Jun 2019 No.426

Pre-Colonial Warfare and Long-Run Development in India

Mark Dincecco, James Fenske, Anil Menon
and Shivaji Mukherjee

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

Pre-Colonial Warfare and Long-Run Development in India*

Mark Dincecco† James Fenske‡ Anil Menon§ Shivaji Mukherjee||

June 14, 2019

Abstract

We analyze the relationship between pre-colonial warfare and long-run development patterns in India. We construct a new, geocoded database of historical conflicts on the Indian subcontinent, from which we compute measures of local exposure to pre-colonial warfare. We document a positive and significant relationship between pre-colonial conflict exposure and local economic development across India today. The main results are robust to numerous checks, including controls for geographic endowments, initial state capacity, colonial-era institutions, ethnic fractionalization, and colonial and post-colonial conflict, and an instrumental variables strategy that exploits variation in pre-colonial conflict exposure driven by the cost distance to the Khyber Pass. Using rich archival and secondary data, we show that early state-making and fiscal development, greater political stability, and basic public goods investments are channels through which pre-colonial warfare has influenced local economic development.

*We thank Pietro Biroli, Traviss Cassidy, Latika Chaudhary, Giorgio Chiovelli, Jeremiah Dittmar, Shashwat Dhar, Bishnupriya Gupta, Mai Hassan, Namrata Kala, Mark Koyama, Alexander Lee, Alexander Persaud, and seminar participants at the Aix-Marseille School of Economics, the London Business School, the National University of Singapore, the New Economic School, the University of Bonn, the University of Gothenburg, UC Louvain, the University of St. Gallen, the American Political Science Association, the Economic History Association, the Midwest Workshop in Empirical Political Science, and the OWL Workshop for helpful comments; Latika Chaudhary, Alexander Lee, and Rinchan Mirza for generous data-sharing; and Justin Huang and Eric Payerle for excellent research help. We gratefully acknowledge financial support from the Department of Political Science at the University of Michigan.

†University of Michigan; dincecco@umich.edu

‡University of Warwick; j.fenske@warwick.ac.uk

§University of Michigan; armenon@umich.edu

||University of Toronto; shivaji.mukherjee@utoronto.ca

1 Introduction

According to a well-known literature, governments undertook institutional reforms in order to enhance their fiscal – and thus military – prowess (Tilly, 1975; Mann, 1984; Brewer, 1989; Besley and Persson, 2011; Gennaioli and Voth, 2015). In time, more powerful government institutions may have helped promote long-run economic development through both greater political stability and the provision of basic public goods. Much of this literature, however, centers on historical state-making in Western Europe. It remains unclear whether this logic applies in other regions.

In this study, we recast the "war makes states" framework in terms of local development patterns across India. We construct a new, geocoded database of historical conflicts on the Indian subcontinent to analyze the pre-colonial military roots of local development outcomes. For hundreds of years prior to European colonial rule, rival states competed for political dominance on the Indian subcontinent (Roy, 1994; Gommans, 1999; de la Garza, 2016). To proxy for local exposure to pre-colonial interstate military competition, we compute a benchmark measure in which a district's exposure is increasing in its physical proximity to pre-colonial conflicts. Our empirical analysis reveals a positive and substantial relationship between pre-colonial conflict exposure and local economic development.

We perform numerous robustness checks. First, we restrict our analysis to within-state variation by including state fixed effects, to show that features specific to Indian states do not drive our results. Second, we control for a range of local geographic features, including climate, terrain ruggedness, soil suitability, disease environments, and waterway access. Third, we do an instrumental variables analysis that exploits variation in pre-colonial conflict exposure driven by the cost distance to the Khyber Pass, the main historical route of invaders from Central Asia into India. Fourth, we show that pre-colonial conflict exposure significantly predicts local development levels today above and beyond the role of colonial-era institutions such as direct British rule and non-landlord revenue systems. We show that this relationship continues to hold after controlling for inter-ethnic relations, colonial-era and post-colonial conflict exposure, initial state development levels, and additional geographic controls. Fifth, we demonstrate that our main results are robust to alternative development outcomes, including district-level GDP per capita, and a variety of alternative indicators of living standards. We control in all specifications for local population density, so our main results are not driven by the greater prevalence of pre-colonial conflict in more populated zones, nor do they simply capture greater population density in conflict-affected zones today.

We next analyze the channels through which pre-colonial warfare in India may have influenced long-run development differences. Here, we draw on rich data from both archival and secondary sources. Consistent with the "war makes states" logic described above, we find evidence for a positive and significant relationship between local exposure to pre-colonial conflict and early local state-making, the development of local fiscal capacity, and greater political stability in the long term. Furthermore, we find that pre-colonial conflict exposure significantly predicts larger investments in irrigation and related improvements in agricultural productivity in the long term, along with more investment in literacy and primary education. We view local public goods investments as functions (at least in part) of more powerful local government institutions and greater political stability.

Our study provides new evidence that the "war makes states" framework applies outside Western Europe, at least in India. Recent literature has explored the relationship between interstate military competition and long-run state development in India (Roy, 2013, 13-38, Gupta, Ma and Roy, 2016, Foa, 2016). We go further in several ways. First, we analyze the long-run implications of historical warfare in India for *economic development*, rather than for state development only. Second, we construct a new, geocoded database of historical conflicts on the Indian subcontinent. We compile a rich array of new data to evaluate pre-colonial and colonial outcomes. Thus, the scope of our analysis is significantly wider than the previous literature.

More generally, our study sheds new light on the historical roots of Indian development patterns. The vast majority of the literature analyzes the role of British colonialism (Banerjee and Iyer, 2005; Iyer, 2010; Bharadwaj and Mirza, 2017; Castelló-Climent, Chaudhary and Mukhopadhyay, 2017; Chaudhary et al., 2018; Lee, 2018). Our empirical analysis considers the importance of colonial factors in several ways, but highlights the potential role of pre-colonial events in India, which this literature tends to overlook.

There is an influential literature about the significance of pre-colonial factors for long-run development (Gennaioli and Rainer, 2007; Hariri, 2012; Michalopoulos and Papaioannou, 2013; Dell, Lane and Querubin, 2018). Among these papers, ours is one of the first to analyze systematically the long-run development consequences of pre-colonial history in India. Furthermore, unlike much of the literature, we focus on the long-run consequences of pre-colonial warfare, rather than pre-colonial levels of state centralization – which, according to our "war makes states" framework, was actually an outcome of prior military conflict. We show evidence in support of this view.

Finally, there is a growing body of quantitatively-oriented research on pre-colonial India (Jha, 2013; Gaikwad, 2014; Iyer, Shrivastava and Ticku, 2017), to which our study brings the role of pre-colonial interstate military competition and warfare on long-run development patterns. We introduce new insights about the pre-colonial military roots of contemporary economic differences across India.

We organize this study as follows. In the next section, we develop our theoretical framework. Section 3 contains the historical background, and Section 4 a description of the data. In Section 5, we present the empirical strategy and main results, and in Section 6 the tests for robustness. In Section 7, we analyze potential channels, and the conclusion is in Section 8.

2 Theoretical Framework

We develop a simple theoretical framework to show how pre-colonial events may influence contemporary economic outcomes.

Interstate warfare is a common explanation for long-run state-making, at least in the Western European context (see Section 1). We can characterize the general logic as follows (Besley and Persson, 2011, 58-9). Protection from foreign attack is a public good typically provided by the government. To improve the government's ability to fend off foreign attacks, individuals may demand new investments in defense, and be willing to pay more tax to fund it. In this manner, the threat of foreign attack may drive higher tax revenues, along with a bigger and more competent public administration to help organize the government's fiscal and military efforts. If there are recurring threats, then institutional reforms may continue in ratchet-like steps (Rasler and Thompson, 2005, 491-3). Once the state has decided to overcome the high fixed costs of increasing its defense capacity, then it should be inexpensive at the margin to maintain its enhanced activity levels. Thus, more powerful government institutions may stay in place even after threats dissipate. Similarly, they may survive regime changes, as new rulers take advantage of current institutional structures, rather than trying to build them from scratch.

In time, a more powerful government may help promote long-run economic development through at least two channels. The first channel is greater political stability. A more powerful government should be better at maintaining law and order (Morris, 2014, 3-26). If there is a reduction in the risk of violence and theft, then individuals will be more willing to make growth-enhancing investments in human and physical capital (North, 1981, 24-6).

The second channel is the provision of other public goods beyond political stability (Dincecco, 2017, 11-13). For example, a more powerful government may provide agricultural infrastructure such as irrigation that improves crop yields. Similarly, it may promote human capital formation by supporting literacy and education. Finally, a more powerful government may facilitate transportation infrastructure such as railways, paved roads, and canals that reduce the costs of trade.

We posit that, if a given zone in India experienced more pre-colonial warfare, then we would expect more powerful local government institutions to have emerged, which in turn would help promote local long-run economic development through the two channels described above (i.e., political stability, basic public goods provision). This is the intuition behind our empirical analysis.

3 Historical Background

We provide a brief overview of warfare and state-making in pre-colonial India in terms of our theoretical framework.

There were numerous independent states on the Indian subcontinent circa 1000, the start year of our analysis (Nag, 2007, 28), and political fragmentation was an enduring feature (de la Garza, 2016, 12). By the early sixteenth century, major rival states included the Delhi Sultanate, the Rajput states, the Deccan Sultanates, and the Vijyanagara Empire (Roy, 1994, 57).

Each of these pre-colonial states was capable of mobilizing a large military (Roy, 1994, 57-70). There are reports that Sultan Alauddin Khilji of Delhi had 475,000 cavalry troops, and that the Vijyanagara Empire had a million-person army. There is also evidence of early institutional change in response to external threats. Under King Krishna Devaraya, for example, the Vijyanagara Empire introduced new weaponry and cavalry, and expanded state control by establishing new military garrisons.

Between 1526 and 1707, the Mughal Empire was the most powerful state on the Indian subcontinent (Richards, 1995, 1, 6-9; de la Garza, 2016, 1). This Empire was established by Babur, who after several failed attempts, finally defeated the Afghan state led by Ibrahim Lodi in 1526. The next year, Babur's relatively small army defeated a large Rajput confederacy of 80,000 cavalry troops and 500 war elephants, helping establish Mughal political control over northern India. Babur's well-honed enveloping tactics were very effective in this campaign.

According to Nath (2018, 245), "The Mughals fought their enemies ceaselessly... war was a constant preoccupation of the Mughal Empire." The Mughal Empire reached new heights under Akbar, who ruled from 1556 to 1605 (Richards, 1995, 12-28). During his long reign, Akbar conquered numerous rival kings and local strongmen, enabling the Mughals to further solidify their control over the northern and western parts of India.

The Mughals committed significant fiscal resources to war-making (de la Garza, 2016, 1; Nath, 2018, 253-5). Describing the 1596 state budget, for example, Richards (1995, 75) writes that "by far the greater part of this budget was devoted to supporting a massive military establishment." More than 80 percent of total state expenditure was granted to Mughal military officials called *mansabdars*, while another 9 percent was devoted to the central military establishment (Richards, 1995, 75-6). By contrast, annual spending on the Mughal imperial household was less than 5 percent.

To help manage Mughal military and fiscal affairs, Akbar implemented centralized administrative structures (Richards, 1995, 58-9; de la Garza, 2016, 6). Under the *mansabdari-jagirdari* institutional system, for example, Akbar granted land to military officials in order to extract as much surplus agricultural output as possible (Nath, 2018, 253-5). Fiscal data available for the late 1680s indicate that the top 6 percent of military officials (roughly 450 persons) were in possession of more than 60 percent of total tax revenue, indicating a high concentration of fiscal control by a small military elite (Qaisar, 1998, 255-6). A good portion of such funds were spent on troop maintenance and other military expenses.

Mughal government officials developed a "pyramid" treasury system that linked the central treasury with those in provincial capitals and towns (Richards, 1995, 69-71). Akbar exploited this system to move funds quickly in times of conflict. Richards (1995, 70) writes that the "swift dispatch of treasure gave his armies the means and morale for victory."

The *zabt* land tax revenue system was another Mughal fiscal innovation (Richards, 1995, 187-90). In the late sixteenth century, the state began to overhaul the land tax revenue system, increasing centralization and introducing more accurate agricultural data. By enabling the state to deal directly with individual farmers, the *zabt* system helped reduce the traditional tax power of local landowners called *zamindars*. The *zabt* system further improved the ability of the Mughal Empire to extract agricultural output and finance the military. The system may also have incentivized farmers to shift production to high-value cash crops, thereby promoting rural economic development.

The death of Emperor Aurangzeb in 1707 helped trigger the decline of the Mughal Empire (Richards, 1995, 253-81). Indigenous states including the Maratha, Mysore, and Travancore began to compete for political control with the British East India Company (Roy, 1994, 37-50). Given heightened interstate military competition, states made new state-building attempts (Stein, 1985, 391; Roy, 1994, 37-50; Ramusack, 2003, 12). In Travancore, for example, King Marthanda Varma established a "warrior state" during the 1730s and 1740s, characterized by a larger bureaucracy and a more centralized tax system capable of extracting more revenue (Foa, 2016, 93-4). Describing this system, Foa (2016, 94) writes that the "flow of revenues to the center allowed the state to build a highly centralized military force, as well as to invest large sums on the construction of fortifications, temples, and palaces." In 1741, the Travancore military defeated the Dutch East India Company (Foa, 2016, 94).

From the time of its victory at the Battle of Plassey in 1757, the British East India Company became a prominent political entity on the Indian subcontinent, marking the end of the pre-colonial era (Dutt, 1950, 1-2; de la Garza, 2016, 12). Over the next century, the EIC systematically defeated its rivals in India, including indigenous states such as the Marathas, Mysores, and Sikhs, along with foreign powers such as the Dutch and French (Dutt, 1950, 1-2; Gommans, 1999, 120).

Even under British colonialism, however, there is good reason to think that pre-colonial institutions continued to play a role in local development outcomes.[1] The total influx of British settlers to India was relatively small (Iyer, 2010, 697). Thus, British colonialists had

[1]Stein (1985) provides a general overview of the institutional continuity between military-fiscal developments in pre-colonial India and British-era colonial rule.

incentives to establish indirect forms of rule, under which traditional leaders retained ample control over internal governance matters – particularly in zones with well-developed pre-colonial institutions (James, 1997, 326-33; Gerring et al., 2011, 380-7; Hariri, 2012, 473-4). Princely states (i.e., "Native" states) ruled by hereditary kings spanned 45 percent of British India (Iyer, 2010, 694), while colonial dependence on "customary" (i.e., traditional) courts was 60 percent (Lange, 2004, 909). Drawing on qualitative work, moreover, Lange (2004, 909) suggests that "the minimal colonial state created local conditions in both the directly and indirectly ruled areas of colonial India that were quite similar to those in indirectly ruled Africa." By relying on existing pre-colonial institutions, the British could reduce overall governance costs. For example, the *zabt* land tax revenue system established by the Mughals was a precursor to the later *raiyatwari* individual cultivator system used by the British (Banerjee and Iyer, 2005, 1192-94).[2] We may therefore expect local pre-colonial institutions to have lasted – to a non-trivial extent – throughout the colonial era and into the post-colonial one. Nonetheless, our empirical analysis controls for colonial institutions such as direct British rule (Iyer, 2010) and non-landlord revenue systems (Banerjee and Iyer, 2005).

Overall, this historical overview links interstate military competition and recurrent warfare in pre-colonial India to lasting local institutional change and state development. By (eventually) providing political stability and basic public goods, such institutions could promote local long-run economic development.

4 Data

4.1 Historical Conflict

4.1.1 Data Construction

Given data limitations, we use historical conflict data to proxy for pre-colonial interstate military competition. The logic is that there was a positive link between the actual prevalence of conflict and threats from military rivals.

To construct our historical conflict database, we rely on the far-reaching book by Jaques

[2]Similarly, the replacement of traditional chiefs with non-local bureaucrats in order to improve tax collection by Tipu Sultan, the late-eighteenth-century ruler under whom "military fiscalism was brought to its highest point" in Mysore, became an important feature of later British colonial governance (Stein, 1985, 401, 406).

(2007). The main goal of this book is to document as many historical conflicts as possible (Jaques, 2007, xi). For inclusion, a conflict must have been written down and cross-referenced with a minimum of two independent sources (Jaques, 2007, xiii). Although this selection criteria will tend to exclude historical conflicts known only through oral history, this potential shortcoming appears to be more severe in pre-colonial Africa than in other world regions (Jaques, 2007, xi).

The conflict information in Jaques' book is organized alphabetically by individual conflict name. For each individual conflict, Jaques provides a paragraph-length description of the conflict, including the type (e.g., land battle), date, approximate duration (e.g., single-day), and approximate location. For example, the first conflict in our database, named "Peshawar," took place on November 27, 1001 as part of the Muslim conquest of Northern India. Here, Mahmud of Ghazni defeated Raja Jaipal of Punjab and his coalition of Hindu princes just outside the city of Peshawar. To proxy for the location of this conflict, we assign the geographical coordinates of Peshawar (34° 1′ 0″ N, 71° 35′ 0″ E). Our database includes all of the individual conflicts – land battles, sieges, naval battles, and other violent conflicts (e.g., mutiny) – on the Indian subcontinent between 1000 and 2010 as recorded by Jaques. By "Indian subcontinent," we mean the modern-day nation of India plus the border nations of Bangladesh, Bhutan, Myanmar, Nepal, Pakistan, and Sri Lanka. We exclude China, since, historically, the Himalayas obstructed China-India interactions. There were few if any historical conflicts anywhere near China's border with India (Dincecco and Wang, 2018, Figure 2). For robustness, we restrict our benchmark conflict exposure measure to within modern India only in the Appendix. The main results remain unchanged. Figure 1 maps the locations of these conflicts, while Appendix Figure A.1 breaks them down by historical sub-period.

To double-check the extent of our historical conflict coverage, we constructed alternative conflict data according to similar procedures from two other independently-produced books, Clodfelter (2002) and Naravane (1997). Clodfelter is a well-regarded source on historical conflicts, and covers the globe from 1500 onward. Here, a key advantage of Jaques is that his conflict coverage extends much further back in time. Nonetheless, the pre-colonial conflict coverage between 1500 and 1757 is similar for Jaques and Clodfelter, providing support for our benchmark measure. Naravane's book focuses on battles in medieval India. While his coverage does expand on Jaques, it lacks details on individual conflicts.[3] Regardless, in the Appendix, we add the non-overlapping pre-colonial data from Clodfelter and Naravane to our benchmark conflict exposure measure. The main results remain significant.[4]

Although we double-check the breadth of our conflict coverage, there may still be measurement error. For example, the quality of the historical conflict data may vary by geographic zone and era. Our regression analysis will account for such potential differences in historical data quality in several ways, including the use of fixed effects for individual Indian states (or by grid cells), controls for initial state capacity, and the exclusion of individual states (or colonial provinces) from our main specification one at a time.

4.1.2 Benchmark Measure

To compute local exposure to individual conflicts, we follow Cassidy, Dincecco and Onorato (2017), defining the conflict exposure of Indian district i as:

$$\sum_{c \in \mathcal{C}} (1 + \gamma distance_{i,c})^{-1}. \quad (1)$$

We measure $distance_{i,c}$ from the centroid of district i to the location of conflict c. The nearer a district is to a particular conflict, the more exposed that district is. A conflict occurring at the district centroid receives a full weight of one; the weight declines as distance increases. The parameter γ controls the speed at which this decay occurs. In our baseline analysis we set γ equal to 1, but show that alternative levels of γ give similar results. To reduce the measure's sensitivity to any single conflict, we add one to $distance_{i,c}$ before taking the inverse.[5]

Our benchmark conflict exposure measure includes pre-colonial land battles between 1000 and 1757 with a cutoff distance of 250 kilometers, beyond which we assume that conflict exposure is zero. Thus, pre-colonial conflict exposure to conflict c will only take a positive, non-zero value for district i if this conflict falls within a 250 kilometers radius from the centroid of this district. In the Appendix, we use an alternative cutoff distance of 5,000

[3]We rely on Appendix B of Naravane's book, which only lists the year, name, victor, and opponent of each medieval battle. To identify conflict locations, we supplemented this information with online research.

[4]Brecke (1999) is another potential alternative source for historical conflict data. Relative to Jaques, however, there are two main shortcomings of this work: (1) similar to Naravane, he does not provide specific information about conflict locations; and (2) his data do not start until 1400.

[5]If we did not add one to this measure, then a district in which a conflict took place very nearby would receive a very large conflict exposure value, regardless of its proximity to any other conflicts

kilometers. Similarly, we use a variable end-date cutoff that allows us to also include exposure to conflicts that took place after 1757 but prior to British conquest of a district, for cases in which Banerjee and Iyer (2005) have coded the date of conquest as taking place after 1757.[6] The coefficient estimates are very similar in magnitude and significance to the main estimates across both checks.[7] We focus on interstate land battles for our benchmark measure, since they were by far the most common pre-colonial conflict type, and because they typically took place in the countryside, thereby reducing the likelihood that physical infrastructure would be destroyed. Still, for robustness, we control for local exposure to (1) pre-colonial sieges and (2) all pre-colonial conflict types in the Appendix. The main results do not change in either case, and there is no significant relationship between pre-colonial siege exposure and current development.

Figure 2 maps pre-colonial conflict exposure across Indian districts.[8] This figure suggests that there were four main geographic zones of pre-colonial conflict: (1) the far north in the vicinity of the state of Punjab; (2) the western coast in the vicinity of Maharashtra; (3) the far east in the vicinity of West Bengal; and (4) the lower southeast in the vicinity of Tamil Nadu. Conflict exposure in the first zone was largely driven by the Mughal conquest of Northern India over the sixteenth century, the Mughal-Sikh Wars of the late 1600s and early 1700s, and to a smaller extent the fourteenth-century Wars of the Delhi Sultanate and the Indian Campaigns of Ahmad Shah during the mid-1700s. Conflict exposure on the western coast was driven by a diverse set of conflicts, including the Wars of the Delhi Sultanate, the Mughal-Ahmadnagar Wars of the early 1600s, and the Mughal-Maratha Wars during the second half of the seventeenth century. In West Bengal, conflict exposure was largely a function of the later Mughal-Maratha Wars. In Tamil Nadu, the Second Carnatic War during the mid-eighteenth century generated most of the conflict exposure.

4.1.3 Alternative Measures

We view the benchmark conflict exposure measure described above as the most straightforward way to analyze whether greater pre-colonial threats from military rivals drove insti-

[6]Specifically, this measure includes conflicts from 1000 to the (potentially post-1757) year of British annexation, which differs by district.

[7]For further robustness, we exclude 155 districts for which our benchmark conflict exposure measures takes a value of zero. The main results continue to hold (not shown to save space).

[8]Similarly, Appendix Figure A.2 maps the residualized conflict exposure measure after controlling for log population density.

tutional change. Still, our benchmark measure may overlook conflicts that were fought far from a pre-colonial state's political center, but nonetheless prompted institutional reforms. Thus, we compute four alternative measures of conflict exposure. We code each major state participant in our pre-colonial conflict database, and identify its capital city. In our first alternative measure, we mimic König et al. (2017) and compute the convex hull for each participant according to the geographical coordinates of the conflicts that participant took part in. We treat all districts that intersect this convex hull as affected by a conflict, whether directly in battle or by troops on the march between battlefield locations. For each district, we count the number of conflicts which are so treated. Oftentimes, Jaques classifies a cluster of conflicts under the same broad title, such as "Later Mughal-Maratha Wars." Thus, as a second alternative measure, we also compute the convex hull for each broad group of conflicts. We treat all districts that intersect this convex hull as affected by this group of conflicts. For each district, we count the number of conflicts which are so treated.

Third, we compute conflict exposure again using Equation 1, replacing the locations of the conflicts with those of the capitals of the pre-colonial states that participated in them. For a conflict with three actors, for example, this measure threats the conflict as if it is three events. Finally, we count the number of conflicts for each pre-colonial state, and assign these conflicts to the district that houses the state's capital. That is: for each district, we count the number of conflicts fought by a state headquartered in that district.

In the Appendix, we show the results for these alternative measures of pre-colonial conflict exposure. Our main results remain robust.

4.2 Economic Development

To proxy for contemporary levels of development in India, we use nighttime luminosity data. Luminosity is a common measure of local economic activity in relatively poor regions (e.g., Henderson, Storeygard and Weil, 2012; Michalopoulos and Papaioannou, 2013; Min, 2015, 51-73). We take these data from the Operational Linescan System of the Defense Meteorological Satellite Program of the US Air Force. Satellite images are taken between 20:30 and 22:00 local time, and are averaged over the year. These are reported in integer values from 0 to 63 for pixels at a 30-second (roughly one square kilometer) resolution. We compute average luminosity across all square kilometer cells within each district for every year between 1992 and 2010, and then take the district averages over the entire 1992-2010 period.[9]

Figure 3 maps average luminosity by district in India. This figure suggests that economic development levels tend to be highest in the vicinity of the four main geographic zones of pre-colonial conflict as described in the previous subsection. Appendix Figure A.3 indicates similar spatial patterns for residualized luminosity after controlling for log population density. Thus, high luminosity levels do not simply depend on densely-packed populations. Nonetheless, to account for this possibility, and to alleviate the concern that conflict locations are driven by the presence of population centers, our regression analysis will always control for population density.

Taken together, Figures 2 and 3 highlight the spatial correlation between pre-colonial conflict exposure and local economic development today. To further test the strength of this relationship, Appendix Figure A.4 plots pre-colonial conflict exposure against luminosity (residualized). There is a strong positive correlation between the two measures.

While district-level GDP per capita data do exist for India, they have been constructed by a private company, and differ from official sources such as the National Sample Surveys in their rankings of districts on economic development outcomes. They are not widely used in the empirical literature (Himanshu, 2009; Castelló-Climent, Chaudhary and Mukhopadhyay, 2017, 5). Nonetheless, we employ them as an alternative development outcome in the Appendix. The main results are similar. Similarly, in Section 7, we take a variety of indicators of living standards as alternative outcomes.

4.3 Channels

4.3.1 Pre-Colonial Variables

We use two principal measures of pre-colonial state building as outcomes: the first is the number of important Mughal sites reported by Schwartzberg (1978). In particular, we georeference plate VI.A.4, "Religious and Cultural Sites of the Mughal Period, 1526-1707" and count the number of sites within each modern district. These sites include a range of public works, such as forts, palaces, and bridges. Second, we use maps of the Mughal Empire digitized by Jha (2013) to identify districts incorporated by Babur (1526-30), Akbar (1556-1605),

[9] In Appendix Figure A.5, we show the main results when we restrict the luminosity data to each year from 1992 to 2010. The coefficient estimates always remain highly significant, although they decline gradually in magnitude over time.

and Aurangzeb (1658-1707). Following others in the literature (e.g., Bockstette, Chanda and Putterman, 2002; Heldring, 2018) we interpret the longevity of state history as a measure of state capacity.

4.3.2 Colonial Variables

We construct colonial fiscal data for the late nineteenth century according to Baness (1881), a secondary archival source. This book contains information on the land tax revenue, physical size, and population for several hundred historical Indian administrative units under direct or indirect British rule. Here, indirect rule refers to major Princely states.[10] To match historical states to modern districts, we rely on the information on provincial and state names in Baness' book.[11] To complement the late nineteenth-century fiscal data, we rely on the 1931 land tax revenue data for districts in British India from Lee (2018).

Beyond the colonial fiscal data, we proxy for colonial institutions in India in two other ways. First, we identify districts that were under direct British rule according to Iyer (2010). Second, we identify the proportion of each district in British India under a non-landlord revenue system according to Banerjee and Iyer (2005).

We take data on irrigation infrastructure at the district level in 1931 from Bharadwaj and Mirza (2017). Similarly, we take district-level literacy data from the 1881 and 1921 censuses, as coded by Fenske and Kala (2017). Finally, we take data on the establishment of the colonial railway from Fenske and Kala (2018). Following a procedure similar to Donaldson (2018), they identify the first date at which a segment of the railroad appears in each district in our data.

4.3.3 Post-Colonial Variables

Political Stability

We use several indicators of district-level political stability. First, we compute our benchmark measure of conflict exposure for the colonial and post-colonial eras. Here, we divide British colonial rule into two distinct sub-periods, 1758-1839 and 1840-1946, with the cut-

[10] Following Iyer (2010, 695), we focus on Princely states that received British ceremonial gun salutes. We identify gun salute status in the late nineteenth century according to the main text of Chakrabarti (1896).

[11] We compute conflict exposure here in terms of the distance from the capital city as recorded by Baness or approximate centroid (if capital city information was not available) of each historical state to each conflict location, and then match them to modern districts.

off marked by the establishment of complete British military and political dominance over the Indian subcontinent by the 1840s (Clodfelter, 2002, 242-50).[12] Second, we measure political violence in terms of fatalities per district between 2010 and 2018 (in hundreds) from the Armed Conflict Location and Event Data Project.[13] Third, we use local Maoist control in 2003 following Mukherjee (2017). In 2006, Prime Minister Manmohan Singh called the Maoist insurgency "India's most important internal security threat" (Mukherjee, 2017, 5). Finally, we compute local linguistic and religious fractionalization levels in 2001, according to the district-wide populations by language and religion in Omid's Peoples of South Asia Database.

Public Good Provision

In line with the literature, we mimic the post-colonial development outcomes in Banerjee and Iyer (2005) and Iyer (2010), who provide district-level data across more than 10 major Indian states on agricultural investment and productivity, literacy and education, health, and transportation infrastructure outcomes. Following Banerjee and Iyer, we average these data between 1956 and 1987 when possible.

4.4 Control Variables

4.4.1 Benchmark Controls

Geographic zones with mild climates and high quality soils may promote dense human settlements and economic development (Ashraf and Galor, 2011). Dense settlements, however, may reduce the cost of collective military action and promote violent conflict (Besley and Reynal-Querol, 2014). Similarly, populated and developed zones may make for attractive targets for attackers (Glaeser and Shapiro, 2002). Local geography, therefore, may influence patterns of both pre-colonial conflict and economic development alike.

To account for this possibility, our regression analysis will control for a wide range of district-level geographic and climate features, including latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and

[12] Alternatively, we may identify 1857 as the cutoff year for the two sub-periods of British colonial rule. This year marked the start of the Sepoy Mutiny (1857-9), along with rule by the British Crown (versus the East India Company). All of the results described in Table 5 remain similar in terms of sign and significance for this alternative cutoff (not shown to save space).

[13] Available at: `https://www.acleddata.com/`.

malaria risk. Here, we focus on control variables that are not likely to be influenced by post-treatment changes after the year 1000. We compute latitude and longitude by identifying district centroids using a polygon file of district boundaries from `gadm.org`. The data for altitude, precipitation, dry rice suitability, wet rice suitability, and wheat suitability are taken from the Food and Agriculture Organization's Global Agro-Ecological Zones (FAO-GAEZ) project. They are originally available as raster data; we compute district-level measures by averaging over raster points within each district. Similarly, we compute ruggedness according to the raster data made available by Nunn and Puga (2012). We take raster data on land quality from Ramankutty et al. (2002). Finally, we take the raster index for the stability of malaria transmission from Kiszewski et al. (2004).

4.4.2 Additional Controls

In Subsection 4.3 above, we interpret the variables for pre-colonial and colonial institutions, political stability, and public goods provision as potential channels that mediate our main results. Alternatively, we may interpret them as potential confounders for which we should control. Our empirical analysis will thus consider both the degree to which they are predicted by pre-colonial conflict and the degree to which controlling for them affects our coefficient estimates in our main results. We also control for additional geographic features (i.e., distance to coast, distance to external border, distance to natural resources, river presence, irrigation potential, and rainfall). We discuss the details of these additional controls in Subsection 6.2.

5 Pre-Colonial Warfare and Economic Development

5.1 Empirical Strategy

To analyze the relationship between pre-colonial conflict exposure and local development outcomes across India, we estimate the following OLS specification:

$$Y_{i,j} = \beta ConflictExposure_{i,j} + \lambda PopDensity_{i,j} + \mu_j + X'_{i,j}\phi + \epsilon_{i,j}, \tag{2}$$

where i indexes districts and j indexes states in peninsular India. $Y_{i,j}$ measures local economic development levels in terms of luminosity. Here, we follow Michalopoulos and Papaioannou (2013) and take the natural logarithm, adding a small number such that $Y_{i,j} \equiv \ln(0.01 + Luminosity_{i,j})$. This log transformation reduces the range of the mean and variance of $Y_{i,j}$, and allows us to make use of all observations. In the Appendix, we show that the main results remain robust, however, if we: (1) take $\ln(1 + Luminosity_{i,j})$ rather than $\ln(0.01 + Luminosity_{i,j})$; (2) keep $Y_{i,j}$ in its original linear form; or (3) take the inverse hyperbolic sine function. $ConflictExposure_{i,j}$ measures pre-colonial conflict exposure, our variable of interest. We always report the original coefficient estimate β and the standardized beta coefficient. As noted in Subsection 4.2, local luminosity levels do not appear to simply reflect population density. Still, to account for this possibility, $PopDensity_{i,j}$ controls for log population density in the most recent year available prior to the year in which the dependent variable is measured.[14] For the main regression analysis, this year is 1990.[15] μ_j is the fixed effect for each of the 36 federal states (more precisely, 29 states and 7 union territories). The vector $X_{i,j}$ denotes the geographic controls as described in Subsection 4.4.1. Finally, $\epsilon_{i,j}$ is the error term.

Our main regression analysis reports robust standard errors and p-values. In the Appendix, we report the p-values obtained according to three alternative treatments of standard errors as robustness checks. Specifically, we report the p-values for: (1) standard errors robust to clustering at the state level; (2) tests of β using the wild cluster bootstrap at the state level (Cameron, Gelbach and Miller, 2008), based on 9,999 replications; and (3) standard errors that allow for general forms of spatial autocorrelation of the error term (Conley, 1999), for a cutoff distance of approximately 250 kilometers. The main results remain highly significant for both cluster-robust standard errors and Conley spatial standard errors, and just miss statistical significance for the wild cluster bootstrap procedure.[16]

Appendix Table A.1 contains the summary statistics for the regression variables used in our analysis.

[14] To account for potential non-linearities in population density, we also include the quadratic and cubic terms as a robustness check. The main results continue to hold (not shown to save space).

[15] When the dependent variable is historical (e.g., 1881), then this year is subject to data availability.

[16] To further account for the possibility that spatial correlation leads to standard errors that are too small, we follow Kelly (2019) and generate artificial spatially-correlated noise placebo variables to replace our variable of interest, reallocating conflict exposure randomly across districts within a state (without replacement). The Moran's I statistic for the full specification with state fixed effects and controls is 0.044, indicating spatial autocorrelation in the regression residuals (Appendix Table A.3). However, the placebo variables nearly always fail to produce treatment effects as large as those of our main coefficient estimates (Appendix Figure A.9).

5.2 Main Results

Table 1 shows the main results for the relationship between pre-colonial conflict exposure and current economic development across Indian districts. In column 1, we report the result for the bivariate correlation after controlling for log population density. The (unstandardized) coefficient estimate for $ConflictExposure_{i,j}$ is 3.713, and is significant at the 1 percent level. Column 2 adds state fixed effects. The coefficient estimate falls to 1.601, but remains significant. In column 3, we add the controls for local geography. The coefficient estimate is similar in size and significance to the previous specification.

Overall, the Table 1 results support the main "reduced-form" prediction of our theoretical framework. Namely, there is a positive and highly significant relationship between local exposure to pre-colonial conflicts and levels of economic development in India today. The coefficient estimate in column 3 indicates that a one standard deviation increase in pre-colonial conflict exposure predicts a 0.10 standard-deviation increase in current luminosity levels. This is similar in magnitude to the effect of pre-colonial political centralization on current luminosity levels in Africa found by Michalopoulos and Papaioannou (2013, 130) (they report a standardized beta coefficient of 0.12). It is also broadly similar in magnitude to the finding by Banerjee and Iyer (2005, 1203) in their study of the relationship between districts in British India under a non-landlord revenue system and post-colonial agricultural productivity (i.e., a main outcome variable of theirs, for which one can compute a standardized beta coefficient of 0.14).

This magnitude is also broadly similar in size that documented by Banerjee and Iyer (2005, 1203) in their study of the relationship between the proportion of a district in British India under a non-landlord revenue system and post-colonial agricultural productivity (i.e., a main outcome variable of theirs, for which one can compute a standardized beta coefficient of 0.14.)

6 Robustness

In this section, we report our main robustness checks. First, we do an instrumental variables analysis that exploits variation in pre-colonial conflict exposure driven by the cost distance to the Khyber Pass, the principal historical route of Central Asian invaders into India. Second, we show robustness to the inclusion of additional control variables, including:

(1) colonial controls; (2) controls related to ethnic and religious fractionalization; (3) controls for post-1757 conflict exposure; (4) controls for initial state capacity; and (5) controls for additional geographic features.

In the Appendix, we show that the main results are generally robust to numerous other checks, several of which we have already mentioned. They include: (1) the estimation of results separately by year (Appendix Figure A.5); (2) excluding federal states from the data one by one (Appendix Figure A.6); (3) excluding colonial provinces from the data one by one (Appendix Figure A.7); (4) including pre-colonial sub-periods from the data one by one (Appendix Figure A.8); (5) generating artificial spatially-correlated noise placebos (Appendix Figure A.9); (6) the use of fixed effects by grid cells (i.e., 4° latitude $\times$ 4° longitude) rather than by Indian states (Appendix Table A.2); (7) the alternative computation of standard errors (Appendix Table A.3); (8) an alternative distance cutoff in computing conflict exposure (Appendix Table A.4); (9) a running cutoff that includes exposure to post-1757 conflicts (Appendix Table A.5); (10) alternative sources of conflict data (Appendix Table A.6); (11) the disaggregation of the results by conflict type (Appendix Table A.7); (12) alternative conflict exposure measures that exploit information about the states that participated in the conflicts (Appendix Table A.8); (13) alternative functional forms for the dependent variable (Appendix Table A.9); and (14) the use of GDP per capita as an alternative measure of economic development (Appendix Table A.10).

6.1 Instrumental Variables

6.1.1 Historical Information

To instrument for conflict exposure, we construct a measure of each district's proximity to the Khyber Pass. As the prime land route connecting Afghanistan to Pakistan, the Khyber Pass has historically served as what Gommans (2003, 23) calls "the most important highway to India." The South Asian subcontinent is naturally protected from invasion by several mountain ranges, including the Himalayas, Hindu Kush, Spin Ghar, and Arakans. Historically, the Khyber Pass has been the principal route for invaders coming from Central Asia to India. Thus, proximity to the Khyber Pass can be treated as a forcing variable that affects a district's exposure to the historical conflicts in our database.

Numerous invaders from Central Asia have either come through the Pass or have sought to control it (Docherty, 2008). Mahmud of Ghazni appears as a participant in the first conflict on our database (see Subsection 4.1). Ruler of the Ghaznavid Empire, Mahmud's invasions of India began in 1001, including clashes with the Shahi Kingdom, along with the trading centers of Multan and Bathinda. Muhammad of Ghor, Sultan of the Ghurid Empire, invaded Multan in 1175. His former slave, Qutb al-Din Aibak, was the first of the Delhi Sultans. Following the death of Genghis Khan, the Delhi Sultinate faced repeated raids from the Chagatai Khanate. Babur's victory at Panipat in 1526 marked the establishment of the Mughal Empire. The Persian ruler Nadir Shah made several attacks on the Mughal Empire, notably entering Delhi in 1739. The Durrani emperor Ahmad Shah Durrani attacked the Mughals repeatedly between 1748 and 1767, via the Khyber Pass. Even after the fall of the Mughal Empire, local Indian rulers such as Ranjit Singh sought to control the Pass to defend against invasions from Afghanistan.

6.1.2 IV Construction

The proximity of each district in India to the Khyber Pass in terms of simple geodesic distance (i.e., "as the crow flies") does not accurately measure how difficult it was to reach the Pass. Despite their proximity to the Pass, for example, mountainous states such as Jammu and Kashmir, Himachal Pradesh, and Uttaranchal remained less accessible to invaders than the flatter regions of Punjab, Haryana, and Uttar Pradesh. Similarly, the Aravalli Mountains were a natural barrier in eastern Rajasthan. We therefore base our measure of proximity on a cost-distance formula.

Following Özak (2010), we construct our measure of cost distance using raster data on ruggedness. Here, we follow Nunn and Puga (2012) and define the ruggedness of a cell as the average difference in absolute elevation between that cell and its eight neighbors. We assume that the cost of crossing a cell is proportionate to the square of its ruggedness. Using the Cost Distance tool in ArcMap, we compute the least-cost path and associated cost of travel between each grid cell in India and the Khyber Pass. Averaging over all cells in a district gives us our main "cost-distance" measure to the Khyber Pass. Figure 4 maps these values.

As distance from the Khyber Pass increases, the relationship between our cost-distance measure and conflict exposure becomes nonlinear, driven by conflicts such as the Carnatic Wars that were unrelated to Central Asian invasions. Hence, in our benchmark measure, we compute the Khyber proximity instrument as a dummy for whether a district is in the set of 50 districts that are closest to the Kyber Pass as measured by cost distance. Figure 5 plots this measure. Ahead, we show robustness to alternative measures of our instrument. First, we show results replacing our cutoff of the closest 50 districts with a cutoff of the closest d districts, where $d \in \{30, 31, ..., 80\}$. A range of cutoff values give results similar to our benchmark. Second, we use several alternative measures of the cost of crossing a grid cell: linear ruggedness, squared slope, linear slope, and a Human Mobility Index based on Özak (2010, 2018). Third, we show robustness to possible violations of the exclusion restriction by (1) controlling for measures of historical trade and (2) reporting placebo results that consider cost-distance proximity to other points of entry into South Asia that conquerers did not historically rely on.

6.1.3 IV Results

Table 2 reports the first-stage and instrumental variables results. Across columns, our first-stage results suggest that being proximate to the Khyber Pass increases exposure to conflict by between 0.08 and 0.21 units. These magnitudes are roughly comparable to the standard deviation of our conflict exposure variable of 0.10. The Kleibergen-Paap F-statistics (KPF) are larger than 10, indicating instrument strength and a relatively low propensity for bias at the second stage. The second-stage coefficient estimates suggest, across columns, that a one-unit change in conflict exposure increases luminosity by between 3.5 and 4.9 units. The former estimate is comparable to the first column of the main results from Table 1, corresponding to a standardized effect size of slightly more than 0.20. Our IV estimates diminish less than our OLS results as state fixed effects and controls are added.

It is apparent from Table 2 that the IV estimates are slightly more than twice as large as the OLS estimates. Four possible explanations of this difference in coefficient magnitudes are: (1) downward bias due to omitted variables in the OLS analysis; (2) differences between compliers and the full sample; (3) measurement error; and (4) violations of the exclusion restriction. Weaker pre-colonial states or those with less capacity for later state-making may have been located closer to the Khyber Pass. Considering compliers in comparison to the full sample, districts exposed to pre-colonial conflict due to invasion from Central Asia may have had a stronger state-making response than districts exposed to pre-colonial conflict for other reasons. The possibility of measurement error is obvious. The list of conflicts in Jaques

(2007) may be incomplete, identification of conflict locations may be inexact, and the Cassidy, Dincecco and Onorato (2017) inverse distance-weighting approach may only approximate the true mapping of conflicts into district-level exposure. We address the exclusion restriction ahead.

6.1.4 IV Robustness

Exclusion Restriction

One objection to our IV strategy is the possibility that the Khyber Pass introduced South Asia to influences beyond conflict exposure that may have been relevant to both pre-colonial state-making and later economic development. Trade was the most notable such influence. We address this possibility in two ways. First, we show that controlling for three measures of access to historical trade does not significantly alter the IV results. The first employs the map of major seventeenth-century trade routes from Raychaudhuri (1982, 334). We code a district as having access to a historical trade route if it is intersected by such a route or contains a major port according to this map. The second is the UNESCO set of Silk Road sites in India.[17] We code a dummy equal to one for districts containing a UNESCO site. The third is the list of major ports of medieval trade from Jha (2013). We code a dummy equal to one for districts containing these ports. Appendix Table A.11 indicates that the IV results are robust to all three trade measures.

Second, we use as placebos alternative points of entry into South Asia that conquerers did not historically rely on. In particular, we consider the important medieval ports of Surat, Kodungallur, Goa, Calicut, and Bombay. For each of them, we compute an analogous cost-distance measure using the same ruggedness-squared approach used to compute cost distance from the Khyber Pass. We then code a placebo "instrument" equal to one for the 50 districts closest to each entry point. Appendix Table A.12 shows that these cost distances generally fail to predict conflict exposure.[18] In the instrumental variables specifications, they generally cannot be used to infer a positive effect of conflict exposure on modern luminosity. The exception is Bombay, with a weak first stage F-statistic of less than four.

Alternative IV Construction

[17] Available at: `https://whc.unesco.org/en/tentativelists/5492/`

[18] Furthermore, their reduced-form correlations with current economic development are negative and only sometimes significant (not shown to save space).

In addition, we show that the specific construction of our instrument does not drive our IV results. Rather than using the squared ruggedness of a grid cell as a measure of the cost of crossing that cell, we use one of four alternatives. They are: linear slope; squared slope; linear ruggedness; and a human mobility index (HMI). The latter measure, developed in Özak (2010, 2018), is based on the speed that a military infantry unit can maintain while walking over different terrain types, accounting for slope, temperature, and humidity. For the first three of these alternatives, we use the same "closest 50 districts" cutoff as in our baseline. For the HMI cost measure, we expand this cutoff to the closest 100 districts. This change makes sense because the HMI assigns a relatively low cost of accessing the mountainous regions of Jammu and Kashmir from the Khyber Pass, even compared to the Punjab and western Uttar Pradesh, and so does not become a robust predictor of conflict exposure unless a larger proximity cutoff is used. Appendix Table A.13 shows that these alternative measures of computing cost distance give results similar to the benchmark IV results using squared ruggedness.

We also show that using a cutoff value of 50 districts does not drive the IV results. Appendix Figure A.10 shows how the coefficient estimate and 95 percent confidence interval on our main conflict exposure measure changes as we vary this cutoff value between 30 and 80. These estimates are taken from the specification with both state fixed effects and the full set of benchmark controls. As the figure makes clear, the coefficient estimates are relatively stable throughout the range of cutoff values, and are statistically significant at the 5 percent level for a range of cutoffs from 36 through 57.

6.2 Additional Controls

6.2.1 Colonial Controls

As described in Section 1, one strand of the current literature highlights the colonial origins of contemporary development in India. Drawing on this literature, we control for the potential role of colonial institutions in two ways. First, following Iyer (2010), we include a dummy variable for direct British rule. Second, following Banerjee and Iyer (2005), we control for the proportion of each district in British India that was under a non-landlord revenue system.

Table 3 presents the results. To establish a "benchmark" coefficient value, we first rerun the main specification for the sub-sample for which the direct rule variable is available in column 1. The coefficient estimate for $ConflictExposure_{i,j}$ is 1.263, and is significant at the 1 percent level. We then add the direct rule measure as a control in column 2. The coefficient estimate for $ConflictExposure_{i,j}$ is very similar in size and significance to the previous specification.[19] In columns 3 and 4, we repeat this exercise for the non-landlord control. Once more, the coefficient estimates for $ConflictExposure_{i,j}$ are positive and highly significant.

Overall, these tests suggest that pre-colonial conflict exposure significantly predicts current local development in India above and beyond the role of colonial institutions.

6.2.2 Fractionalization

Another strand of the literature emphasizes the historical role of inter-ethnic relations in India. In Table 4, we control this factor in three ways. First, following Jha (2013), column 1 includes a dummy variable for districts that had major medieval ports. Next, we include, separately, the controls for local linguistic and religious fractionalization levels in 2001 in columns 2 and 3. In column 4, we include both fractionalization controls together. The coefficient estimates for $ConflictExposure_{i,j}$ remain positive and highly significant across all four specifications, with values between 1.422 and 1.462. These robustness checks imply that inter-ethnic relations do not confound our main results.[20]

6.2.3 Post-1757 Conflict

We account for the potential roles of colonial and post-colonial conflict exposure in Table 5. We include controls for colonial conflict exposure between 1758 and 1839 (column 1) and between 1840 and 1946 (column 3), and post-colonial conflict exposure between 1947 and 2010. In column 4, we include all three controls together. The coefficient estimates for $ConflictExposure_{i,j}$ are always highly significant, with values between 1.461 and 1.492. These results suggest that local exposure to colonial and post-colonial conflicts do not diminish the predictive importance of pre-colonial conflict exposure. Importantly, colonial

[19] For robustness, we re-ran this specification after hand-coding the direct rule variable for the missing 30-odd observations. The coefficient estimate for $ConflictExposure_{i,j}$ remains very similar to the main result in column 3 of Table 1 (not shown to save space).

[20] For further robustness, we control for the contemporary share of each district that is Muslim, and recent district-wide religious polarization levels (computed according to Montalvo and Reynal-Querol, 2005), respectively, from the Indian Census. The main results continue to hold (not shown to save space).

conflict exposure between 1840 and 1946 predicts significantly *lower* local development levels in India today, although the magnitude of this coefficient estimate is less than half the size of the main estimate. Nonetheless, this result suggests that the nature of post-1840 colonial warfare was different from pre-colonial warfare. We view this result as broadly in line with Besley and Persson (2008), who argue that intra-state conflict (i.e., versus inter-state conflict) may hinder state development.

6.2.4 Initial State Capacity

Fiscal development during the pre-colonial era may derive from previous state-making efforts. Table 6 thus controls for "initial" state development by district in several ways. First, we georeference and count the number of Indian settlements during the Neolithic and Chalcolithic Ages, respectively, according to Nag (2007, 4, 6). Second, we georeference and count the number of important Indian cultural sites between 300-700 CE and the eighth through twelfth centuries from Schwartzberg (1978, 28, 34). Third, we control for the natural logarithm of (one plus) the total urban population in the year 1000 according to Chandler (1987). Finally, we georeference and count the presence of a major Indian state between the tenth through eleventh and eleventh through twelfth centuries based on Nag (2007, 28, 30). The coefficient estimates for $ConflictExposure_{i,j}$ remain robust.[21]

6.2.5 Additional Geographic Features

In Appendix Table A.14, we repeat the main analysis after taking into account several additional geographic controls beyond those included in the benchmark specification (i.e., latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk.) We compute the natural logarithm of (one plus) distance from each district to the coast in meters. We compute each district's distance from the border as the minimum of its distances from Bangladesh, Bhutan, Burma, China, Nepal, and Pakistan, again in meters. As with distance from the coast, we then take the natural logarithm of one plus this distance. We report a "river" dummy that captures whether a district is intersected by one of the major rivers as reported in the Natural Earth Data.[22] We

[21] As another way to control for past development, we include a measure of colonial real wages between 1873-1906 at the nearest market according to data from Fenske and Kala (2017). The main results continue to hold (not shown to save space).

[22] Available at: `https://www.naturalearthdata.com/`. Alternatively, we control for whether a district is intersected by the Ganges River. The main results continue to hold (not shown to save space).

compute each district's irrigation potential using data from Bentzen, Kaarsen and Wingender (2017). To account for the prevalence of drought, we control for the mean and coefficient of variation of rainfall as reported by Matsuura and Willmott (2009). Using data from Tollefsen, Strand and Buhaug (2012), we control for the natural logarithm of (one plus) distances from the district centroid to deposits of petroleum, diamonds, gems, and gold. Columns 1 to 6 of Table A.14 include each additional geographic control independently, while column 7 includes all of them together. The coefficient estimate for $ConflictExposure_{i,j}$ remains highly significant across each specification.[23]

7 Channels

The results in Sections 5 and 6 provide support for the main "reduced-form" prediction of our argument, namely that the relationship between pre-colonial conflict exposure and current economic development levels in India is positive and highly significant. Drawing on our theoretical framework from Section 2, we now analyze the different channels through which pre-colonial warfare may have influenced long-run development.

To review, the intuition for our argument is that if a given zone in India experienced greater pre-colonial warfare, then we would expect more powerful local government institutions to have emerged there, which in turn would help promote local long-run economic development through both greater political stability and the provision of basic public goods (e.g., agricultural infrastructure).

7.1 Pre-Colonial Institutions

We start the channels analysis by testing for the link from pre-colonial conflict exposure to early state-making efforts. In Table 7, we regress each of our principal measures of pre-colonial state-making on pre-colonial conflict exposure. Column 1 indicates that there is a positive and significant relationship between pre-colonial conflict exposure and important Mughal sites, including public works (e.g., forts, palaces, bridges). Columns 2 to 4 take the longevity of state history as the outcome variable, which we operationalize in terms of districts incorporated into the Mughal Empire by Babur (1526-30), Akbar (1556-1605),

[23] Additionally, we compute distance to a presidency city as the minimum of the distance from a district's centroid to Bombay, Calcutta, or Madras. As an alternative, we compute the minimum of this value and distance from Delhi. The main result is robust (not shown to save space).

and Aurangzeb (1658-1707). Here, we follow the literature (e.g., Bockstette, Chanda and Putterman, 2002; Heldring, 2018) and interpret state longevity as a proxy for state capacity. There is a positive and significant relationship between pre-colonial conflict exposure and early state development under both Babur and Akbar (there is no statistically significant relationship for Aurangzeb). Overall, we view this evidence as consistent with our intuition, namely that pre-colonial conflict exposure played a significant role in pre-colonial state-making.

7.2 Colonial Institutions

The previous subsection suggests that pre-colonial conflict exposure significantly predicts early state-making. To complement this analysis, we regress a variety of measures of colonial institutions on pre-colonial conflict exposure in Table 8. Column 1 takes the dummy variable for districts that were under direct British rule as the outcome variable, while Column 2 takes the proportion of each district under a non-landlord revenue system in British India. Pre-colonial conflict exposure does not significantly predict either colonial outcome. In columns 3 to 8, we take log land tax revenue in 1881 as our dependent variable. We scale these data in two different ways, by area and by persons. Furthermore, we divide them up by British direct rule or indirect rule (i.e., Princely states). There is a positive and significant relationship between pre-colonial conflict exposure and colonial fiscal outcomes, particularly for districts that were under direct British rule. In the final two columns, we take log land tax revenue for districts in British India in 1931 (scaled by area and by persons) as the outcome variables. The coefficient estimates for $ConflictExposure_{i,j}$ remain positive, but do not attain statistical significance. Given that the number of sample districts for which fiscal data are available differs between 1881 and 1931, we use caution in interpreting the differences between these results. Nevertheless, when taken together, they suggest that districts that experienced greater pre-colonial conflict exposure were "early movers" in the development of colonial fiscal capacity, but that historical fiscal differences between them later diminished. We view the Table 8 results as broadly in line with Lee (2018), who highlights the importance of colonial differences in local fiscal capacity in explaining long-run development in India. Relative to Lee, our results suggest that pre-colonial conflict exposure was a significant determinant of early colonial fiscal levels.

7.3 Political Stability

Political stability was another channel through which pre-colonial conflict exposure could promote long-run economic development.

In Table 9, we regress local exposure to colonial and post-colonial conflicts on pre-colonial conflict exposure. There is a positive and highly significant relationship between pre-colonial and colonial conflict exposure between 1758 and 1839, indicating that districts that experienced greater pre-colonial conflict exposure continued to experience conflict during the first sub-period of British colonial rule. This relationship, however, is not significant for the second sub-period of British colonial rule between 1840 and 1946, and turns *negative* and highly significant for the post-colonial era. Districts that experience more pre-colonial conflict exposure, therefore, experienced significantly less conflict between 1947 and 2010.

We take several other measures of political violence as outcome variables in Table 10. Column 1 regresses the number of district-level fatalities between 2010 and 2018 according to the ACLED Project on pre-colonial conflict exposure. Here, we find a *negative* and highly significant relationship between pre-colonial conflict exposure and contemporary political violence in terms of fatalities. Similarly, in Column 2, we show that pre-colonial conflict exposure predicts a significantly lower likelihood of local control by Maoist insurgents at the start of the 2000s. Overall, these results are consistent with our theoretical framework that previous conflict exposure may pave the way for domestic peace in the long term (Morris, 2014, 3-26).[24]

Finally, we regress linguistic and religious fractionalization on pre-colonial conflict exposure in columns 3 and 4. Pre-colonial conflict predicts significantly less linguistic fractionalization today (there is no statistically significant relationship for religious fractionalization). This result suggests that a reduction in linguistic heterogeneity – via the homogenizing effects of historical conquest, for example – may be one long-run outcome of pre-colonial warfare that helps explain the anti-persistence of conflict in India which we observe.

[24] According to our theoretical framework, we would not expect to observe the anti-persistence of conflict until a dominant political entity (e.g., the post-1840 British colonial state and/or the post-colonial Indian state) was able to establish a widespread monopoly over violence. In the interim, however, we would expect warfare to persist so long as there was continuous interstate military competition. During the pre-colonial era, in fact, we find evidence for conflict persistence from one century to the next (results not shown to save space).

7.4 Other Public Goods

According to our theoretical framework, the provision of basic public goods is another channel through which pre-colonial conflict exposure could improve local development outcomes in the long run.

7.4.1 Agricultural Investment and Productivity

Table 11 takes a variety of measures of colonial and post-colonial agricultural investment as our dependent variables. Columns 1 and 2 indicate that there is a positive and significant relationship between pre-colonial conflict exposure and the proportion of agricultural land within a district that is irrigated across both the late colonial and post-colonial periods. We view these results as consistent with our framework, namely that greater conflict exposure may promote local investments in physical infrastructure. We do not find, however, any statistically significant relationship for fertilizer usage or the proportion of agricultural land sown with high-yield varieties of rice, wheat, or other cereals (columns 3-7).

In Table 12, we take local yields across all major crops, and for rice, wheat, and other cereals individually, as the outcome variables. Pre-colonial conflict exposure significantly predicts greater overall yields (column 1). This result appears to run through significant production gains in wheat, rather than in rice or other cereals (columns 2-4). The significant result for irrigation from Table 11 suggests that improvements in local agricultural infrastructure may help explain this increase in crop production.

7.4.2 Literacy and Education

Table 13 takes the local literacy rate in 1881, in 1921, averaged between 1961 and 1991, in 2001, and in 2011 as our dependent variables. While there is no significant relationship between pre-colonial conflict exposure and district-level literacy rates under British colonial rule, this relationship turns positive and significant across all three post-colonial observation points. We think these results are consistent with our theoretical framework. Namely, once newly-independent India began to pursue a state-led industrialization policy (Gupta, 2018, 2), then those districts that had been more exposed to historical conflict – and hence had developed more powerful local government institutions, and achieved local political stability – may have been better placed to invest in basic human capital.

In Table 14, we regress a variety of local education measures on pre-colonial conflict exposure. Column 4 indicates that pre-colonial conflict exposure significantly predicts greater primary school attendance in 2001. The coefficient estimate for $ConflictExposure_{i,j}$ is also positive and significant for our 1981 measure of primary education (namely, the proportion of villages having a primary school) in column 1. Taken together, we view these results as consistent with the significant results for literacy from Table 13. By contrast, the coefficient estimate is negative and significant when the outcome variable is the proportion of villages having a high school in 1981. Similarly, the coefficient estimate for middle schools in 1981 is negative, although not statistically significant.

Overall, these results suggest that greater conflict exposure may eventually promote investments in human capital (e.g., literacy), but may actually run counter to more advanced human capital investments.

7.4.3 Health

Appendix Table A.15 takes a variety of measures of local public health as the outcome variables. Pre-colonial conflict exposure significantly predicts lower infant mortality rates in 1991 (column 1). By contrast, there is a negative and significant relationship between pre-colonial conflict exposure and the proportion of villages having a health center or subcenter in 1981 (columns 2-3). When taken in combination with the results from Table 14, the significant result for infant mortality appears to reflect local investments in basic human capital, rather than improvements in physical health infrastructure.

7.4.4 Transportation Infrastructure

Finally, we take a variety of local transportation infrastructure measures as our dependent variables. One historical measure of transportation infrastructure that is systematically available is the year in which the first railroad connection was made within a district. Column 1 of Appendix Table A.16 indicates that there is a positive and significant relationship between pre-colonial conflict exposure and early railroad development. Appendix Figure A.11 shows that the first wave of colonial railroad construction in the late 1850s and early 1860s drives this result. By contrast, we find a negative and significant relationship between pre-colonial conflict exposure and our post-colonial measure of railroad access (namely, the average proportion of villages within a district having access to a railroad between 1981 and 1991) in column 2 of Appendix Table A.16. Given that the number of sample districts for which railroad data are available differs between the colonial and post-colonial eras, we interpret the differences between these results with caution. Nevertheless, in combination, they suggest that there was a role reversal in railroad access for districts that experienced greater pre-colonial conflict exposure. Column 3 indicates that there is a positive, although not statistically significant relationship, between pre-colonial conflict exposure and post-colonial road access (i.e., the average proportion of villages within a district having access to a paved road between 1981 and 1991). Finally, pre-colonial conflict exposure predicts significantly lower average canal access between 1981 and 1991 (column 4).

7.5 Section Summary

In terms of our theoretical framework, the results in this section suggest that the positive relationship between pre-colonial conflict exposure and current economic development in India runs through the following channels: (1) early state-making and fiscal development; (2) greater political stability in the long term; (3) greater investments in irrigation and related improvements in agricultural productivity in the long term; and (4) greater investments in literacy and primary education in the long term. Following our theoretical framework, we view local investments in agriculture and human capital as functions (at least in part) of more powerful local government institutions and greater political stability.

8 Conclusion

We have analyzed the role of pre-colonial history – and in particular the role of warfare – in long-run development outcomes across India. We have argued that, if a given district in India experienced more pre-colonial warfare, then more powerful local government institutions were likely to emerge there, which in turn helped promote local long-run economic development through both greater political stability and the provision of basic public goods.

To evaluate the predictions of this argument, we have exploited a new, geocoded database of historical conflicts on the Indian subcontinent. We have shown evidence for a positive, significant, and robust relationship between pre-colonial conflict exposure and local economic development in India today. Consistent with our theoretical framework, we have found that early local state-making and the development of local fiscal capacity, more political stability, and investment in basic public goods help explain this relationship.

Our study shows that the "war makes states" framework applies beyond the paradigmatic case of Western Europe. This parallel between Western Europe and India makes sense, given that two key historical factors in the European context – namely, enduring political fragmentation and interstate military competition – were also important features of the pre-colonial Indian landscape. In Imperial China, by contrast, there was political centralization and internal warfare (e.g., mass rebellion). This dynamic altered the consequences of violent conflict for institutional reforms (Dincecco and Wang, 2018). Furthermore, unlike in pre-colonial Africa, historical population density in pre-colonial India was high enough – as in Western Europe – to make territorial acquisition through warfare worthwhile (Herbst, 2000, 13-16). Low population density meant that a traditional goal of African warfare was to capture slaves (Herbst, 2000, 20), weakening the relationship between warfare and state-making.[25] Recently, scholars have shown negative correlations between pre-colonial conflict levels and long-run development outcomes in Africa (Besley and Reynal-Querol, 2014; Dincecco, Fenske and Onorato, 2019). Finally, in contrast to pre-twentieth-century states in Latin America, those in pre-colonial India were able to establish early political authority, and relied on domestic extraction (e.g., versus external debt) to fund their military efforts (Centeno, 2002, 165-6; Queralt, 2019). Institutional reforms in response to warfare were more likely under such circumstances. Overall, our study helps to clarify the conditions under which the "war makes states" logic is most likely to hold.

[25] Osafo-Kwaako and Robinson (2013), for example, do not find any significant correlation between warfare and state centralization in pre-colonial Africa.

References

Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101(5):2003–41.

Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4):1190–1213.

Baness, J. Frederick. 1881. *Geographicus Indicus*. Edward Stanford.

Bentzen, Jeanet, Nicolai Kaarsen and Asger Wingender. 2017. "Irrigation and Autocracy." *Journal of the European Economic Association* 15(1):1–53.

Besley, Timothy and Marta Reynal-Querol. 2014. "The Legacy of Historical Conflict: Evidence from Africa." *American Political Science Review* 108(2):319–336.

Besley, Timothy and Torsten Persson. 2008. "Wars and State Capacity." *Journal of the European Economic Association* 6(2-3):522–30.

Besley, Timothy and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.

Bharadwaj, Prashant and Rinchan Ali Mirza. 2017. "Displacement and Development: Partition of India and Agricultural Development." *Working Paper, University of Namur* .

Bockstette, Valerie, Areendam Chanda and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth* 7(4):347–69.

Brecke, Peter. 1999. "Violent Conflicts 1400 A.D. to the Present in Different Regions of the World." *Working Paper, Meeting of Peace Science Society Annual Meeting* .

Brewer, John. 1989. *The Sinews of Power: War, Money, and the English State, 1688-1783*. Harvard University Press.

Cameron, A. Colin, Jonah Gelbach and Douglas Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3):414–427.

Cassidy, Traviss, Mark Dincecco and Massimiliano Onorato. 2017. "The Economic Legacy of Conflict: Evidence from European Regions." *Working Paper, University of Michigan* .

Castelló-Climent, Amparo, Latika Chaudhary and Abhiroop Mukhopadhyay. 2017. "Higher Education and Prosperity: From Catholic Missionaries to Luminosity in India." *Economic Journal* .

Centeno, Miguel. 2002. *Blood and Debt: War and the Nation-State in Latin America*. Pennsylvania State University Press.

Chakrabarti, Jadab. 1896. *The Native States of India*. Luzac.

Chandler, Tertius. 1987. *Four Thousand Years of Urban Growth: A Historical Census*. St. David's University Press.

Chaudhary, Latika, Jared Rubin, Sriya Iyer and Anand Shrivastava. 2018. "Cultural Transmission and the Colonial Legacy: Evidence from Public Good Games Along a Historical Border." *Working Paper, Naval Postgraduate School* .

Clodfelter, Michael. 2002. *Warfare and Armed Conflicts: A Statistical Reference to Casualty and Other Figures, 1500-2000*. McFarland.

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

de la Garza, Andrew. 2016. *The Mughal Empire at War*. Routledge.

Dell, Melissa, Nathan Lane and Pablo Querubin. 2018. "The Historical State, Local Collective Action, and Economic Development in Vietnam." *Forthcoming in Econometrica* .

Dincecco, Mark. 2017. *State Capacity and Economic Development: Present and Past*. Cambridge University Press.

Dincecco, Mark, James Fenske and Massimiliano Onorato. 2019. "Is Africa Different? Historical Conflict and State Development." *Economic History of Developing Regions* .

Dincecco, Mark and Yuhua Wang. 2018. "Violent Conflict and Political Development over the Long Run: China versus Europe." *Annual Review of Political Science* 21:341–358.

Docherty, Paddy. 2008. *The Khyber Pass: A History of Empire and Invasion*. Union Square Press.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108(4-5):899–934.

Dutt, Romesh. 1950. *The Economic History of India under Early British Rule*. Routledge.

Fenske, James and Namrata Kala. 2017. "Linguistic Distance and Market Integration in India." *Centre for Competitive Advantage in the Global Economy Working Paper 331* .

Fenske, James and Namrata Kala. 2018. "Railroads and Cities in India." *Working Paper, University of Warwick* .

Foa, Roberto Stefan. 2016. *Ancient Polities, Modern States*. PhD Dissertation, Harvard University.

Gaikwad, Nikhar. 2014. "East India Companies and Long-Term Economic Change in India." *Working Paper, Columbia University* .

Gennaioli, Nicola and Hans-Joachim Voth. 2015. "State Capacity and Military Conflict." *Review of Economic Studies* 82(4):1409–1448.

Gennaioli, Nicola and Ilia Rainer. 2007. "The modern impact of precolonial centralization in Africa." *Journal of Economic Growth* 12(3):185–234.

Gerring, John, Daniel Ziblatt, Johan Van Gorp and Julián Arévalo. 2011. "An Institutional Theory of Direct and Indirect Rule." *World Politics* 63(3):377–433.

Glaeser, Edward L. and Jesse M. Shapiro. 2002. "Cities and Warfare: The Impact of Terrorism on Urban Form." *Journal of Urban Economics* 51(2):205 – 224.

Gommans, Jos. 1999. Warhorse and Gunpowder in India, c. 1000-1850. In *War in the Early Modern World*, ed. Jeremy Black. Routledge chapter 5, pp. 105–28.

Gommans, Jos. 2003. *Mughal Warfare: Indian Frontiers and Highroads to Empire 1500-1700*. Routledge.

Gupta, Bishnupriya. 2018. "Falling Behind and Catching up: India's Transition from a Colonial Economy." *Working Paper, University of Warwick* .

Gupta, Bishnupriya, Debin Ma and Tirthankar Roy. 2016. States and Development: Early Modern India, China, and the Great Divergence. In *Economic History of Warfare and State Formation*, ed. Jari Eloranta, Eric Golson, Andrei Markevich and Nikolaus Wolf. Springer chapter 2, pp. 51–69.

Hariri, Jacob. 2012. "The Autocratic Legacy of Early Statehood." *American Political Science Review* 106(3):471–494.

Heldring, Leander. 2018. "The Origins of Violence in Rwanda." *Working Paper, University of Bonn* .

Henderson, Vernon, Adam Storeygard and David Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102(2):994–1028.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.

Himanshu. 2009. "Electoral Politics and the Manipulation of Statistics." *Economic and Political Weekly* 44(19):31–35.

Iyer, Lakshmi. 2010. "Direct versus Indirect Colonial Rule in India: Long-Term Consequences." *Review of Economics and Statistics* 92(4):693–713.

Iyer, Sriya, Anand Shrivastava and Rohit Ticku. 2017. "Holy Wars? Temple Desecrations in Medieval India." *Working Paper, University of Cambridge* .

James, Lawrence. 1997. *Raj: The Making and Unmaking of British India*. Little, Brown, and Co.

Jaques, Tony. 2007. *Dictionary of Battles and Sieges: A Guide to 8,500 Battles from Antiquity through the Twenty-First Century*. Greenwod Press.

Jha, Saumitra. 2013. "Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia." *American Political Science Review* 107(4):806–32.

Kelly, Morgan. 2019. "The Standard Errors of Persistence." *Centre for Economic Policy Research Discussion Paper 13783* .

Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs and Jeffrey Sachs. 2004. "A Global Index Representing the Stability of Malaria Transmission." *The American Journal of Tropical Medicine and Hygiene* 70(5):486–498.

König, Michael, Dominic Rohner, Mathias Thoenig and Fabrizio Zilibotti. 2017. "Networks in Conflict: Theory and Evidence From the Great War of Africa." *Econometrica* 85(4):1093–132.

Lange, Matthew. 2004. "British Colonial Legacies and Political Development." *World Development* 32(6):905–22.

Lee, Alexander. 2018. "Land, State Capacity, and Colonialism: Evidence From India." *Comparative Political Studies* pp. 1–33.

Mann, Michael. 1984. "The Autonomous Power of the State: Its Origins, Mechanisms and Results." *European Journal of Sociology* 25(2):185–213.

Matsuura, Kenji and Cort Willmott. 2009. "Terrestrial Air Temperature: 1900-2008 Gridded Monthly Time Series." *Center for Climatic Research, University of Delaware* .

Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1):113–152.

Min, Brian. 2015. *Power and the Vote: Elections and Electricity in the Developing World*. Cambridge University Press.

Montalvo, José and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review* 95(3):796–816.

Morris, Ian. 2014. *War! What Is It Good For?: Conflict and the Progress of Civilization from Primates to Robots*. Princeton University Press.

Mukherjee, Shivaji. 2017. "Colonial Origins of Maoist Insurgency in India: Historical Institutions and Civil War." *Journal of Conflict Resolution* .

Nag, Prithvish. 2007. *Historical Atlas of India*. NATMO.

Naravane, M.S. 1997. *Battles of Medieval India: AD 1295-1850*. APH Publishing.

Nath, Pratyay. 2018. "Through the Lens of War: Akbar's Sieges (1567-69) and Mughal Empire-Building in Early Modern North India." *Journal of South Asian Studies* 41(2):245–258.

North, Douglass. 1981. *Structure and Change in Economic History*. W.W. Norton.

Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics* 94(1):20–36.

Osafo-Kwaako, Philip and James Robinson. 2013. "Political Centralization in Pre-Colonial Africa." *Journal of Comparative Economics* 41(1):534–564.

Özak, Ömer. 2010. "The Voyage of Homo-Economicus: Some Economic Measures of Distance." *Working Paper, Brown University* .

Özak, Ömer. 2018. "Distance to the Pre-Industrial Technological Frontier and Economic Development." *Journal of Economic Growth* 23(2):175–221.

Qaisar, Ahsan. 1998. Distribution of the Revenue Resources of the Mughal Empire among the Nobility. In *The Mughal State, 1526-1750*, ed. Alam Muzaffar and Sanjay Subrahmanyam. Oxford University Press chapter 7, pp. 252–258.

Queralt, Didac. 2019. "The Legacy of War on Fiscal Capacity." *International Organization* .

Ramankutty, Navin, Jonathan Foley, John Norman and Kevin McSweeney. 2002. "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change." *Global Ecology and Biogeography* 11(5):377–392.

Ramusack, Barbara. 2003. *The Indian Princes and their States*. Cambridge University Press.

Rasler, Karen and William Thompson. 2005. "War Making and State Making: Governmental Expenditures, Tax Revenues, and Global Wars." *American Political Science Review* 79(2):491–507.

Raychaudhuri, Tapan. 1982. Inland Trade. In *The Cambridge Economic History of India, Volume 1: c.1200-c.1750*, ed. Tapan Raychaudhuri and Irfan Habib. Cambridge: Cambridge University Press chapter 11, pp. 325–359.

Richards, John. 1995. *The Mughal Empire*. Second ed. Cambridge University Press.

Roy, Madhabi. 1994. *Politics, War, and State Formation in Early Modern India*. PhD Dissertation, Harvard University.

Roy, Tirthankar. 2013. *An Economic History of Early Modern India*. Routledge.

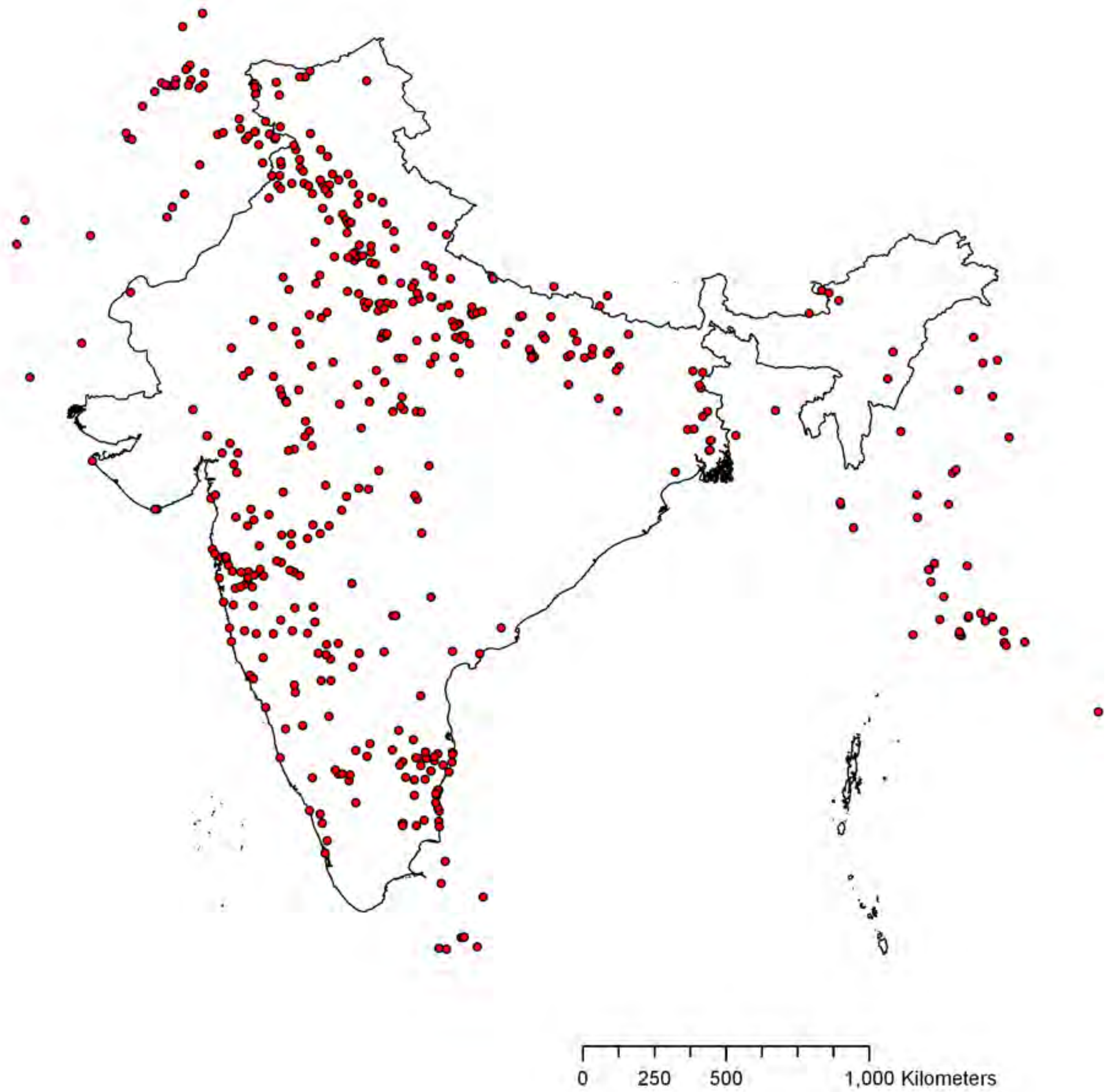
Schwartzberg, Joseph. 1978. *A Historical Atlas of South Asia*. University of Chicago Press.

Stein, Burton. 1985. "State Formation and Economy Reconsidered." *Modern Asian Studies* 19(3):387–413.

Tilly, Charles. 1975. Reflections on the History of European State-Making. In *The Formation of States in Western Europe*, ed. Charles Tilly. Princeton University Press chapter 1, pp. 3–83.

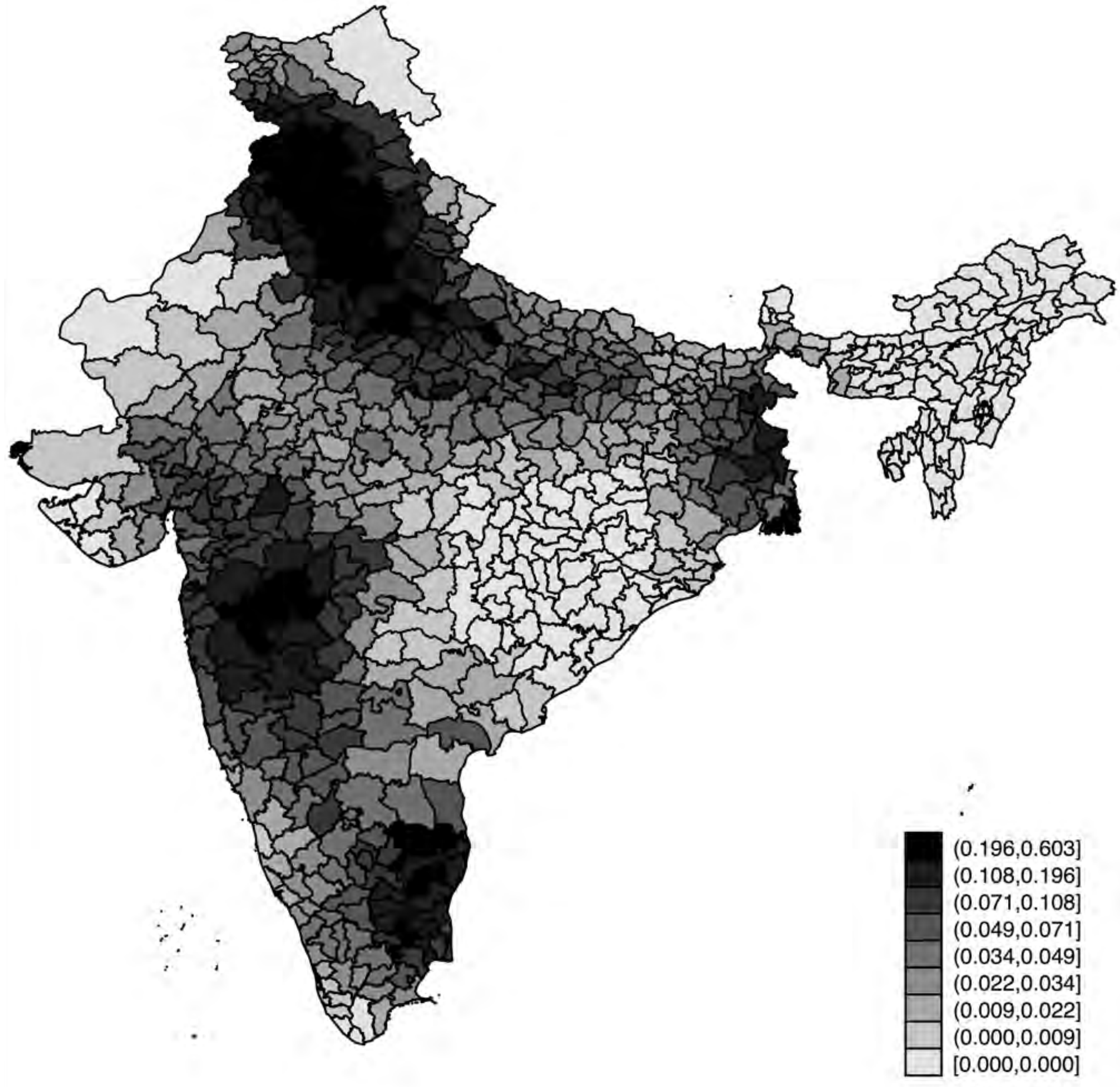
Tollefsen, Andreas, Håvard Strand and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49(2):363–74.

Figure 1: Conflict Locations, 1000-2010



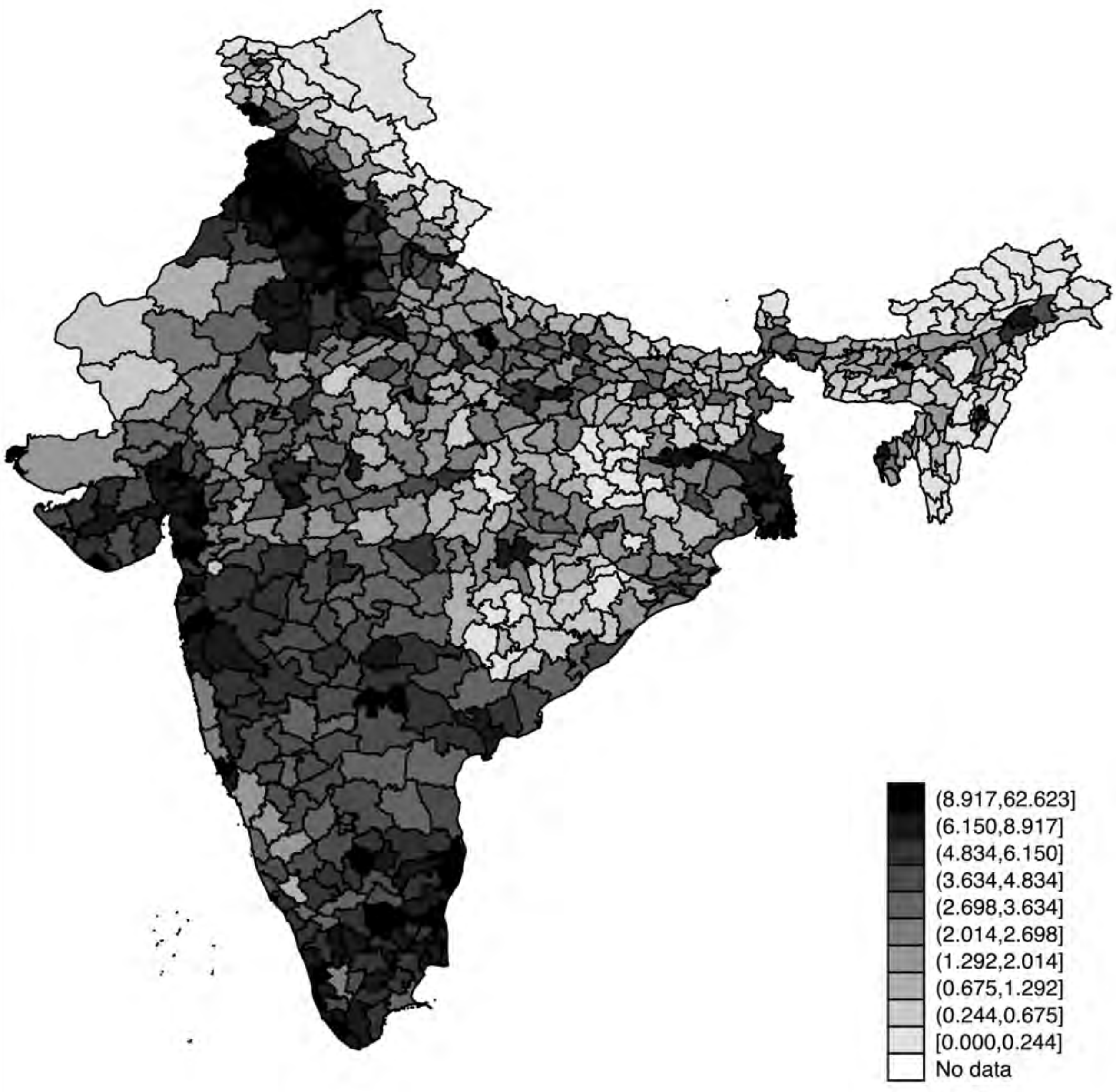
Notes. This figure shows the location of each recorded military conflict on the Indian subcontinent between 1000-2010.

Figure 2: Pre-Colonial Conflict Exposure by Indian Districts



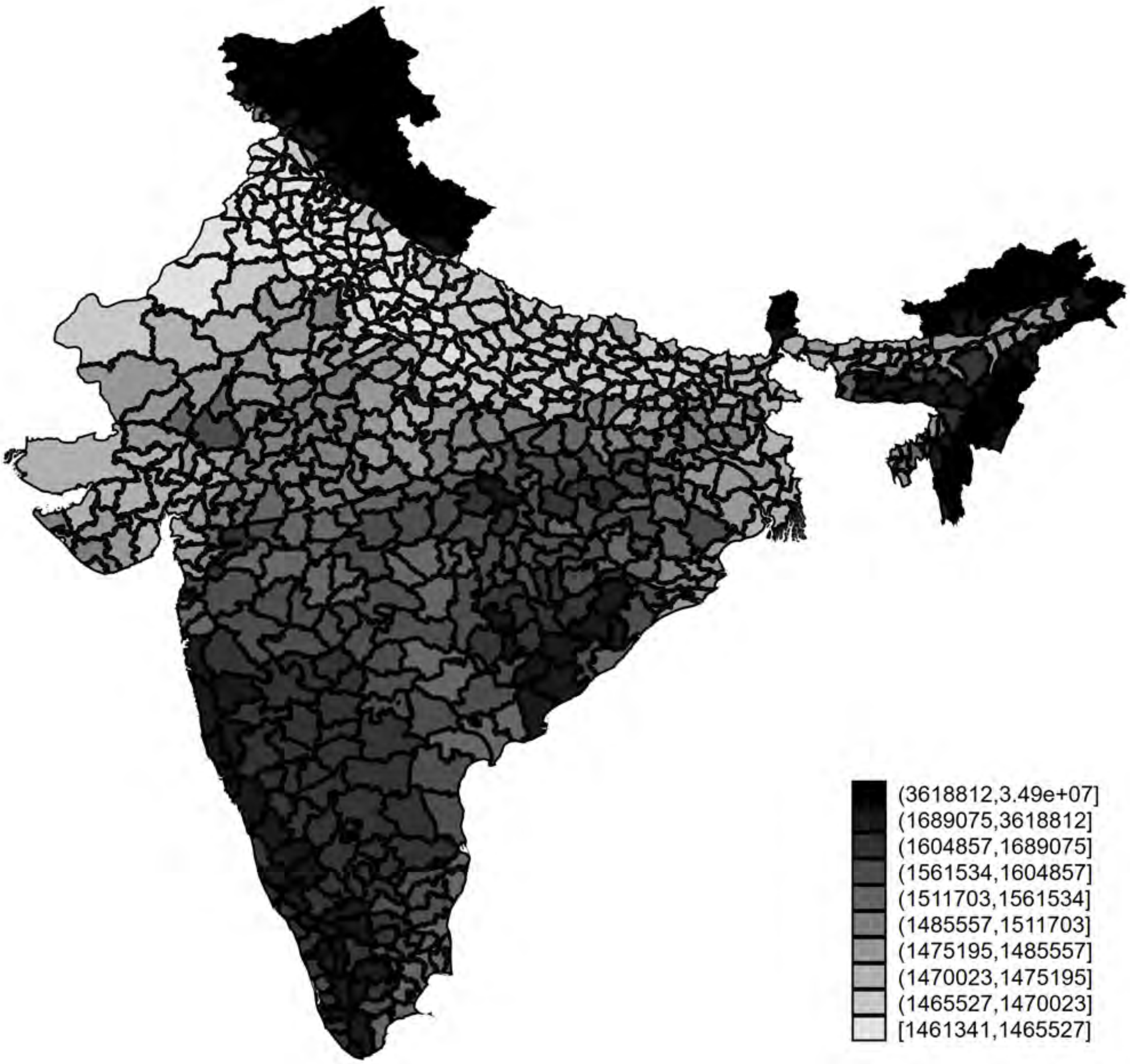
Notes. This figure shows pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers by district in India. Districts are shaded by decile: districts in the top decile receive the darkest shade.

Figure 3: Luminosity by Indian Districts



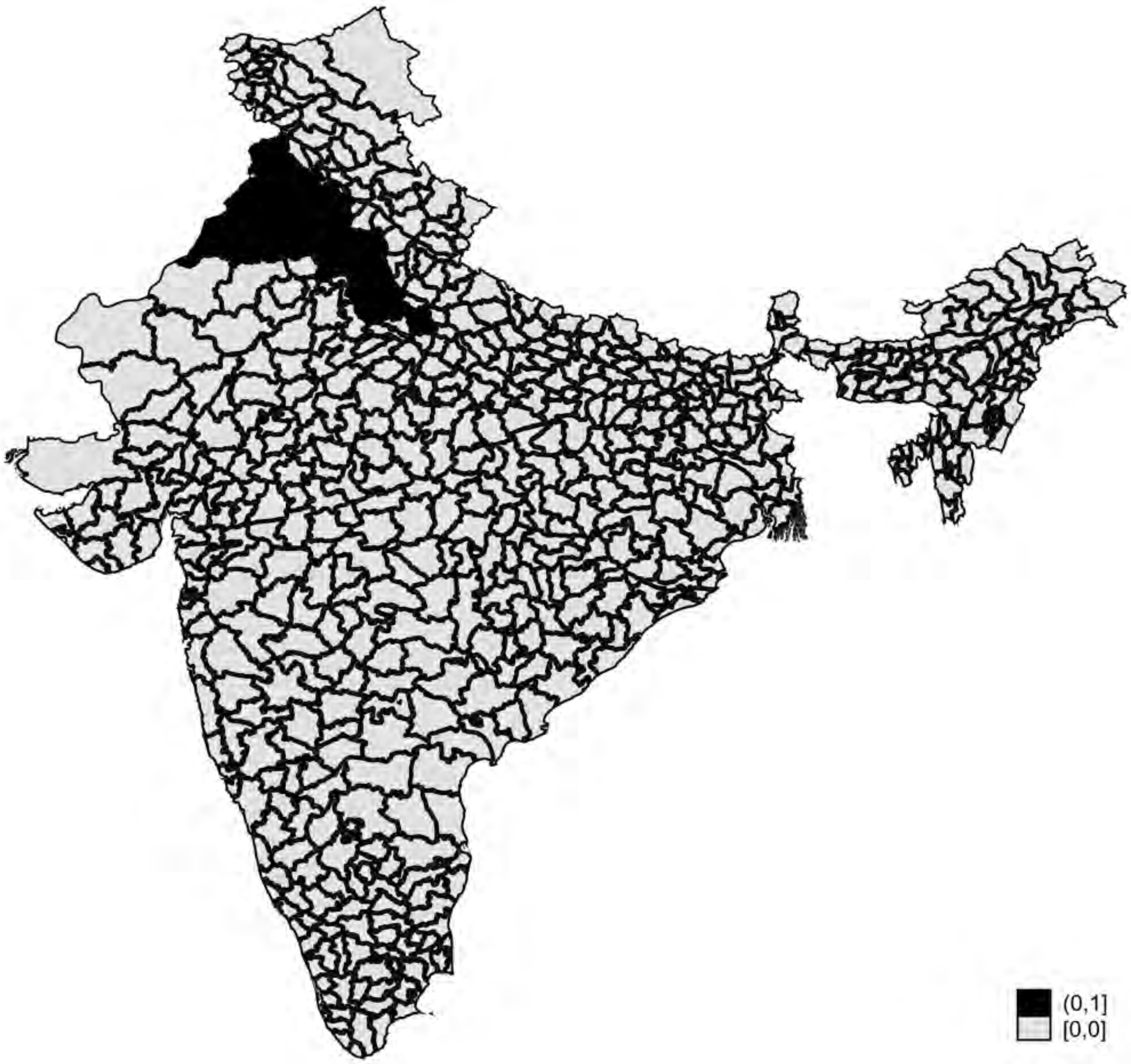
Notes. This figure shows average luminosity between 1992-2010 by district in India. Districts are shaded by decile: districts in the top decile receive the darkest shade.

Figure 4: Cost Distance from Khyber Pass by Indian Districts



Notes. This figure shows the average cost distance of each district in India from the Khyber Pass, where we assume that the cost of crossing a grid cell is proportional to its squared ruggedness. Districts are shaded by decile: districts in the top decile receive the darkest shade.

Figure 5: Khyber Proximity IV



Notes. This figure shows the values of the Khyber Proximity instrument by district in India.

Table 1: Pre-Colonial Conflict and Economic Development: Main Results

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	3.713***	1.601***	1.465***
	(0.305)	(0.380)	(0.370)
	[0.000]	[0.000]	[0.000]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.240	0.104	0.095
R^2	0.598	0.829	0.849
Observations	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 2: Pre-Colonial Conflict and Economic Development: Instrumental Variables

Panel A: First Stage

Dependent variable:	Pre-Colonial Conflict Exposure		
	(1)	(2)	(3)
Cost distance to Kyhber Pass	0.206***	0.097***	0.080***
	(0.018)	(0.025)	(0.024)
	[0.000]	[0.000]	[0.001]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
R^2	0.429	0.649	0.669
Observations	660	660	660

Panel B: Second Stage

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	4.930***	4.626***	3.482**
	(0.609)	(1.291)	(1.389)
	[0.000]	[0.000]	[0.012]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Anderson-Rubin p-value	0.000	0.000	0.012
Kleibergen-Paap Wald rk F-statistic	131.275	14.444	10.693
Observations	660	660	660

Notes. Estimation method is 2SLS. Unit of analysis is district. In Panel A (first stage), dependent variable is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers, while variable of interest is cost distance to Khyber Pass (computed as squared ruggedness). In Panel B (second stage), dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010, while variable of interest is pre-colonial conflict exposure between 1000-1757 with a cutoff distance of 250 kilometers, as instrumented by cost distance to Khyber Pass. Geographic controls for both first and second stages include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1750 for first stage, and in 1990 for second stage. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 3: Pre-Colonial Conflict and Economic Development: Colonial Controls

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	1.263***	1.265***	0.922**	0.951**
	(0.377)	(0.379)	(0.406)	(0.417)
	[0.001]	[0.001]	[0.025]	[0.024]
Direct rule		-0.085		
		(0.071)		
		[0.230]		
%Non-landlord				-0.052
				(0.111)
				[0.640]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.091	0.091	0.102	0.105
R^2	0.817	0.817	0.856	0.856
Observations	634	634	166	166

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. $DirectRule$ is a dummy variable that equals 1 for direct British rule. $\%NonLandlord$ measures the proportion of a district under a non-landlord revenue system in British India. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 4: Pre-Colonial Conflict and Economic Development: Fractionalization Controls

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	1.455***	1.437***	1.462***	1.422***
	(0.373)	(0.369)	(0.371)	(0.372)
	[0.000]	[0.000]	[0.000]	[0.000]
Medieval port	-0.032			
	(0.089)			
	[0.721]			
Linguistic fractionalization		-0.134		-0.157
		(0.138)		(0.139)
		[0.332]		[0.257]
Religious fractionalization			0.038	0.122
			(0.257)	(0.261)
			[0.883]	[0.641]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.094	0.093	0.095	0.092
R^2	0.849	0.849	0.849	0.849
Observations	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. $MedievalPort$ is a dummy variable that equals 1 for the presence of a major medieval port. $LinguisticFractionalization$ is 1 minus the Herfindahl index of language population shares in 2001. $ReligiousFractionalization$ is 1 minus the Herfindahl index of religion population shares in 2001. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 5: Pre-Colonial Conflict and Economic Development: Post-1757 Conflict Controls

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	1.483***	1.492***	1.461***	1.489***
	(0.393)	(0.375)	(0.389)	(0.418)
	[0.000]	[0.000]	[0.000]	[0.000]
Colonial conflict exposure (1758-1839)	-0.109			0.005
	(0.735)			(0.742)
	[0.882]			[0.995]
Colonial conflict exposure (1840-1946)		-0.679*		-0.679*
		(0.406)		(0.408)
		[0.095]		[0.097]
Post-colonial conflict exposure (1947-2010)			-0.136	-0.055
			(3.139)	(3.151)
			[0.965]	[0.986]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.096	0.096	0.094	0.096
R^2	0.849	0.849	0.849	0.849
Observations	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. All conflict exposure variables measure conflict exposure to land battles with a cutoff distance of 250 kilometers. Variable of interest is pre-colonial conflict exposure between 1000-1757. The first colonial conflict exposure variable spans 1758-1839, while the second spans 1840-1946. The post-colonial conflict exposure variable spans 1947-2010. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 6: Pre-Colonial Conflict and Economic Development: Initial State Capacity Controls

Dependent variable:	Ln(0.01+Luminosity)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-colonial conflict exposure	1.480***	1.458***	1.406***	1.464***	1.467***	1.531***	1.402***
	(0.371)	(0.369)	(0.379)	(0.370)	(0.370)	(0.419)	(0.368)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Neolithic settlements	Yes	No	No	No	No	No	No
Chalcolithic settlments	No	Yes	No	No	No	No	No
Cultural sites (300-700 CE)	No	No	Yes	No	No	No	No
Cultural sites (8th-12th centuries)	No	No	No	Yes	No	No	No
Ln(1+Urban population in 1000)	No	No	No	No	Yes	No	No
Major Indian states (10th-11th centuries)	No	No	No	No	No	Yes	No
Major Indian states (11th-12th centuries)	No	No	No	No	No	No	Yes
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.096	0.094	0.091	0.095	0.095	0.099	0.091
R^2	0.849	0.849	0.849	0.849	0.849	0.851	0.852
Observations	660	660	660	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. $Neolithic$ and $Chalcolithic$ control for the number of Neolithic and Chalcolithic settlements. $CulturalSite$ controls for the number of cultural sites between 300-700 CE and the eighth-twelfth centuries. $UrbanPop$ controls for the natural logarithm of (one plus) the total urban population in 1000 CE. $MajorState$ controls for the presence of a major Indian state between the tenth-eleventh centuries and the eleventh-twelfth centuries. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 7: Pre-Colonial Conflict and Early State-Making

Dependent variable:	Important Mughal Sites	State History		
		Babur	Akbar	Aurangzeb
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure (benchmark)	0.954*			
	(0.497)			
	[0.056]			
Pre-colonial conflict exposure (1000-1526)		0.513**		
		(0.229)		
		[0.025]		
Pre-colonial conflict exposure (1000-1556)			0.723***	
			(0.262)	
			[0.006]	
Pre-colonial conflict exposure (1000-1658)				-0.080
				(0.173)
				[0.642]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.199	0.041	0.068	-0.012
R^2	0.122	0.768	0.715	0.718
Observations	659	659	659	659

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable in column 1 is number of important Mughal-era sites. Dependent variables in columns 2-4 are state longevity in terms of districts incorporated into the Mughal Empire by Babur (1526-30), Akbar (1556-1605), and Aurangzeb (1658-1707). Variable of interest is pre-colonial conflict exposure to land battles with a cutoff distance of 250 kilometers. It spans 1000-1757 in column 1, 1000-1526 in column 2, 1000-1556 in column 3, and 1000-1658 in column 4. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1500. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 8: Pre-Colonial Conflict and Colonial Institutions

Dependent variable:	Direct Rule	%Non-Landlord	1881						1931	
			All		British		Princely			
			Ln(Tax/Area)	Ln(Tax/Person)	Ln(Tax/Area)	Ln(Tax/Person)	Ln(Tax/Area)	Ln(Tax/Person)	Ln(Tax/Acre)	Ln(Tax/Person)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pre-colonial conflict exposure	0.017	0.539	2.246***	1.245***	2.208***	1.157***	6.386*	1.612	1.310	0.835
	(0.251)	(0.361)	(0.550)	(0.382)	(0.516)	(0.354)	(3.524)	(2.228)	(0.911)	(0.705)
	[0.946]	[0.138]	[0.000]	[0.001]	[0.000]	[0.001]	[0.076]	[0.473]	[0.153]	[0.238]
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.003	0.120	0.256	0.173	0.281	0.227	0.496	0.210	0.135	0.098
R^2	0.537	0.676	0.468	0.518	0.596	0.545	0.606	0.408	0.731	0.696
Observations	634	166	270	274	200	200	70	74	145	144

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. *DirectRule* is a dummy variable that equals 1 for direct British rule. *%NonLandlord* measures the proportion of a district under a non-landlord revenue system in British India. $ln(Tax/Area)$, 1881 and $ln(Tax/Person)$, 1881 measures land revenue in 1,000 rupees per square kilometer or per capita, in 1881 for districts under direct British rule and/or indirect rule (i.e., major Princely states). $ln(Tax/Acre, 1931)$ and $ln(Tax/Person, 1931)$ measures average land revenue in rupees per acre or per capita, in 1931 for districts in British India. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1750 in columns 1 and 2, 1850 in columns 3 to 8, and 1930 in columns 9 and 10. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 9: Pre-Colonial Conflict versus Colonial and Post-Colonial Conflict

Dependent variable:	Colonial Conflict Exposure (1758-1839)		Colonial Conflict Exposure (1840-1946)		Post-Colonial Conflict Exposure	
	Land Battles	All Conflicts	Land Battles	All Conflicts	Land Battles	All Conflicts
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-colonial conflict exposure	0.170***	0.441***	0.040	0.316	-0.025***	-0.030***
	(0.036)	(0.090)	(0.039)	(0.302)	(0.005)	(0.007)
	[0.000]	[0.000]	[0.308]	[0.295]	[0.000]	[0.000]
Population density	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.350	0.429	0.044	0.206	-0.129	-0.107
R^2	0.568	0.571	0.740	0.562	0.816	0.874
Observations	660	660	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is colonial conflict exposure to land battles between 1758-1839 with a cutoff distance of 250 kilometers in column 1 and to all conflict types in column 2. Similarly, it is colonial conflict exposure between 1840-1946 in columns 3-4 and post-colonial conflict exposure between 1947-2010 in columns 5-6. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 10: Pre-Colonial Conflict and Political Violence

Dependent variable:	Political Violence	Maoist Control	Fractionalization	
			Linguistic	Religious
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	-0.241**	-0.381**	-0.209*	0.080
	(0.102)	(0.163)	(0.113)	(0.071)
	[0.019]	[0.020]	[0.065]	[0.260]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	-0.119	-0.129	-0.073	0.048
R^2	0.408	0.281	0.570	0.557
Observations	660	395	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable in column 1 is *Fatalities*, defined as fatalities per district between 2010-18 (in hundreds). Dependent variable in column 2 is *MaoistControl*, a dummy variable that equals 1 for Maoist control in 2003. Dependent variable in column 3 is *LinguisticFractionalization*, defined as 1 minus the Herfindahl index of language population shares in 2001. Dependent variable in column 4 is *ReligiousFractionalization*, defined as 1 minus the Herfindahl index of religion population shares in 2001. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 11: Pre-Colonial Conflict and Economic Development: Agricultural Investment

Dependent variable:	%Irrigated		Fertilizer	%HYV Area			
	1931	1956-87	1956-87	1956-87			
				All	Rice	Wheat	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-colonial conflict exposure	21.275**	37.413**	7.927	8.241	-14.654	-28.524	25.674
	(10.357)	(15.758)	(12.922)	(8.845)	(10.489)	(25.172)	(30.558)
	[0.041]	[0.018]	[0.540]	[0.352]	[0.164]	[0.258]	[0.402]
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.194	0.173	0.041	0.064	-0.073	-0.069	0.042
R^2	0.391	0.611	0.654	0.600	0.620	0.294	0.563
Observations	257	271	271	271	271	271	271

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. *%Irrigated* measures the proportion of area sown with canal irrigation in 1931 (column 1) and the proportion of gross cropped area that is irrigated averaged between 1956-87 (column 2). *Fertilizer* measures fertilizer use in terms of kilograms per hectare averaged between 1956-87. *%HYVArea* measures the proportion of area sown with high-yield varieties (HYV) of rice, wheat, and other cereals, averaged between 1956-87. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1900 in column 1, and in 1950 in columns 2-7. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 12: Pre-Colonial Conflict and Economic Development: Agricultural Productivity

Dependent variable:	Ln(Yield)			
	Major	Rice	Wheat	Other
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	0.712*	0.137	0.444*	0.727
	(0.370)	(0.213)	(0.233)	(0.454)
	[0.055]	[0.522]	[0.057]	[0.111]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.129	0.029	0.089	0.139
R^2	0.758	0.678	0.737	0.580
Observations	271	266	260	271

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are averaged between 1956-87, and are as follows. $\ln(Yield)$ measures the total yield across 15 major crops, rice, wheat, or other cereals. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls always include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Additionally, column 1 controls for the shares of area under rice, wheat, and other cereals cultivation. Population density is $\ln(PopulationDensity)$ in 1950. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 13: Pre-Colonial Conflict and Economic Development: Literacy

Dependent variable:	%Literacy				
	1881	1921	1961-91	2001	2011
	(1)	(2)	(3)	(4)	(5)
Pre-colonial conflict exposure	-1.933	-5.635	11.796*	15.040***	10.146**
	(3.188)	(3.772)	(6.888)	(5.280)	(4.119)
	[0.545]	[0.136]	[0.088]	[0.005]	[0.014]
Population density	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	-0.047	-0.095	0.112	0.120	0.103
R^2	0.464	0.556	0.623	0.634	0.599
Observations	251	303	271	589	626

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. $\%Literacy$, 1881 is the proportion of "literate" persons in 1881. $\%Literacy$, 1921 is the proportion of persons that can read and write in 1921. $\%Literacy$, 1961-91 is the literacy rate averaged between 1961-91. $\%Literacy$, 2001 and $\%Literacy$, 2011 measure the adult literacy rate across both rural and urban populations for ages 15-plus in 2001 and for ages 7-plus in 2011. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1850 in column 1, 1900 in column 2, 1950 in column 3, and 1990 in columns 4-5. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table 14: Pre-Colonial Conflict and Economic Development: Education

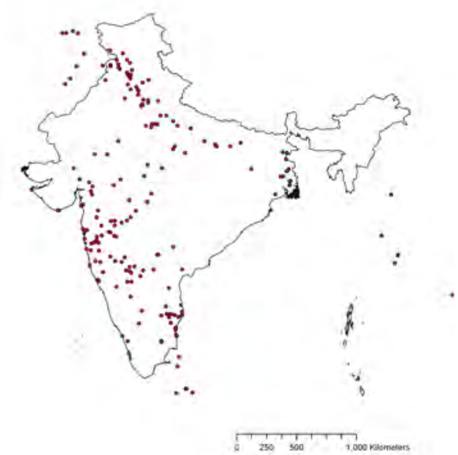
Dependent variable:	1981			2001
	%Primary	%Middle	%High	%School
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	18.683*	-14.559	-16.094**	10.493**
	(11.150)	(9.845)	(6.553)	(4.446)
	[0.096]	[0.141]	[0.015]	[0.019]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.099	-0.085	-0.139	0.077
R^2	0.712	0.786	0.840	0.747
Observations	203	195	187	589

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. $\%Primary$, $\%Middle$, and $\%High$ measure the proportion of villages having a primary, middle, or high school in 1981. $\%School$ measures the proportion of children between the ages of 6-10 that attended school in 2001. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1950 in columns 1-3, and 1990 in column 4. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

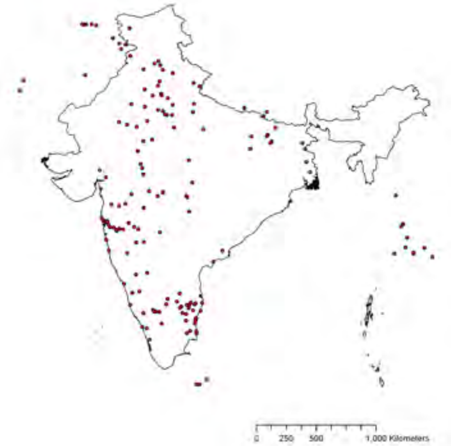
Online Appendix for

Pre-Colonial Warfare and Long-Run Development in India

Figure A.1: Conflict Locations by Sub-Period



(a) 1000-1757



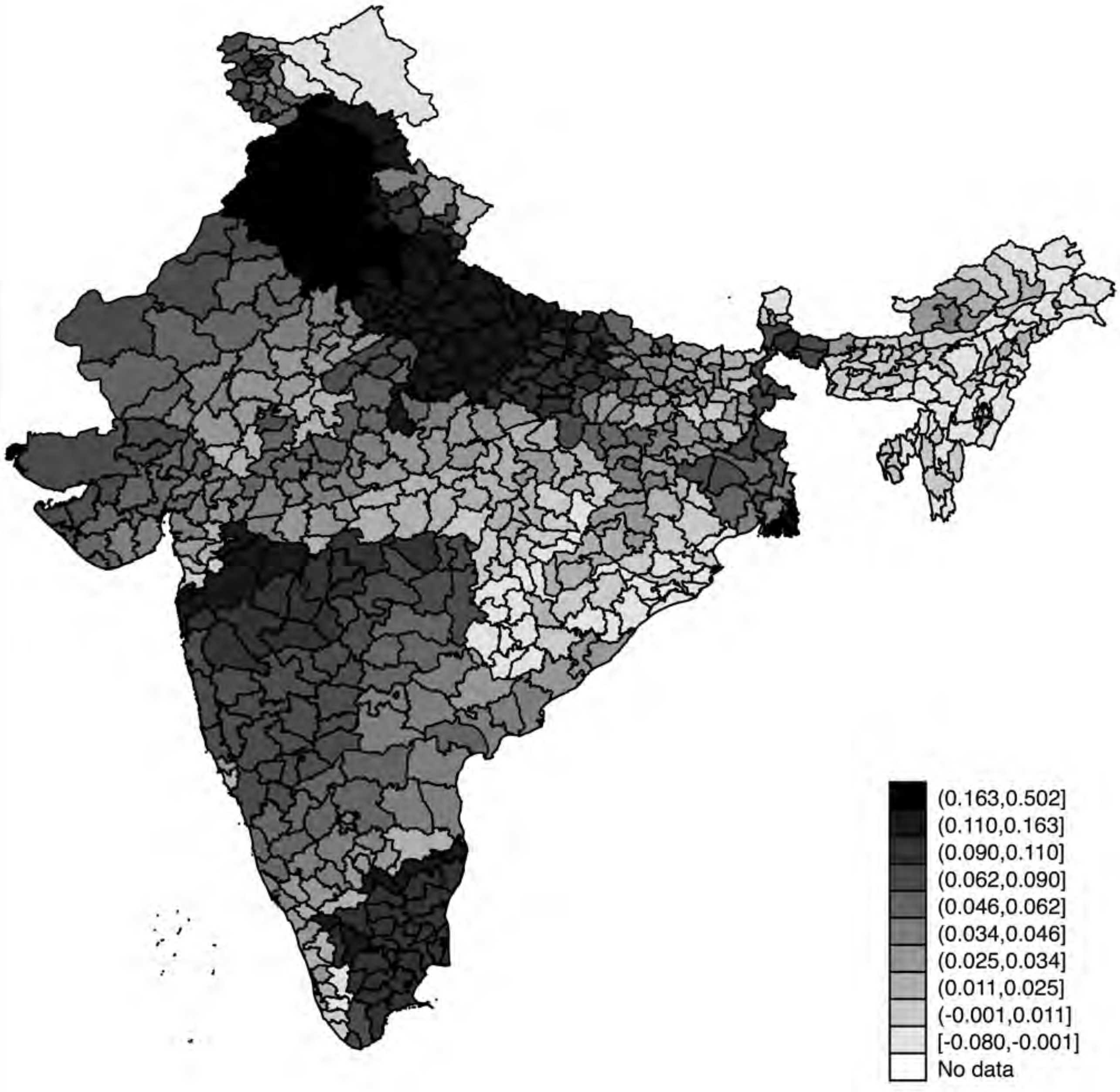
(b) 1758-1839

(c) 1840-1946

(d) 1947-2010

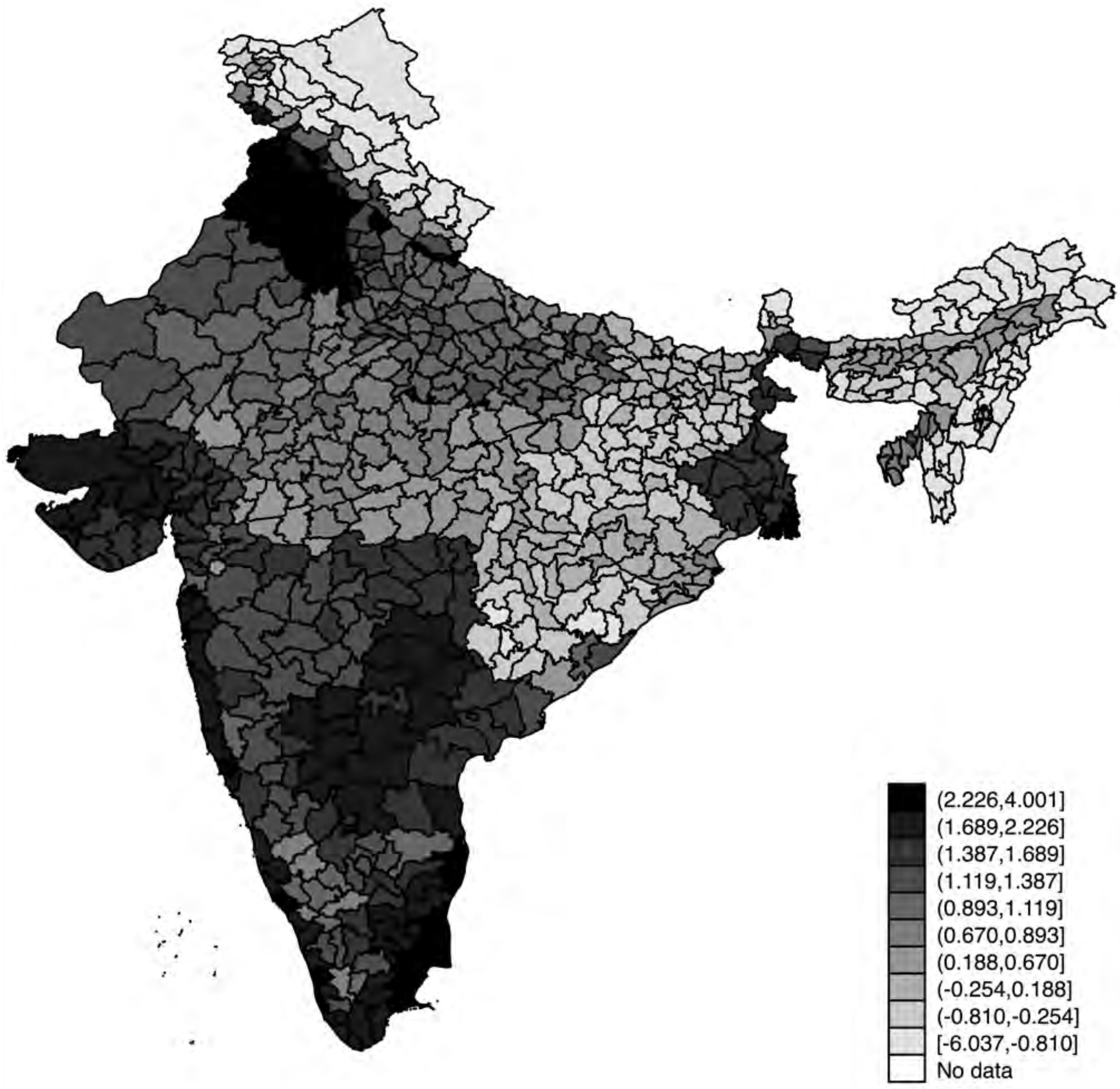
Notes. This figure shows the location of each recorded military conflict on the Indian subcontinent between 1000-2010 by four sub-periods: (a) pre-colonial (1000-1757); (b) colonial (1758-1839); (c) colonial (1840-1946); and (d) post-colonial (1947-2010).

Figure A.2: Residualized Pre-Colonial Conflict Exposure by Indian Districts



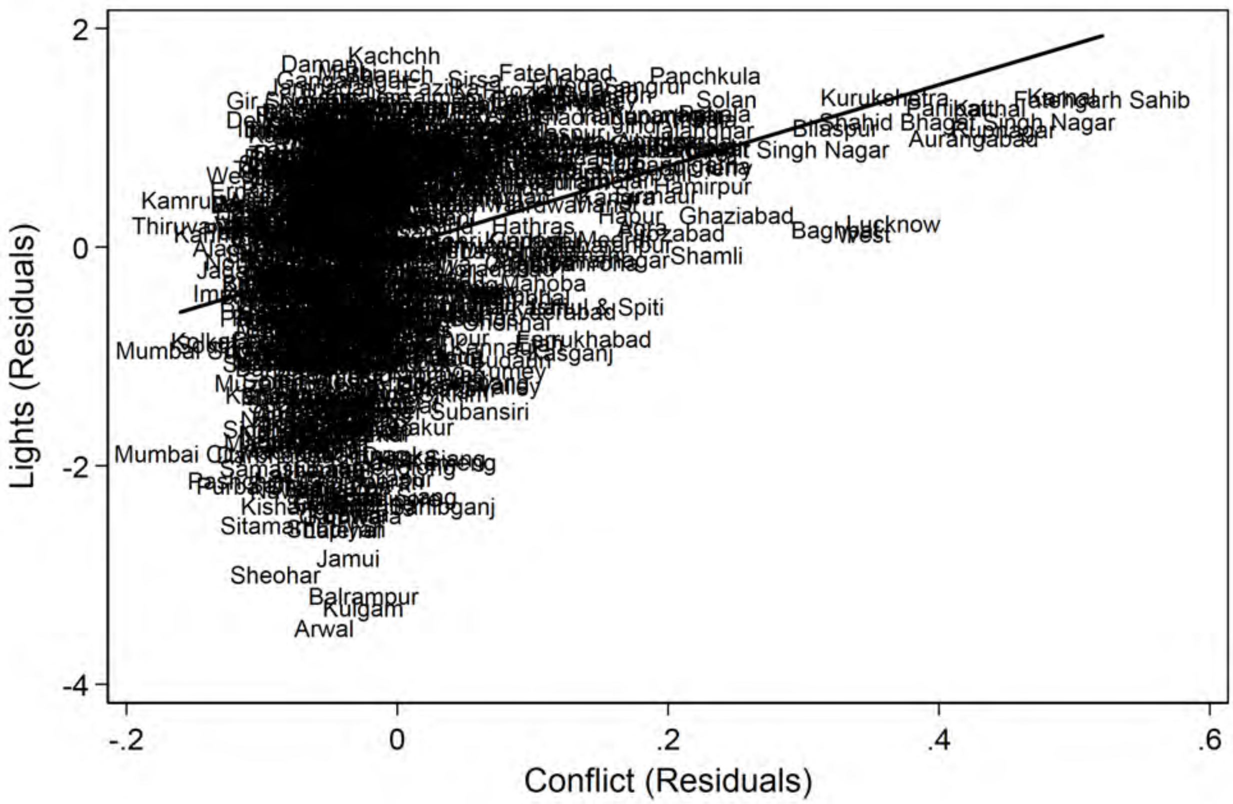
Notes. This figure shows residualized pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers by district in India after controlling for $\ln(PopulationDensity)$ in 1990. Districts are shaded by quintile, whereby districts in the top quintile receive the darkest shade.

Figure A.3: Residualized Luminosity by Indian Districts



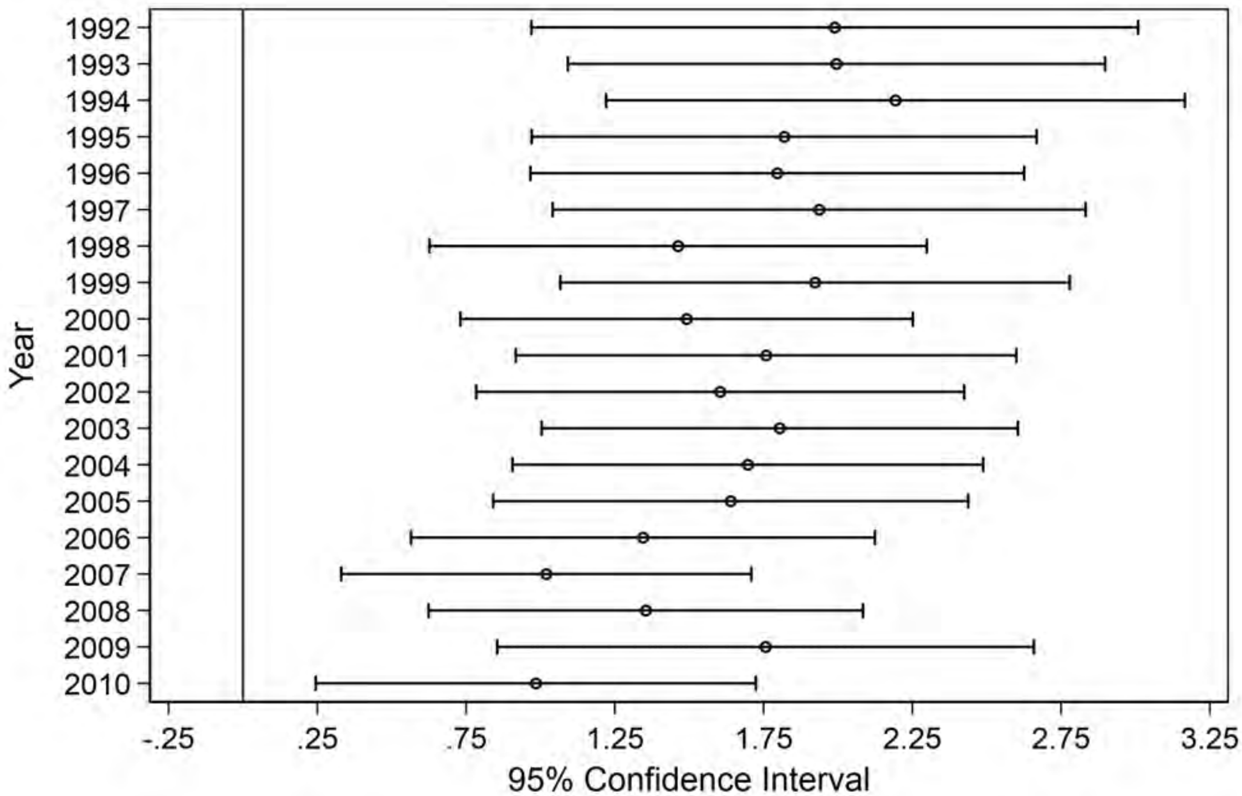
Notes. This figure shows residualized average luminosity between 1992-2010 by district in India after controlling for $\ln(PopulationDensity)$ in 1990. Districts are shaded by quintile, whereby districts in the top quintile receive the darkest shade.

Figure A.4: Pre-Colonial Conflict Exposure and Luminosity by Indian Districts (Residualized)



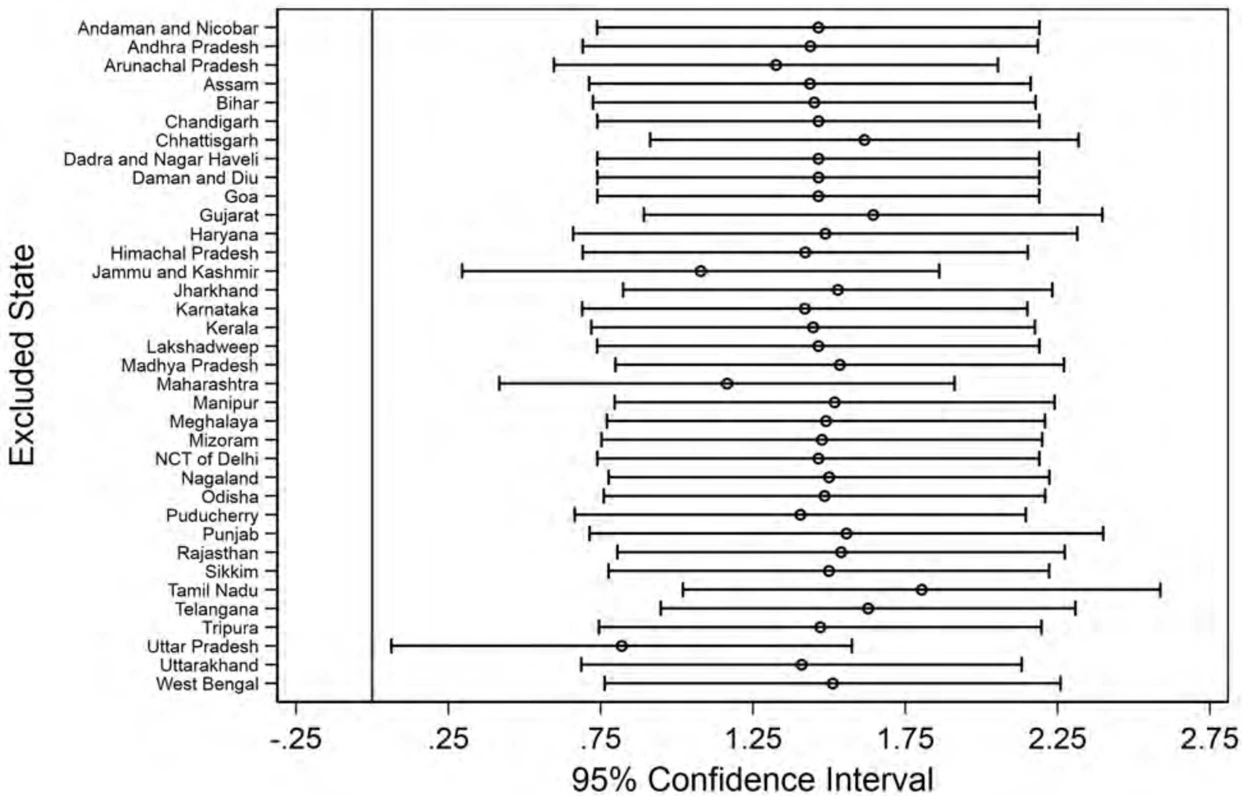
Notes. This figure plots residualized pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers against residualized average luminosity between 1992-2010 by district in India. Both variables are residualized by controlling for $\ln(PopulationDensity)$ in 1990.

Figure A.5: Pre-Colonial Conflict and Economic Development: 95% Confidence Intervals



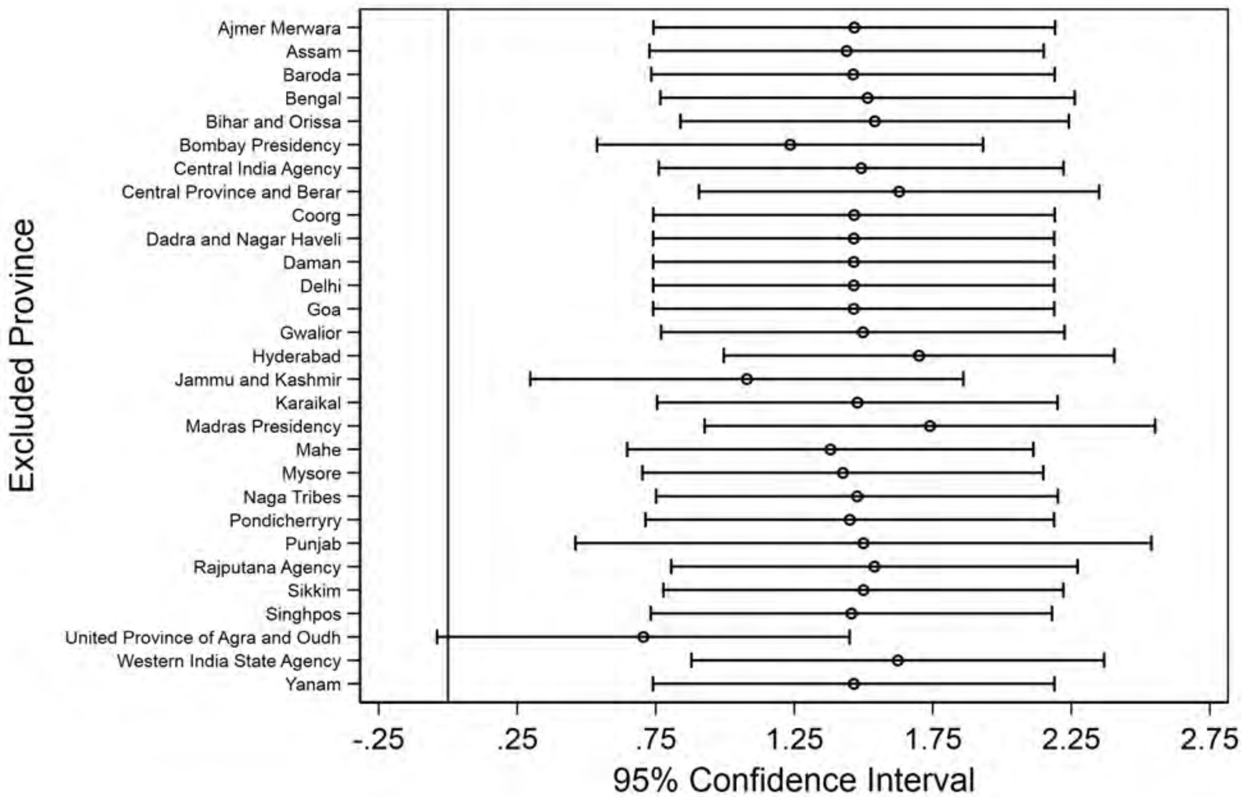
Notes. Each hollow dot represents the point estimate for the regression model in column 3 of Table 1 when the dependent variable is $\ln(0.01 + \mathit{Luminosity})$ for each year between 1992-2010. Horizontal bars indicate 95% confidence intervals.

Figure A.6: Pre-Colonial Conflict and Economic Development: Exclude States One by One



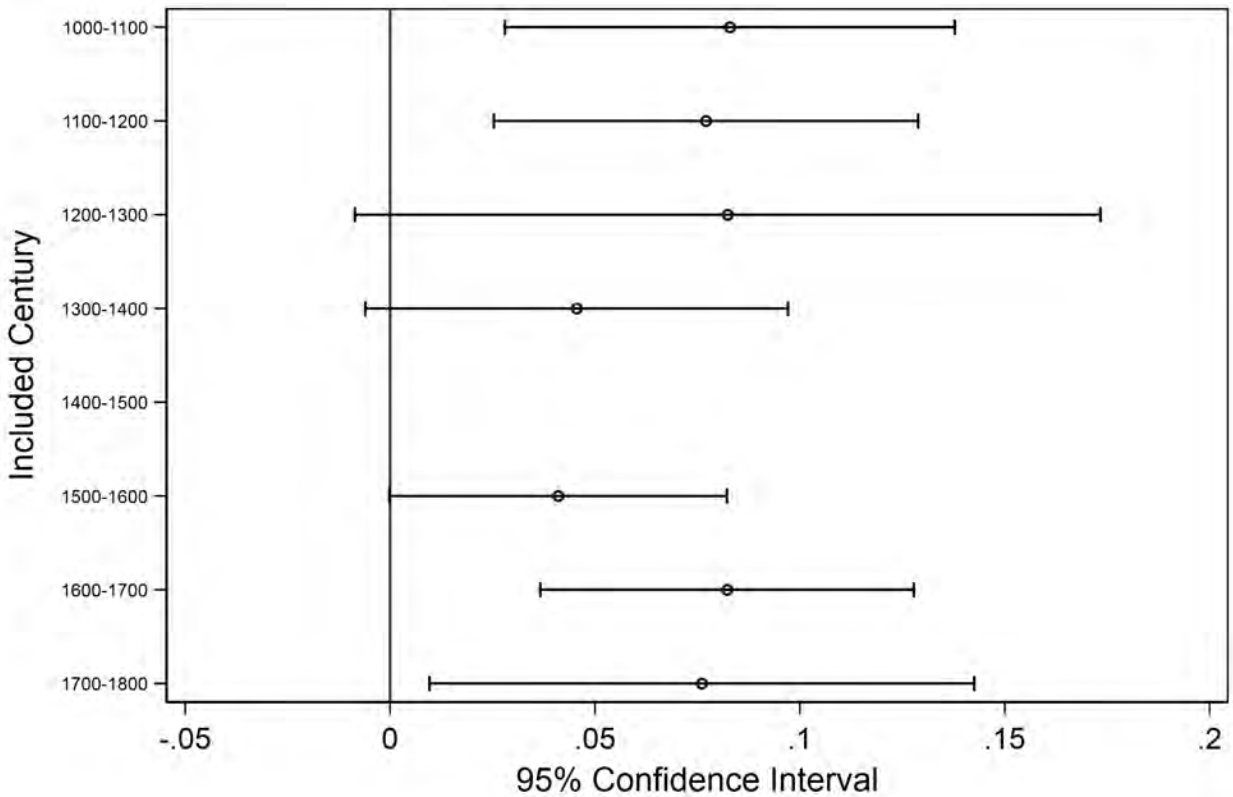
Notes. Each hollow dot represents the point estimate for the regression model in column 3 of Table 1 when we exclude each state or union territory one by one. Horizontal bars indicate 95% confidence intervals.

Figure A.7: Pre-Colonial Conflict and Economic Development: Exclude Colonial Provinces



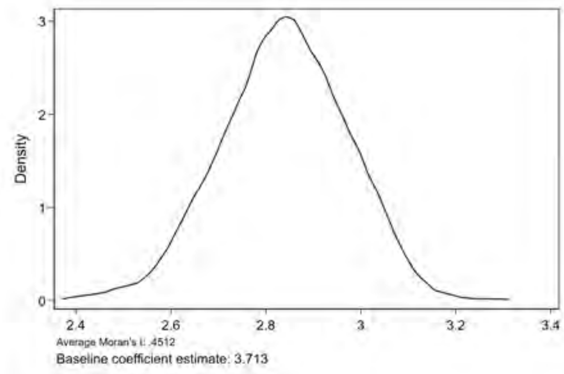
Notes. Each hollow dot represents the point estimate for the regression model in column 3 of Table 1 when we exclude each colonial province one by one. Horizontal bars indicate 95% confidence intervals.

Figure A.8: Pre-Colonial Conflict and Economic Development: Include Each Century One by One

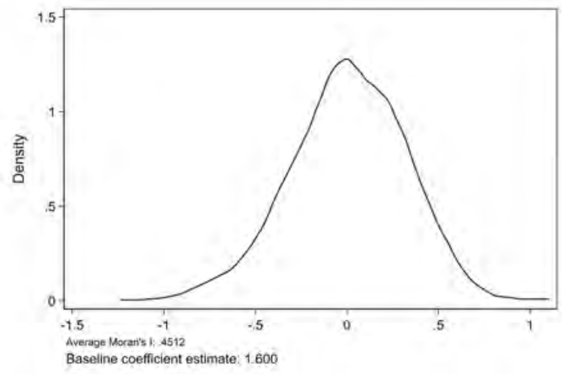


Notes. Each hollow dot represents the (standardized beta) point estimate for the variable of interest from the regression model in column 3 of Table 1 when we include each century of pre-colonial conflict exposure one by one. Horizontal bars indicate 95% confidence intervals. The data do not include any recorded land battles over 1400-1500 (there are five other types of conflict events over this century). Thus, our benchmark conflict exposure measure for 1400-1500 is zero.

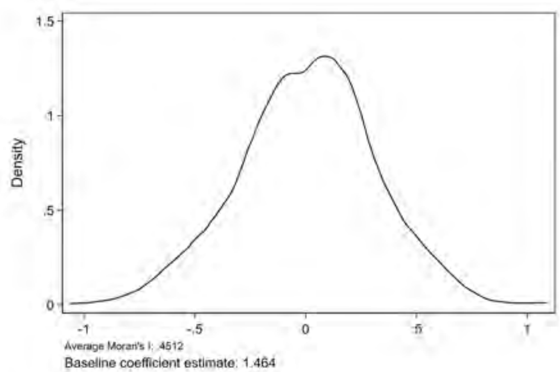
Figure A.9: Artificial Spatially-Correlated Noise Placebo Variables



(a) Column 1, Table 1



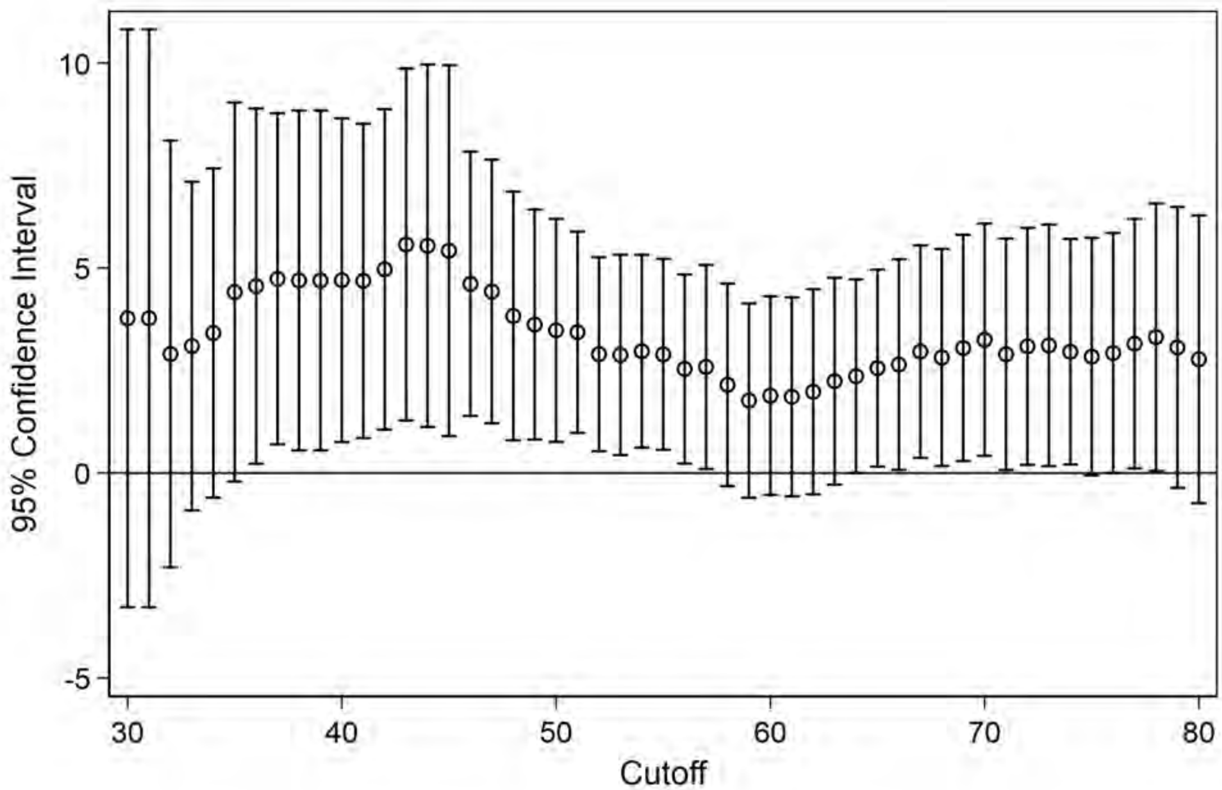
(b) Column 2, Table 1



(c) Column 3, Table 1

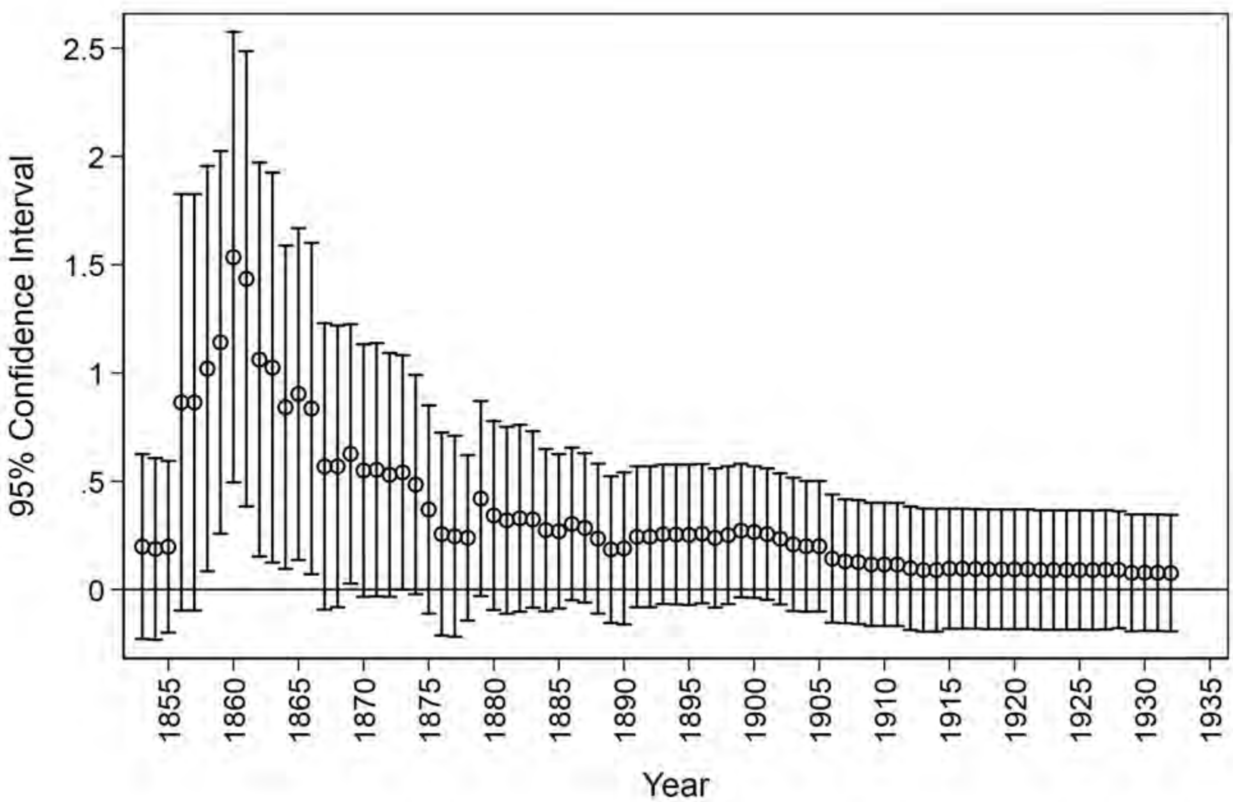
Notes. This figure shows the results of tests that generate artificial spatially-correlated noise placebo variables to replace our variable of interest, reallocating conflict exposure randomly across districts within a state (without replacement) for each of the regression models in Table 1.

Figure A.10: Sensitivity of IV Results to Alternative Cutoffs



Notes. Each hollow dot represents the point estimate for the regression model in column 3 of Table 2 for different cutoff values used to define the instrument. Horizontal bars indicate 95% confidence intervals.

Figure A.11: Pre-Colonial Conflict and Colonial Railroad Establishment



Notes. Each hollow dot represents the (standardized beta) point estimate for the variable of interest from the regression model in column 3 of Table 1 when the dependent variable is year of colonial establishment of first railroad connection within a district (if any). Horizontal bars indicate 95% confidence intervals.

Table A.1: Summary Statistics

	Mean	Std Dev	Min	Max	N
ln(0.01+Luminosity)	0.68	1.49	−4.61	4.14	664
Pre-colonial conflict exposure (1000-1757)	0.07	0.10	0.00	0.60	666
Colonial conflict exposure (1758-1839)	0.04	0.05	0.00	0.31	666
Colonial conflict exposure (1840-1946)	0.05	0.09	0.00	0.54	666
Post-colonial conflict exposure (1947-2010)	0.01	0.02	0.00	0.15	666
Khyber Proximity	0.08	0.26	0.00	1.00	660
ln(Population density), 1990	5.47	1.15	−1.44	10.61	665
Latitude	23.38	5.72	7.53	34.53	666
Longitude	81.12	6.39	69.47	96.83	666
Altitude	471.46	702.10	−200.73	4914.91	665
Ruggedness	98518.33	161662.39	0.00	851959.50	666
Precipitation	1370.74	698.71	200.22	4486.95	665
Land quality	0.45	0.29	0.00	0.97	662
Dry rice suitability	629.84	589.89	0.00	1722.67	665
Wet rice suitability	1439.98	797.25	0.00	2826.93	665
Wheat suitability	628.43	574.14	0.00	2914.67	665
Malaria risk	0.10	0.34	0.00	2.81	664
Neolithic settlements	0.36	1.54	0.00	20.00	666
Chacolithic settlements	0.29	1.38	0.00	19.00	666
Cultural sites (300-700 CE)	0.16	0.48	0.00	4.00	666
Cultural sites (8th-12th centuries)	0.66	1.23	0.00	10.00	666
Ln(1+ Urban population in 1000)	0.08	0.96	0.00	11.51	666
Medieval port	0.07	0.25	0.00	1.00	666
Linguistic fractionalization	0.46	0.27	0.01	4.21	666
Religious fractionalization	0.26	0.16	0.01	0.72	666
British direct rule	0.64	0.48	0.00	1.00	638
% Non-landlord, British India	50.81	42.68	0.00	100.00	166
ln(Tax/Area), All, 1881	−1.31	1.09	−4.84	1.17	274
ln(Tax/Area), British India, 1881	−1.46	1.08	−4.84	1.17	201
ln(Tax/Area), Princely states, 1881	−0.88	0.99	−3.05	1.06	73
ln(Tax/Person), All, 1881	0.30	0.91	−2.90	2.83	279
ln(Tax/Person), British India, 1881	−0.07	0.72	−2.90	1.86	201
ln(Tax/Person), Princely states, 1881	1.26	0.59	−0.21	2.83	78
ln(Tax/Area), 1931	−0.41	0.93	−4.20	1.39	145
ln(Tax/Person), 1931	0.32	0.81	−3.09	2.07	144
% Irrigated, 1931	4.76	9.54	0.00	60.99	257
% Irrigated, 1956-87	24.16	20.18	0.04	99.92	271
Fertilizer use, 1956-87 (Kg/Hectare)	2003.55	1798.55	0.66	10636.59	271
% HYV crop, 1956-87 (All)	23.10	12.05	0.00	59.54	271
% HYV crop, 1956-87 (Rice)	20.75	18.69	0.00	109.52	271
% HYV crop, 1956-87 (Wheat)	44.77	38.73	0.00	298.09	271
% HYV crop, 1956-87 (Other)	21.16	57.26	0.00	687.14	271
ln(Yield), Major crops, 1956-87	−0.16	0.51	−2.61	1.16	271
ln(Yield), Rice, 1956-87	−0.08	0.44	−1.19	0.85	266
ln(Yield), Wheat, 1956-87	−0.11	0.47	−1.39	0.91	260
ln(Yield), Other cereals, 1956-87	−0.39	0.49	−3.00	2.34	271
% Villages with primary schools, 1981	73.98	19.50	19.13	100.00	203
% Villages with middle schools, 1981	22.13	18.00	3.54	99.13	195
% Villages with high schools, 1981	9.76	12.05	0.56	81.74	187
Literacy rate, 1881	3.38	3.97	0.00	48.62	252
Literacy rate, 1921	6.33	5.46	0.73	36.39	322
Literacy rate, 1961-91	29.15	9.81	9.99	60.60	271
Literacy rate, 2001	59.70	13.69	25.40	96.60	591
Literacy rate, 2011	72.27	10.53	36.10	97.91	630
School attendance, 2001	70.70	14.96	27.50	94.90	591
Maoist control, 2003	0.09	0.29	0.00	1.00	395
Political violence, 2010-18 (100s of fatalities)	0.06	0.19	0.00	2.32	666

Notes. See text for variable descriptions and data sources.

Table A.2: Pre-Colonial Conflict and Economic Development: Grid Cell Fixed Effects

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	3.713***	1.875***	1.871***
	(0.305)	(0.524)	(0.390)
	[0.000]	[0.000]	[0.000]
Population density	Yes	Yes	Yes
Grid cell FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.240	0.121	0.121
R^2	0.598	0.777	0.814
Observations	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Grid cell fixed effects are 4° latitude × 4° longitude. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.3: Pre-Colonial Conflict and Economic Development: Alternative Standard Errors

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	3.713	1.601	1.465
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.240	0.104	0.095
State clustered p-value	0.012	0.038	0.034
Wild clustered bootstrap p-value	0.007	0.139	0.129
Conley spatial p-value	0.000	0.004	0.004
Moran's I statistic	0.508	0.070	0.044
R^2	0.598	0.829	0.849
Observations	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. We report the p-values for the coefficient estimates of the variable of interest for robust standard errors clustered by state, the wild cluster bootstrap procedure, and Conley spatial standard errors. The p-values for the wild cluster bootstrap procedure use 9,999 replications, while the Conley spatial standard errors use a cutoff distance of approximately 250 kilometers. Additionally, we report the Moran's I statistics.

Table A.4: Pre-Colonial Conflict and Economic Development: Alternative Cutoff Distance

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure (5,000 km cutoff)	4.080***	1.536***	1.378***
	(0.315)	(0.404)	(0.395)
	[0.000]	[0.000]	[0.001]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.278	0.105	0.094
R^2	0.615	0.828	0.848
Observations	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 5,000 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.5: Pre-Colonial Conflict and Economic Development: Running End-Date Cutoff

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure (running end-date)	3.289***	1.247***	1.081***
	(0.245)	(0.265)	(0.310)
	[0.000]	[0.000]	[0.001]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.392	0.149	0.129
R^2	0.480	0.814	0.825
Observations	377	377	377

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with variable end-date cutoff that includes exposure to conflicts that took place after 1757 but prior to British conquest of a district. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.6: Pre-Colonial Conflict and Economic Development: Alternative Conflict Data

Dependent variable:	Ln(0.01+Luminosity)	
	(1)	(2)
Pre-colonial conflict exposure (plus Clodfelter)	1.483***	
	(0.369)	
	[0.000]	
Pre-colonial conflict exposure (plus Clodfelter and Narvane)		1.227***
		(0.357)
		[0.001]
Population density	Yes	Yes
State FE	Yes	Yes
Geographic controls	Yes	Yes
Standardized beta coefficient	0.096	0.082
R^2	0.849	0.848
Observations	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Column 1 adds any conflicts from Clodfelter (2002) that do not already appear in the benchmark conflict database (i.e., Jaques, 2007), while column 2 adds non-overlapping conflicts from both Clodfelter (2002) and Narvane (1996). Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.7: Pre-Colonial Conflict and Economic Development: Conflict Types

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure (benchmark)	1.573***			
	(0.374)			
	[0.000]			
Pre-colonial conflict exposure (sieges)	-0.328			
	(0.289)			
	[0.257]			
Pre-colonial conflict exposure (single-day)		1.326***		
		(0.438)		
		[0.003]		
Pre-colonial conflict exposure (multi-day)		2.208***		
		(0.454)		
		[0.000]		
Pre-colonial conflict exposure (internal)			1.481***	
			(0.368)	
			[0.000]	
Pre-colonial conflict exposure (all)				0.681***
				(0.250)
				[0.007]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.102	0.074	0.096	0.066
R^2	0.849	0.849	0.849	0.847
Observations	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure between 1000-1757 with a cutoff distance of 250 kilometers. "Benchmark" restricts the conflict sample to land battles, while "siege" restricts it to sieges. "Single-day" and "multi-day" restrict this sample to land battles which lasted up to one day or multiple days. "Internal" restricts this sample to land battles internal to India. "All" includes the following conflict types: land battles, sieges, naval battles, and other conflict events (e.g., rebellion), whether single- or multi-day. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.8: Pre-Colonial Conflict and Economic Development: Alternative Exposure Measures

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure (# conflicts in convex hull, by participant)	0.005*** (0.002) [0.006]			
Pre-colonial conflict exposure (# conflicts in convex hull, by conflict group)		0.011*** (0.003) [0.000]		
Pre-colonial conflict exposure (Benchmark measure, by location of participant capitals)			0.424*** (0.126) [0.001]	
Pre-colonial conflict exposure (# conflicts in district of state capital)				0.014* (0.008) [0.094]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.099	0.115	0.081	0.023
R^2	0.847	0.848	0.848	0.846
Observations	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + \mathit{Luminosity})$ averaged between 1992-2010. Variable of interest is one of four alternative measures of pre-colonial conflict exposure. In the first alternative, we compute the convex hull for each participant according to the geographical coordinates of the conflicts in which that participant took part, treating all districts that intersect this convex hull as "affected" by a conflict. For each district, we count the number of conflicts by which they are so treated. In the second alternative, we compute the convex hull for each broad group of conflicts (e.g., "Later Mughal-Maratha Wars") as classified by Jaques (2007). In the third alternative, we compute conflict exposure again using Equation 1, replacing the locations of conflicts with those of the capitals of the pre-colonial states that participated in them. In the fourth alternative, we count the number of conflicts for each pre-colonial state, and assign these conflicts to the district that houses the state's capital. For further details of the construction of these variables, see Subsection 4.1.3. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(\mathit{PopulationDensity})$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.9: Pre-Colonial Conflict and Economic Development: Alternative Specifications of Dependent Variable

Dependent variable:	Ln(1+Luminosity)			Luminosity (levels)			Luminosity (IHS)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-colonial conflict exposure	3.625***	1.495***	1.313***	20.060***	10.725**	12.582***	3.595***	1.777***	1.791***
	(0.322)	(0.409)	(0.394)	(3.494)	(4.258)	(3.793)	(0.261)	(0.295)	(0.286)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.012]	[0.001]	[0.000]	[0.000]	[0.000]
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Geographic controls	No	No	Yes	No	No	Yes	No	No	Yes
Standardized beta coefficient	0.228	0.094	0.082	0.298	0.159	0.187	0.350	0.173	0.174
R^2	0.580	0.818	0.838	0.389	0.657	0.728	0.534	0.811	0.839
Observations	657	657	657	660	660	660	660	660	660

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(1 + Luminosity)$ in columns 1-3, *Luminosity* (levels) in columns 4-6, and *Luminosity* (inverse hyperbolic sine function) in columns 7-9. All dependent variables are averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.10: Pre-Colonial Conflict and Economic Development: GDP as Dependent Variable

Dependent variable:	Ln(GDP per Capita)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	1.782***	0.536**	0.448*
	(0.300)	(0.255)	(0.230)
	[0.000]	[0.036]	[0.052]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
Standardized beta coefficient	0.339	0.102	0.085
R^2	0.111	0.687	0.732
Observations	512	512	512

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(GDPperCapita)$ averaged between 1999-2007. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.11: Pre-Colonial Conflict and Economic Development: IV: Control for Trade Access

Panel A: First Stage

Dependent variable:	Pre-Colonial Conflict Exposure		
	(1)	(2)	(3)
Cost distance to Kyhber Pass	0.078***	0.080***	0.080***
	(0.024)	(0.024)	(0.024)
	[0.001]	[0.001]	[0.001]
Historical trade route	Yes	No	No
Silk Road site	No	Yes	No
Medieval trade port	No	No	Yes
Population density	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
R^2	0.685	0.670	0.673
Observations	660	660	660

Panel B: Second Stage

Dependent variable:	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	3.534**	3.465**	3.483**
	(1.412)	(1.388)	(1.392)
	[0.012]	[0.013]	[0.012]
Historical trade route	Yes	No	No
Silk Road site	No	Yes	No
Medieval trade port	No	No	Yes
Population density	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
Anderson-Rubin p-value	0.012	0.012	0.012
Kleibergen-Paap Wald rk F-statistic	10.376	10.572	10.654
Observations	660	660	660

Notes. Estimation method is 2SLS. Unit of analysis is district. In Panel A (first stage), dependent variable is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers, while variable of interest is cost distance to Khyber Pass. In Panel B (second stage), dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010, while variable of interest is pre-colonial conflict exposure between 1000-1757 with a cutoff distance of 250 kilometers, as instrumented by cost distance to Khyber Pass. *HistoricalTradeRoute* is a dummy variable that equals 1 for the presence of a major historical trade route or major port according to UNESCO. *SilkRoad* is a dummy variable that equals 1 for the presence of a Silk Road site. *MedievalPort* is a dummy variable that equals 1 for the presence of a major medieval port according to Jha (2013). Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. All previous controls are for both first and second stages. Population density is $\ln(PopulationDensity)$ in 1750 for first stage, and in 1990 for second stage. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.12: Pre-Colonial Conflict and Economic Development: IV: Placebo Entry Points

Panel A: First Stage

Dependent variable:	Pre-Colonial Conflict Exposure				
	(1)	(2)	(3)	(4)	(5)
Cost distance to placebo entry point	-0.041*	-0.002	0.007	-0.000	-0.047**
	(0.025)	(0.013)	(0.010)	(0.013)	(0.024)
	[0.097]	[0.853]	[0.462]	[1.000]	[0.049]
Placebo entry point	Surat	Kodungallur	Goa	Calicut	Bombay
Population density	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
R^2	0.656	0.651	0.651	0.651	0.657
Observations	660	660	660	660	660

Panel B: Second Stage

Dependent variable:	Ln(0.01+Luminosity)				
	(1)	(2)	(3)	(4)	(5)
Pre-colonial conflict exposure	7.315	34.071	-18.723	-113.599	7.994*
	(5.554)	(173.545)	(15.519)	(1925.976)	(4.715)
	[0.188]	[0.844]	[0.228]	[0.953]	[0.090]
Population density	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes
Anderson-Rubin p-value	0.073	0.287	0.002	0.294	0.018
Kleibergen-Paap Wald rk F-statistic	2.380	0.036	1.523	0.003	3.508
Observations	660	660	660	660	660

Notes. Estimation method is 2SLS. Unit of analysis is district. In Panel A (first stage), dependent variable is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers, while variable of interest is cost distance to placebo entry point (i.e., Surat, Kodungallur, Goa, Calicut, and Bombay). In Panel B (second stage), dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010, while variable of interest is pre-colonial conflict exposure between 1000-1757 with a cutoff distance of 250 kilometers, as instrumented by cost distance to placebo entry point. Geographic controls for both first and second stages include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1750 for first stage, and in 1990 for second stage. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.13: Pre-Colonial Conflict and Economic Development: IV: Alternative Cost Distance

Panel A: First Stage

Dependent variable:	Pre-Colonial Conflict Exposure			
	(1)	(2)	(3)	(4)
Cost distance to Khyber Pass (alternative)	0.089***	0.053**	0.106***	0.080***
	(0.022)	(0.022)	(0.022)	(0.023)
	[0.000]	[0.015]	[0.000]	[0.000]
Alternative cost distance	Linear slope	Squared slope	Linear ruggedness	HMI
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
R^2	0.668	0.661	0.678	0.667
Observations	660	660	660	660

Panel B: Second Stage

Dependent variable:	Ln(0.01+Luminosity)			
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	5.232***	3.106*	5.042***	6.208***
	(1.668)	(1.872)	(1.284)	(1.750)
	[0.002]	[0.097]	[0.000]	[0.000]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Anderson-Rubin p-value	0.001	0.114	0.000	0.000
Kleibergen-Paap Wald rk F-statistic	14.646	5.415	20.938	10.483
Observations	660	660	660	660

Notes. Estimation method is 2SLS. Unit of analysis is district. In Panel A (first stage), dependent variable is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers, while variable of interest is alternative cost distance to Khyber Pass. In Panel B (second stage), dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010, while variable of interest is pre-colonial conflict exposure between 1000-1757 with a cutoff distance of 250 kilometers, as instrumented by alternative cost distance to Khyber Pass. In column 1, alternative cost distance is computed as linear slope. In column 2, it is computed as squared slope. In column 3, it is computed as linear ruggedness. In column 4, it is computed based on human mobility index (HMI) according to Özak (2010, 2018). Geographic controls for both first and second stages include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1750 for first stage, and in 1990 for second stage. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.14: Pre-Colonial Conflict and Economic Development: Additional Geographic Features

Dependent variable:	Ln(0.01+Luminosity)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-colonial conflict exposure	1.400***	1.250***	1.451***	1.435***	1.464***	1.115***	0.747**
	(0.367)	(0.363)	(0.380)	(0.381)	(0.365)	(0.361)	(0.381)
	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.002]	[0.050]
Ln(1+Distance to coast)	Yes	No	No	No	No	No	Yes
Ln(1+Distance to border)	No	Yes	No	No	No	No	Yes
River dummy	No	No	Yes	No	No	No	Yes
Irrigation potential	No	No	No	Yes	No	No	Yes
Rainfall variation	No	No	No	No	Yes	No	Yes
Ln(1+Distance to resource deposit)	No	No	No	No	No	Yes	Yes
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls (benchmark)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.051	0.268	0.004	0.027	0.042	-0.032	0.056
R^2	0.849	0.850	0.849	0.847	0.849	0.856	0.857
Observations	660	660	660	652	660	660	652

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variable is $\ln(0.01 + Luminosity)$ averaged between 1992-2010. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Benchmark geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Additional geographic controls include natural logarithm of (one plus) distance to coast, natural logarithm of (one plus) distance to border, river presence, irrigation potential, rainfall variation, and natural logarithm of (one plus) distance to resource deposits. Population density is $\ln(PopulationDensity)$ in 1990. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.15: Pre-Colonial Conflict and Economic Development: Health

Dependent variable:	%Infant Mortality	%Health Center	%Health Subcenter
	(1)	(2)	(3)
Pre-colonial conflict exposure	-35.283**	-3.372*	-8.290*
	(14.405)	(1.971)	(4.509)
	[0.015]	[0.089]	[0.068]
Population density	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
Standardized beta coefficient	-0.112	-0.100	-0.120
R^2	0.674	0.788	0.644
Observations	270	203	188

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. $\%InfantMortality$ is the infant mortality rate in 1991. $\%HealthCenter$ and $\%HealthSubcenter$ measure the proportion of villages having a primary health center or subcenter in 1981. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1950. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.

Table A.16: Pre-Colonial Conflict and Economic Development: Transportation

Dependent variable:	Pre-1934	1981-91		
	Railroad establishment	%Railroad	%Road	%Canal
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	2.319*	-4.646***	13.200	-17.080***
	(1.241)	(1.524)	(9.798)	(5.999)
	[0.062]	[0.002]	[0.179]	[0.005]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.086	-0.190	0.049	-0.193
R^2	0.573	0.488	0.867	0.350
Observations	660	391	391	391

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. $RailroadEstablishment$ is decade of establishment of first railroad connection within a district before 1934 (if any), computed as $(1934 - EstablishmentYear)/10$. $\%Railroad$, $\%PavedRoad$, and $\%Canal$ are averaged between 1981 and 1991. They measure the proportion of villages having access to a railroad, paved road, or navigable canal. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757 with a cutoff distance of 250 kilometers. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is $\ln(PopulationDensity)$ in 1900 in column 1, and in 1950 in columns 2-4. Robust standard errors in parentheses, followed by p-values in brackets. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level.