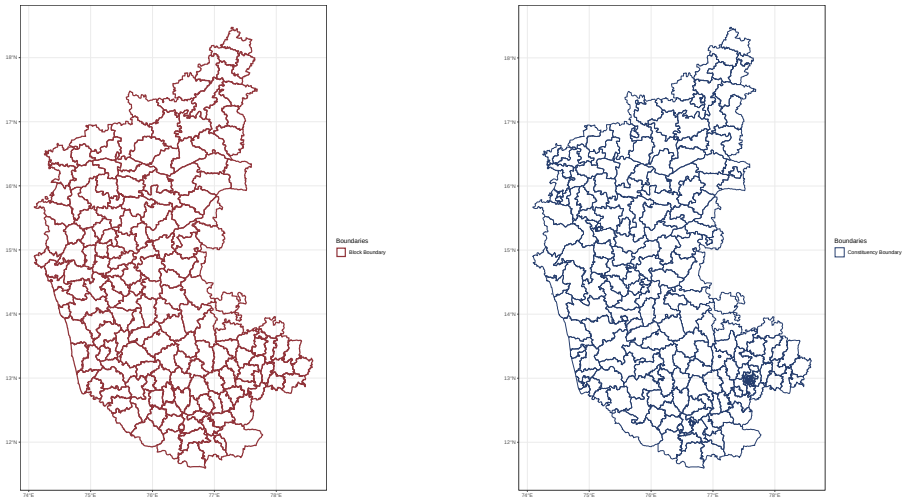
Specialized Business Activity Licenses (Section 67)

- **Legal Mandate**: Section 67 mandates that "No place within the jurisdiction of Grama Panchayat shall be used for the purpose of any trade, business or industry which the Government may, by notification declare to be offensive or dangerous, except under a licence granted or renewed by the Grama Panchayat."
- **Coverage**: Government-notified hazardous, offensive, or dangerous trades and industries.
- **Downstream Requirements**: Environmental clearances, KSPCB authorizations.

General Business NOCs (Section 60)

- **Legal Mandate**: Section 60 grants that "Grama Panchayat shall have powers to do all acts necessary for or incidental to the carrying out of the functions entrusted, assigned or delegated to it."
- **Coverage**: No Objection Certificates and other supporting documentation required by state-level regulatory authorities as prerequisites for higher-level licenses.
- **Downstream Requirements**: Bank Accounts, Land conversion Approvals, Factory Plan Approval, Electricity Connection.

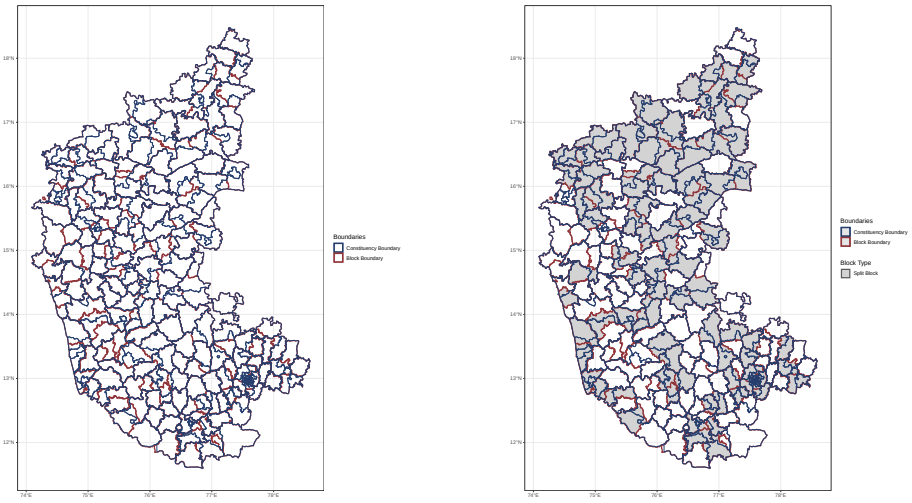
Figure A.2: Administrative and Electoral Boundaries



(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.

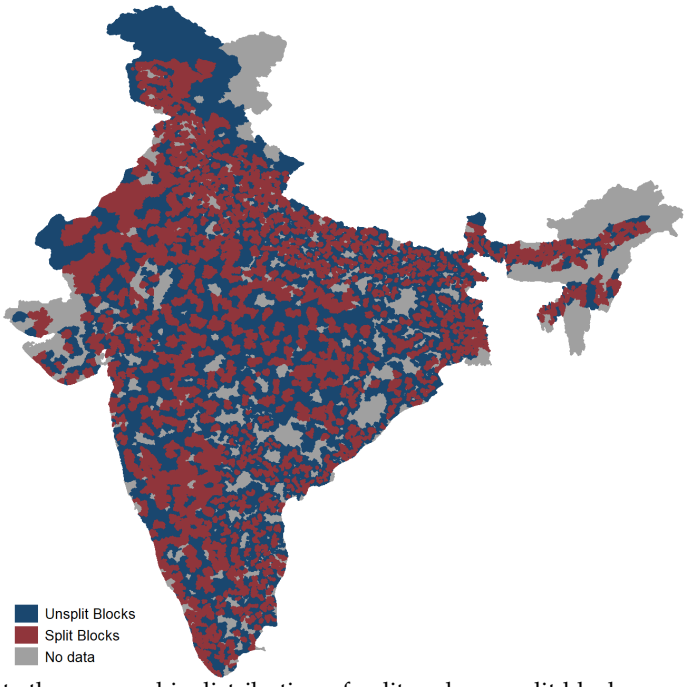
Figure A.3: Overlap of Administrative and Electoral Boundaries



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

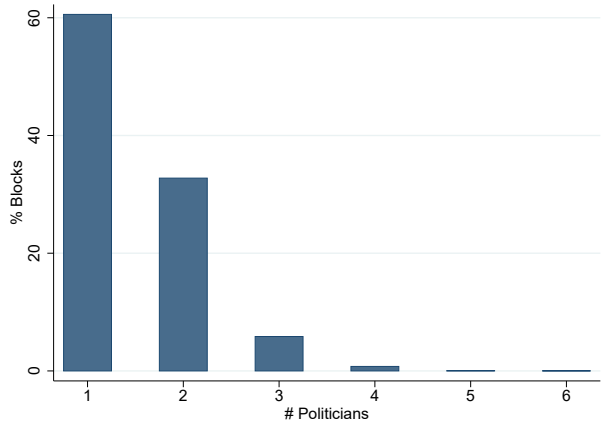
The figure presents the overlap of the administrative and electoral boundaries, as an example, from the state of Karnataka and the formation of split and unsplit blocks. The administrative boundaries in red are based on the basic administrative unit - block. The electoral boundaries shown in blue are 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 Data

B.1 Firm Geolocation Methodology

This subsection 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.1.1 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.[39]
- **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.

[39] 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[40]
    - 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

[40]Our estimates are robust to this choice, and we further test the robustness of this selection method in subsection B.1.3

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.1.2 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.1.3 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

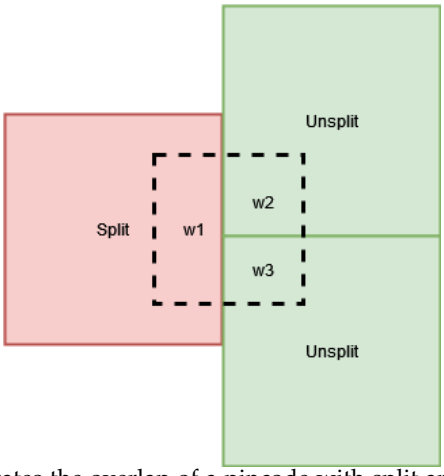
overlap and measure the fraction of the area occupied by split blocks for each pincode using the following expression:

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

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.

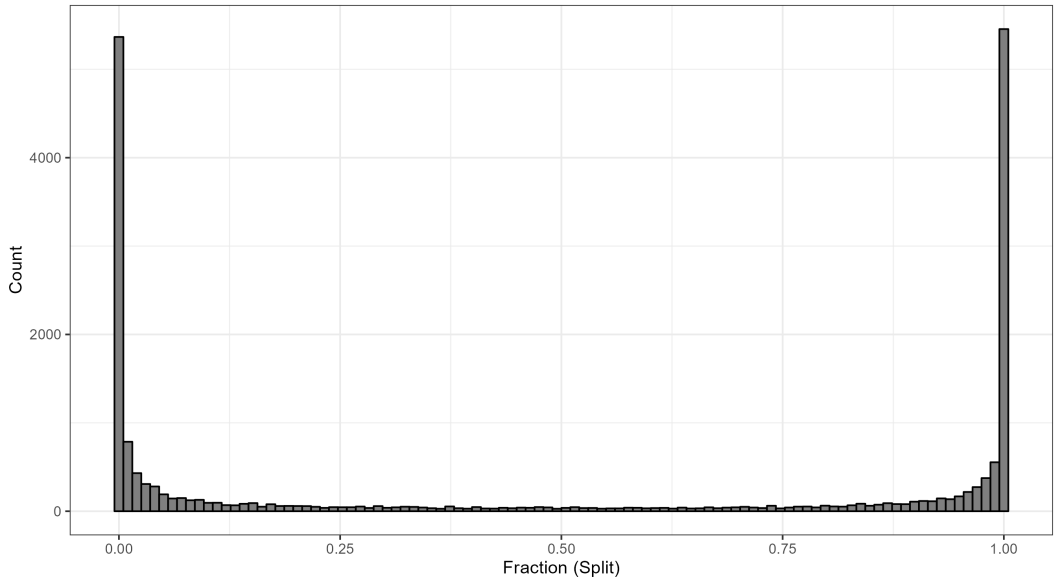
Figure B.1 provides a visual representation of this calculation. The figure shows how a pincode overlaps with both split and unsplit blocks. The shaded squares represent different block types: the red square indicates a split block, while the green squares represent unsplit blocks. The dashed outline 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) > 0.9				
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

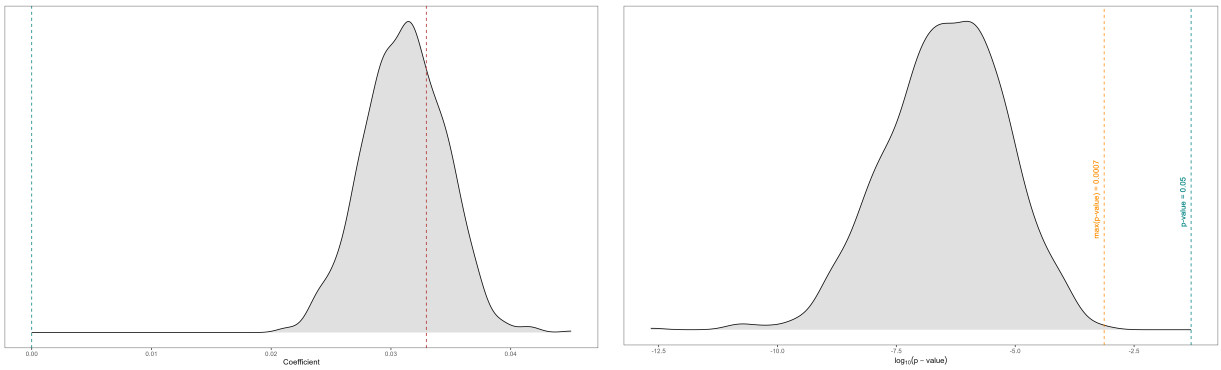
Panel C: Pincodes with Fraction(Split) < 0.01 or Fraction(Split) > 0.99				
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
g10_split	0.1054***	0.1021***	0.0966***	0.0929***
	(0.0240)	(0.0197)	(0.0178)	(0.0169)
R^2	0.28650	0.23506	0.19889	0.17754
Observations	76,044	122,538	165,610	183,343
Bandwidth	5 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



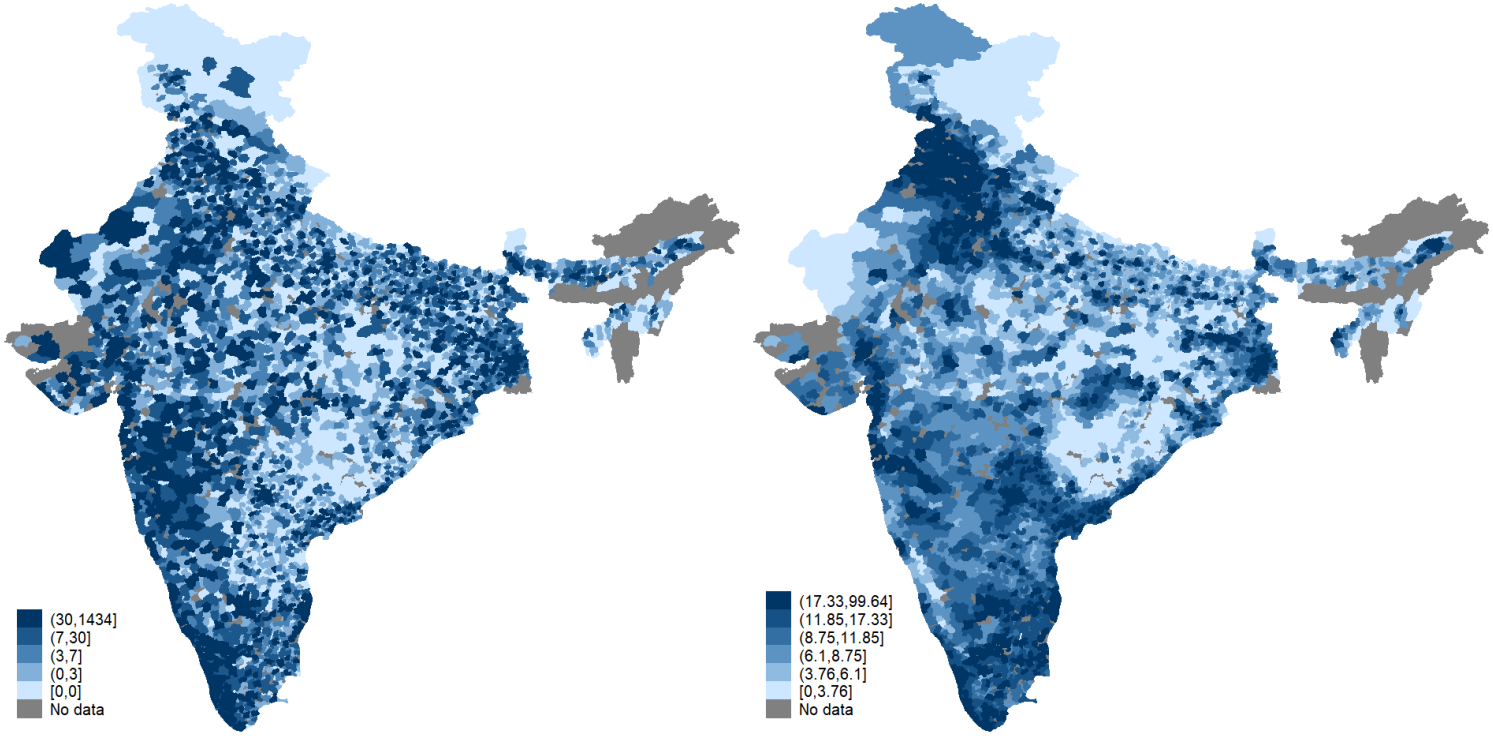
(a) Distribution of coefficient estimates

(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.

Figure B.4: Geography of Firm Entry and Nightlights

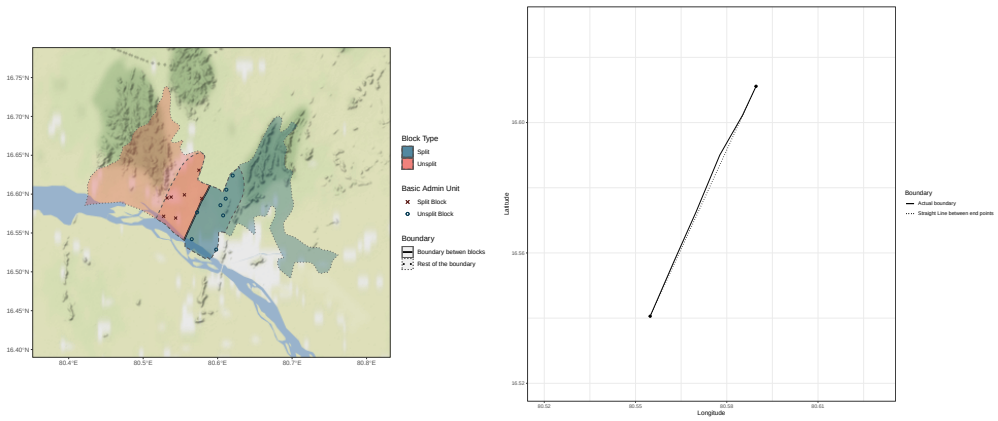


(a) # New Private Firms

(b) Average Nightlights

The figure panel B.4a plots the geographic distribution of the number of new private firms registered across all blocks in India from 2003 to 2016. The figure panel B.4b 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 B.5: Sinuosity of Straight-Line-Like Block Boundary Segment

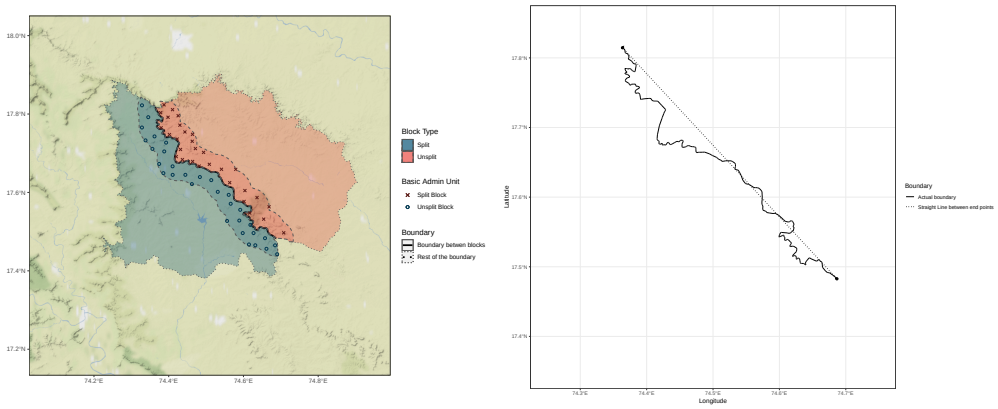


(b) Schematic Diagram

(a) Representative Example

Figure B.5 illustrates the measure of straightness, *Sinuosity*, using a representative example of straight-line-like boundary. Figure B.5a illustrates the boundary between the blocks, Ibrahimpatnam and Vijayawada (Rural), located in Krishna District, Andhra Pradesh. Figure B.5b 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 B.6: Sinuosity of a Wiggly Block Boundary Segment



(b) Schematic Diagram

(a) Representative Example

Figure B.6 illustrates the measure of straightness, *Sinuosity*, using a representative example of wiggly boundary segment. Figure B.6a illustrates the boundary between the blocks, Khatav and Phaltan located in Satara District, Maharashtra. Figure B.6b 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 B.6: 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.

B.2 Caste-Based Patronage Detection

B.2.1 Overview

This appendix outlines our methodology for identifying firms potentially connected to politicians through shared identity. In India, patronage networks often operate along lines of jati — endogamous groups that shape social organization. Indian surnames frequently carry markers of jati affiliation, providing a basis for detecting such network connections. However, inferring jati from surnames faces substantial challenges arising from heterogeneity in naming conventions and cultural nuances. We address these challenges by implementing a computational pipeline combining machine learning techniques and rich ethnographic documentation. We attempt to go beyond simple surname matching to capture social affiliation through which patronage relationships might realistically operate.

B.2.2 Data and Sample

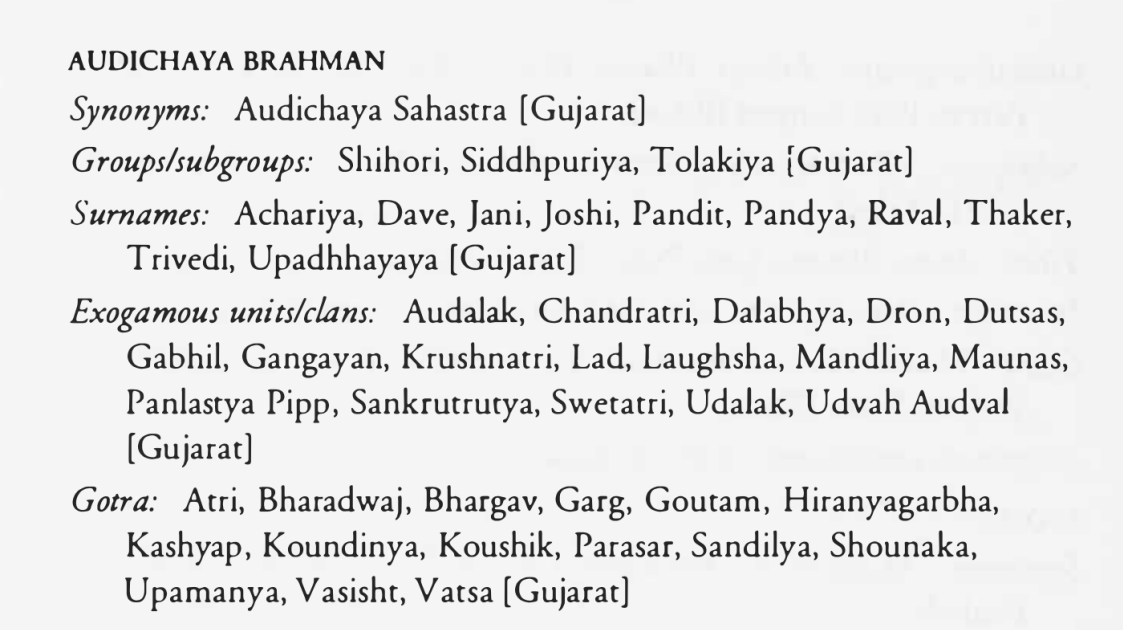
Data Sources :

Administrative Data. The names of politicians and their electoral constituency assignments come from the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG) database (Asher, Lunt, Matsuura, and Novosad, 2021; Jensenius and Verniers, 2017b). We obtain the names of firm directors from the Company Register maintained by the Ministry of Corporate Affairs, India. We match firms to politicians based on geographic location, ensuring that our analysis examines potential patronage connections only within the constituencies that politicians represent.

People of India Project Database. The People of India Project (Singh, 1996) provides ethnographic documentation of Indian communities. Conducted state by state, this project systematically catalogued *jatis*, documenting for each community its associated surnames, regional synonyms, hierarchical sub-groups, exogamous marriage circles, and patrilineal descent groups (gotras). A representative entry for the community (*jati*) "Audichaya Brahman", belonging to the state of Gujarat, is illustrated in Figure B.7.

The relationship between surnames and *jatis* exhibits systematic many-to-many correspondence rather than one-to-one mapping. The methodology followed in the People of India Project documented surnames as communities actually use them, capturing substantial variation within single *jatis* and overlap across related groups. This pattern reflects historical processes—closely related *jatis* often share surnames due to fission from common ancestor communities. Geographic dispersion also produces different surname usage for the same endogamous group, as regional conventions and migration patterns create surname diversity within caste boundaries. Identical surnames may therefore indicate membership in multiple related sub-castes within a given region, while individuals sharing caste identity may carry different surnames.

Figure B.7: People of India Project Database Structure



AUDICHAYA BRAHMAN

Synonyms: Audichaya Sahastra [Gujarat]

Groups/subgroups: Shihori, Siddhpuriya, Tolakiya [Gujarat]

Surnames: Achariya, Dave, Jani, Joshi, Pandit, Pandya, Raval, Thaker, Trivedi, Upadhhayaya [Gujarat]

Exogamous units/clans: Audalak, Chandratri, Dalabhya, Dron, Dutsas, Gabhil, Gangayan, Krushnatri, Lad, Laughsha, Mandliya, Maunas, Panlastya Pipp, Sankrutrutya, Swetatri, Udalak, Udvah Audval [Gujarat]

Gotra: Atri, Bharadwaj, Bhargav, Garg, Goutam, Hiranyagarbha, Kashyap, Koundinya, Koushik, Parasar, Sandilya, Shounaka, Upamanya, Vasisht, Vatsa [Gujarat]

The figure displays a representative entry from the People of India Project database for the community (*jati*) "Audichaya Brahman" from the state of Gujarat. The entry illustrates the hierarchical structure of ethnographic documentation: *Synonyms* captures alternative community names used regionally; *Groups/subgroups* documents related sub-castes within broader classificatory hierarchies; *Surnames* lists family names associated with the community; *Exogamous units/clans* identifies marriage circles that define permissible unions; and *Gotra* records patrilineal descent groups. The database documents 2,205 such *jatis* across Indian states, with each entry containing multiple surnames and related communities, creating the many-to-many mapping between surnames and *jati*

Sample Construction :

Cultural nuances in Indian naming conventions restrict our analysis to contexts where surname-based *jati* identification operates most reliably.

Religious Restriction. Surnames carry caste markers primarily for Hindus, where *jati* organization functions most systematically within social structure. We therefore restrict attention to Hindu politicians and firm directors.

Geographic Restriction. North Indian naming practices emphasize hereditary surnames that transmit consistently across generations and carry *jati* affiliation. South Indian conventions prioritize patronymics (father's given name) and village references that change generationally, making surname-based caste identification unreliable. Hence we limit the analysis to eleven North Indian states: Himachal Pradesh, Haryana, Rajasthan, Uttar Pradesh, Bihar, Jharkhand, Orissa, Chhattisgarh, Madhya Pradesh, Gujarat, and Maharashtra.

Surname-Based Restriction. Some surnames (such as Kumar and Singh) function as generic honorifics or titles that diffused across multiple communities, weakening their reliability as caste markers. We exclude politicians with such caste-neutral surnames from the sample.

These restrictions yield a final sample of 915 Hindu politicians from an initial universe of 1,805, matched to 86,767 unique firms operating in their constituencies. This sample provides complete coverage of contexts where our methodology enables reliable patronage detection, avoiding false connections that might arise from uncertain caste assignments or unreliable naming conventions.

B.2.3 Technical Implementation

Overview In this subsection we detail the methodology used to process politician and director names, following a sequential pipeline to address technical challenges that commonly arise when handling Indian names. Stage 1 classifies religion to identify individuals with Hindu names, for whom *jati* represents a salient identity marker. Stage 2 extracts surnames from complex naming conventions that vary substantially across Indian states. Stage 3 handles transliteration differences and non-standard spellings through embedding-based fuzzy matching with the People of India database, allowing us to infer the *jati* of the politician from their surname despite orthographic variations. Stage 4 implements expanded caste mapping by first building surname inventories for each politician's *jati*, then matching directors against these expanded surname lists to identify potential identity-based connections.

Stage 1: Religion Classification We build a classifier to identify individuals with Hindu names for whom surname-based *jati* inference operates systematically. The classifier

employs a two-stage approach balancing computational efficiency with accuracy. Initial screening uses character-level Long Short-Term Memory neural networks trained on proprietary matrimonial datasets containing 1.3 million entries on persons' names and religion.[41] We set a high classification standard where only names receiving high-confidence predictions above 0.9 probability proceed automatically. Names with prediction probabilities below this threshold require more sophisticated analysis and are routed to an ensemble of large language models comprising *Qwen 2.5*, *GPT-4o Mini*, and *Gemini 2.0*. These models analyze names using cultural and linguistic knowledge beyond character patterns, with structured output schemas preventing classification errors.[42]

Table B.7 reports classification performance metrics on holdout validation data. High precision (96.3%) ensures that downstream caste inference operates on reliably identified Hindu names, while high recall (98.2%) minimizes systematic exclusion of Hindu politicians from the analysis sample. This filtering produces a subsample of Hindu politicians where subsequent surname extraction and caste mapping achieve maximum reliability.

Table B.7: Religion Classification Performance

(a) Confusion Matrix

	Actual Religion	
Predicted	Hindu	Non-Hindu
Hindu	751	29
Non-Hindu	14	206
Total	765	235

(b) Performance Metrics

Metric	**Value**
Accuracy	95.7%
Precision	96.3%
Recall	98.2%
F1-Score	97.2%
Sample Size	1,000

Panel A displays the confusion matrix for binary religion classification where Hindu identification serves as the positive class. Cell entries represent counts of names classified in each category. **Panel B** reports key performance metrics validated on a holdout sample of 1,000 names from matrimonial data that the LSTM models have not seen during training, ensuring out-of-sample performance assessment. *Accuracy* measures the proportion of correctly classified names across both categories. *Precision* indicates the percentage of names predicted as Hindu that are actually Hindu (type I error control). *Recall* captures the percentage of actual Hindu names correctly identified (type II error control). *F1-Score* provides the harmonic mean of precision and recall, balancing both error types.

Stage 2: Surname Extraction We extract surnames from complex Indian names that exhibit substantial variation across regional naming conventions. Different states employ distinct naming patterns—some emphasizing hereditary surnames, others incorporating village names, patronymic elements, or clan affiliations in varying positions within names. Names such as "Rajendra Pratap Singh alias Moti" incorporate aliases that may override formal surname conventions, while "Jadeja Kandhalbhai Sarmanbhai" reflects

[41] We implement three specialized binary classifiers—one each for Hindu, Muslim, and Christian identification—rather than a single multi-class model to capture religion-specific linguistic patterns more effectively. Each model processes character sequences through bidirectional LSTM layers with dropout regularization.

[42] Final classifications employ majority voting across the ensemble, with validation on out-of-sample matrimonial data ensuring reported performance metrics reflect genuine generalization rather than overfitting.

state-specific (here, Gujarat) patterns that differ markedly from conventions in other states. Additional complexity arises from honorific titles, professional designations, and administrative variations that obscure genealogically relevant components. Naive extraction of the last name component therefore fails to capture caste-informative elements consistently across India's diverse naming landscape.

To address this challenge, we leverage large language models that acquire knowledge of cross-cultural naming conventions during pretraining on diverse text corpora, enabling them to distinguish surnames from other name components across different cultural contexts. We provide each model with the politician's home state as contextual information to guide culturally appropriate surname extraction. Our ensemble comprises five models—*Mistral Nemo*, *Grok 3 Mini*, *Gemini 2.0*, *Qwen 2.5*, and *GPT-4o Mini* — pretrained on large-scale datasets [43]. Models analyze name structure independently, with consensus achieved through majority voting to identify surname components most likely to carry caste information. This approach enables systematic extraction of caste-relevant surname candidates across India's heterogeneous naming conventions.

Stage 3: Fuzzy Surname Matching Extracted surnames must be matched against the People of India database to infer *jati* associations. However, identical surnames exhibit substantial orthographic variation in administrative records. Transliteration from regional scripts to Roman characters produces inconsistent spellings, where the same genealogical identity appears with different spellings due to transcription practices.[44] Exact string matching would systematically miss valid connections between politicians and database entries, undermining inference of *jati*.

We address this challenge through embedding-based fuzzy matching that captures not just semantic but also contextual similarity across spelling variants. The approach converts surnames to high-dimensional vectors where orthographically different but contextually identical surnames cluster in proximity. We then enable similarity-based retrieval rather than exact correspondence. We rely on vector databases through Pinecone, which indexes People of India surnames as embeddings and retrieves contextual candidate names. We then use *LLMs-as-a-Judge* to validate candidate matches by confirming whether spelling variants represent identical surname identities versus merely similar-sounding but distinct names, preventing false connections while capturing genuine orthographic variations.[45]

Stage 4: Expanded Caste Mapping Given a politician's surname, we seek to identify directors who belong to the same community. A surname maps to multiple related *jatis* that share common ancestry, regional location, and social position. For each politician's surname, we construct a comprehensive surname inventory by extracting all surnames associated with these related *jatis*, creating an expanded list that broadly captures names

[43] Huh, Cheung, Wang, and Isola (2024) demonstrate that AI models trained on large corpora develop converging representations across domains, justifying ensemble approaches for cross-cultural tasks.

[44] Common examples include "Chaudhary," "Choudhary," and "Choudhari," or "Agarwal," "Aggarwal," and "Agrawal"—all representing identical surname identities with different romanization conventions.

[45] The LLM-as-a-Judge approach employs large language models to evaluate and validate system outputs against semantic criteria, providing more nuanced assessment than simple statistical metrics by capturing contextual appropriateness and semantic equivalence.

indicating membership across the interconnected subcommunities. This inventory represents coherent networks for social interaction and mutual support rather than arbitrary surname groupings. Using this expanded surname inventory, we flag potential connections by matching director names against any surname in the comprehensive list. This approach recognizes that patronage operates through broader community-based networks that reflect the actual social structures through which caste identity functions in Indian society.

Illustrative Application: Abhishek Mishra, MLA (Sarojini Nagar, Uttar Pradesh) :

We illustrate our methology of inferring of patronage using the example of Abhishek Mishra, an MLA representing the Sarojini Nagar constituency in Uttar Pradesh. Querying the People of India database for "Mishra" within Uttar Pradesh reveals that this surname associates with four distinct but related Brahman subcommunities [46]:

- JUJHAUTIYA BRAHMAN
- SAKALDWIPI BRAHMAN
- SANADHYA BRAHMAN
- SARAYUPARIA BRAHMAN

We extract the complete surname inventory across these four related *jatis* which yields 85 distinct surnames: *Arajariya, Atri, Awasthi, Bharadwaj, Chaube, Dikshit, Dwivedi, Mishra, Pandey, Sharma, Tiwari, Tripathi, Upadhyaya*, among others. Any firm director in Sarojini Nagar constituency carrying any of these 85 surnames would be flagged as a potential patronage connection with Abhishek Mishra. Simple surname matching would identify only directors sharing the "Mishra" surname, systematically missing the broader caste network through which preferential treatment might operate. The expanded approach captures more potential connections while trying to remain grounded with ethnographic boundaries that reflect the genuine community relationship.

[46] External validation from news sources confirms his Brahmin community affiliation. See https://www.news18.com/news/politics/its-bjps-outsider-versus-sps-brahmin-face-in-lucknows-sarojini-nagar-constituency-4729361.html

B.3 Data Sources and Variables

B.3.1 Geographic and Political Boundaries

1. **Block Shapefiles**

   Source: MLInfomap

   Description: Administrative block boundaries as of the 2001 Census of India. Blocks are sub-district administrative units that serve as the primary level for rural development program implementation. The dataset contains polygon geometries for 5,823 blocks across India. We use these boundaries to identify split blocks — blocks governed by multiple MLAs due to overlap with assembly constituency boundaries — and to construct our primary treatment variable.

2. **Assembly Constituency Shapefiles**

   Source: DataMeet

   Description: Electoral boundaries for state assembly constituencies corresponding to the 2008 and 1977 delimitations. Assembly constituencies define the electoral districts for Members of Legislative Assembly (MLAs) in India's state legislatures. The 2008 delimitation covers 4,182 constituencies across Indian states. The 1977 delimitation covers 4,109 constituencies. We overlay these electoral boundaries with block boundaries to identify split blocks and to construct our primary treatment variable.

3. **Village Shapefiles**

   Source: Socioeconomic Data and Applications Center (SEDAC) (Meiyappan, Roy, Soliman, Li, Mondal, Wang, and Jain, 2018)

   Description: Village-level polygon boundaries corresponding to the 2001 Census of India, covering approximately 640,000 villages. We use these boundaries for two purposes: (1) aggregate outcomes for firms and nightlights, and (2) construction of distance-to-boundary measures (running variables) for villages near block boundaries in our regression discontinuity design.

4. **River Segment Data**

   Source: HydroATLAS Global River Database (Linke, Lehner, Ouellet Dallaire, Ariwi, Grill, Anand, Beames, Burchard-Levine, Maxwell, Moidu, Tan, and Thieme, 2019)

   Description: Hydrological data on river courses across India, including river segment geometries and average discharge measurements. We use this data to classify block boundaries delimited by major water bodies — those that align with major rivers (discharge $\geq 200\ m^3/s$) for at least 80% of their length. This classification enables constructing subsamples where boundary placement is plausibly exogenous to local economic and political conditions.

B.3.2 Firm Registration Data

1. **Firm Registration Records**

   Source: Ministry of Corporate Affairs (MCA)

   Description: Statutory records for all companies incorporated under the Companies Act, covering 2003-2016 with approximately 825,980 firms. The data include incorporation dates, operational addresses, and industry classification codes. We aggregate firm entry to the village level, constructing both cross-sectional and time-varying panel measures of entrepreneurial activity.

2. **Director Records**

   Source: Ministry of Corporate Affairs (MCA)

   Description: Director names and Director Identification Numbers (DINs) from company incorporation documents. We use director surnames to infer social identity connections between firms and politicians.

B.3.3 Economic Activity Measures

1. **Nightlight Data**

   Source: Harmonized DMSP and VIIRS (Li, Zhou, Zhao, and Zhao, 2020)

   Description: Harmonized nighttime light intensity data combining Defense Meteorological Satellite Program (DMSP, 1992-2013) and Visible Infrared Imaging Radiometer Suite (VIIRS, 2012-2020) satellite imagery. We compute mean nightlight radiance within each village boundary, normalize the values within each year to a 0-100 scale. The nightlight index serves as a proxy for local economic activity and development.

2. **NREGA Application Data**

   Source: National Rural Employment Guarantee Act (NREGA) Program

   Description: Application-level records from the NREGA employment guarantee program covering seven states (Bihar, West Bengal, Jharkhand, Chhattisgarh, Gujarat, Maharashtra, Karnataka) from 2016-2020. Each record contains individual work applications with days of work demanded for every village. The number of days for which applicants seek work through NREGA provides a measure of demand for unemployment assistance.

3. **CMIE CapEx**

   Source: Centre for Monitoring Indian Economy (CMIE)

   Description: Project-level capital expenditure database tracking industrial and infrastructure projects announced between 2003-2016. The data include project announcement dates, implementation commencement dates, and geographic coordinates. We compute distance to split block boundaries (running variable) using spatial coordinates and construct project delay (days from announcement to commencement) as a proxy for efficiency regulatory processing.

4. **Economic Census**

   Source: Government of India, 6th Economic Census 2013 (Asher, Lunt, Matsuura, and Novosad, 2021; Central Statistics Organisation, Ministry of Statistics and Programme Implementation, Government of India, 2013)

   Description: Census of all non-agricultural establishments in India. The data are available at the village level. We use total non-farm employment as a measure of local economic activity.

B.3.4 Political and Electoral Data

1. **Politician Data**

   Source: Trivedi Center for Political Data (Jensenius and Verniers, 2017a), accessed through SHRUG (Asher, Lunt, Matsuura, and Novosad, 2021)

   Description: Election results and candidate information for state legislative assembly elections across India, including winner names, party affiliations, constituencies, and victory margins. We use party affiliations to construct measures of political non-alignment in split blocks (whether multiple MLAs governing the same block belong to different parties). We also use candidate names to infer social identities and identify patronage connections to entrepreneurs.

2. **Electoral Affidavit Data**

   Source: Association for Democratic Reforms (ADR), accessed through MyNeta.info, scraped data available at https://github.com/Vonter/india-election-affidavits

   Description: Digitized sworn affidavits filed by candidates with the Election Commission of India, covering state legislative assembly elections. The affidavits include self-reported assets and liabilities. We use asset declarations at the start and end of electoral terms to compute percentage changes in politician net wealth as a measure of private returns to office.

B.3.5 Public Goods Provision

1. **PMGSY Road Construction Data**

   Source: Pradhan Mantri Gram Sadak Yojana (PMGSY), accessed through SHRUG (Asher and Novosad, 2020; Asher, Lunt, Matsuura, and Novosad, 2021)

   Description: Administrative data from the rural road construction program covering projects from 2001-2014. The data include estimated costs and actual costs for road construction projects. We construct cost overruns as the log difference between actual and estimated costs to measure efficiency in public goods provision.

2. **Population Census**

   Source: Government of India Census 2001 (Office of the Registrar General and Census Commissioner, India, 2011), accessed through SHRUG (Asher, Lunt, Matsuura, and Novosad, 2021)

Description: Village-level demographic, infrastructure, and geographic data from the 2001 Population Census. Demographic variables include total population, female population, and marginalized group (SC/ST) population. Infrastructure measures include health facilities (hospitals and dispensaries), educational facilities (schools and colleges), and other amenities (banks, cinema halls, auditoriums). Geographic characteristics include land area and forest coverage. We use these variables to assess balance on pre-existing characteristics between split and unsplit blocks.

3. **Swachh Bharat Mission**

   Source: Ministry of Drinking Water and Sanitation

   Description: Village-level disbursement data from the Swachh Bharat Mission sanitation program covering 2015-2018. We use disbursement amounts as a proxy for implementation efficacy in public goods provision.

4. **IHDS Survey**

   Source: India Human Development Survey (IHDS), 2001 and 2011 rounds

   Description: Nationally representative household survey measuring confidence in various institutions including local politicians, police, panchayats, and state government. Respondents rate their confidence on a three-point scale (no confidence, intermediate, high confidence). We aggregate responses to the district level using survey weights to construct district-average measures of institutional confidence.

B.3.6 Auxiliary Data

1. **People of India Database**

   Source: People of India Project (Singh, 1996)

   Description: Ethnographic documentation of Indian communities (jatis) conducted state by state, cataloguing surnames, regional synonyms, hierarchical sub-groups, marriage circles, and patrilineal descent groups (gotras). The database exhibits many-to-many correspondence between surnames and jatis, reflecting historical fission processes and geographic dispersion. We use this database to map director and politician surnames to social identity groups.

2. **Matrimony Website Data**

   Source: Proprietary data from Indian matrimony websites

   Description: User profiles containing names (first and last) and self-reported religion, comprising approximately 2.8 million records. We use this data as training data for LSTM models that learn to predict religious identity from naming patterns. The trained models enable inference of social identity connections between firm directors and politicians.

B.4 Industry Classification

Table B.8: 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 B.9: 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 C Robustness Tests

Table C.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 C.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 C.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 C.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 C.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 C.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 C.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 C.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 C.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 C.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 C.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 C.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

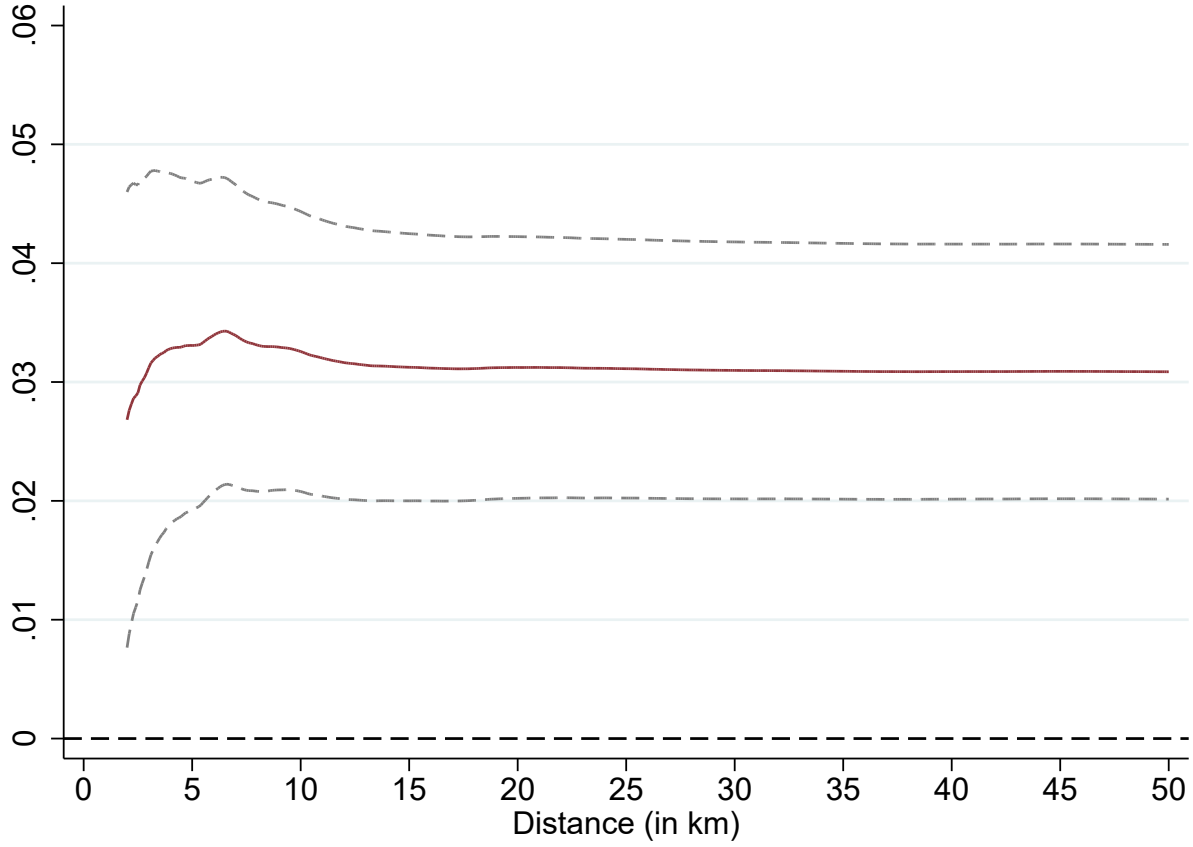
This table presents estimates for specification (1) using the natural logarithm of 0.001 plus firm entry as the dependent variable. We use different bandwidths of 5, 10, 20, and 50 km on either side of the boundary, separating a split block from an unsplit block in columns (1), (2), (3), and (4), respectively. 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 C.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 C.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.

Table C.14: 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.15: P-value Adjusted for Multiple Hypothesis Testing

		(1)	(2)	(3)	(4)	(5)	(6)
			5km			10km	
		Firms	Nightlight	Employment	Firms	Nightlight	Employment
Adjusted (List et al)	p-value	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted (Bonferroni)	p-value	0.001	0.001	0.001	0.001	0.001	0.001
Adjusted (Holm)	p-value	0.001	0.000	0.000	0.001	0.001	0.000

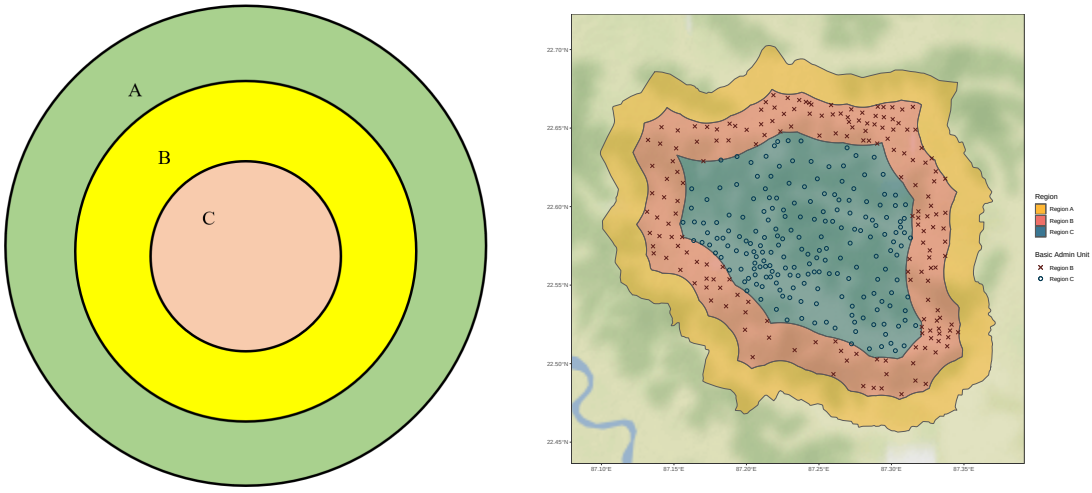
This table explores various adjustments due to the multiplicity of outcomes and reports the p-values for β from the specification (1) using the natural logarithm of 0.001 plus firm entry, nightlight intensity, and employment as the dependent variables. We use different bandwidths of 5 and 10 kms. The columns of the table reports the various ways to adjust standard errors for the multiplicity of the null hypothesis and reports the p-values post the error correction. In First row, we present p-values based on adjustment discussed in List, Shaikh, and Xu (2019) and List, Shaikh, and Vayalinkal (2023). The second row presents p-values adjusted based on Bonferroni. The third row presents p-value adjusted based on Holm. We use code from List, Shaikh, and Vayalinkal (2023) to implement the adjustment.

s

C.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 C.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 C.2: Falsification Test: Sample Construction



(a) Schematic Diagram

(b) Representative Example

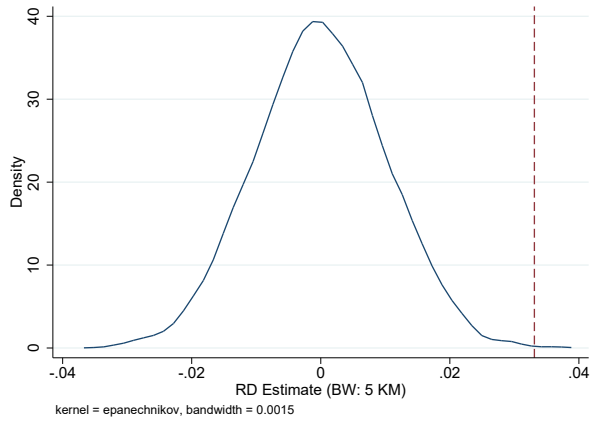
Figure C.2a presents the schematic diagram of a block that is split into three regions A, B and C. These regions are equal in area and account for one-third of the entire block area. Our falsification exercise compares villages in region B with villages in region C within a narrow bandwidth of the boundary separating the two regions. Figure C.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 C.16: Falsification Test: Results

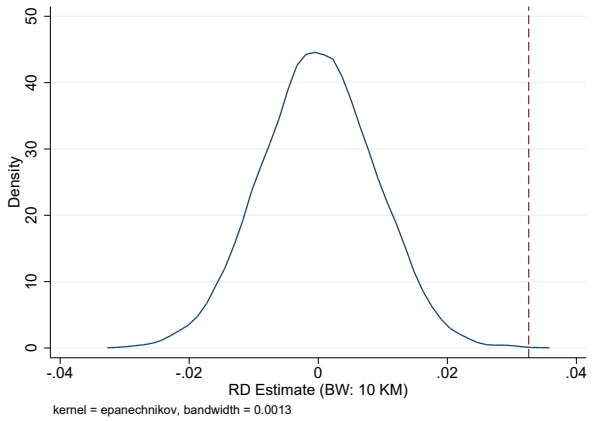
Dep Var: LN(0.001+# New Firms)	(1)	(2)	(3)	(4)
Inside Region C (=1)	-0.0007	-0.0037	-0.0058	-0.0047
	(0.0088)	(0.0067)	(0.0058)	(0.0055)
# Obs	200,843	278,097	320,484	332,161
R^2	0.1587	0.1306	0.1190	0.1157
Bandwidth	5 KM	10 KM	20 KM	50 KM
Boundary FE	Yes	Yes	Yes	Yes

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 C.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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

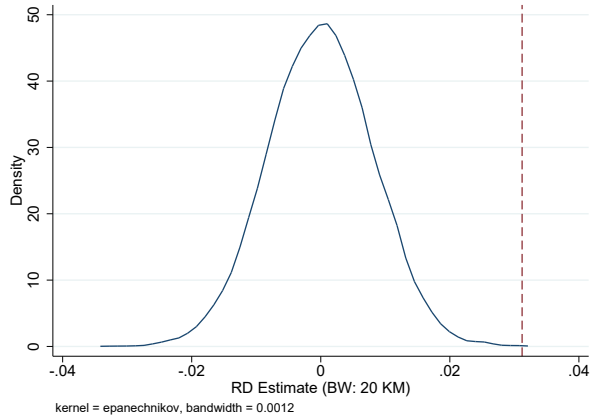
Figure C.3: Placebo Test



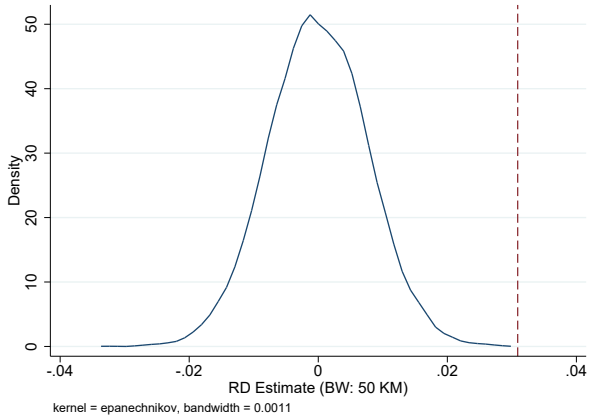
(a) Bandwidth: 5 KM



(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 C.3a), 10 (Figure C.3b), 20 (Figure C.3c), and 50 km (Figure C.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 C.17: 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 D Mechanism

Table D.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 D.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 D.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 D.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 D.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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.6: 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.