

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

Pre-Colonial Warfare and Development in India

Mark Dincecco<sup>1,\*</sup>, James Fenske<sup>2</sup>, Anil Menon<sup>1</sup>, and Shivaji Mukherjee<sup>3</sup>

**Abstract:** Does pre-colonial history – and in particular the role of interstate warfare – help explain long-run development patterns across India? To address this question, we construct a new geocoded database of historical conflicts on the Indian subcontinent. We document a robust positive relationship between pre-colonial conflict exposure and local economic development today. Drawing on archival and secondary data, we show that districts that were more exposed to pre-colonial conflict experienced greater early state-making, followed by lower political violence and higher investments in physical and human capital in the long term.

**Keywords:** Warfare, Economic Development, State Capacity, Public Goods, India, History

**Classification:** N45, O11, P48, H11

---

\* **Correspondence address:** Department of Political Science, University of Michigan, 5700 Haven Hall, 505 South State Street, Ann Arbor, MI 48109, USA. Email: dincecco@umich.edu.

We thank Editor Ekaterina Zhuravskaya, four anonymous referees, Jeanet Bentzen, Pietro Biroli, Traviss Cassidy, Latika Chaudhary, Giorgio Chiovelli, Jeremiah Dittmar, Shashwat Dhar, Edgar Franco Vivanco, Bishnupriya Gupta, Mai Hassan, Namrata Kala, Mark Koyama, Alexander Lee, Alexander Persaud, Priya Satia, Yuri Zhukov, and audiences at the Aix-Marseille School of Economics, London Business School, National University of Singapore, New Economic School, Sorbonne, University of Bonn, University of Bristol, University of Chicago, University of East Anglia, University of Gothenburg, University of Laval, UC Louvain, University of Michigan, UQAM, University of St. Gallen, University of York, and several conferences for helpful comments; Latika Chaudhary, Saumitra Jha, Alexander Lee, and Rinchan Mirza for data-sharing; and Justin Huang and Eric Payerle for research assistance.

## 1 Introduction

According to a well-known literature, governments undertook institutional reforms in order to enhance their military prowess in the face of interstate competition (Mann, 1984; Brewer, 1989; Tilly, 1992; Gennaioli and Voth, 2015). In time, more powerful government institutions helped promote long-run economic development through the greater provision of domestic security and other basic public goods (Besley and Persson, 2011; Morris, 2014). Much of this literature, however, centers on the historical experience of Western Europe.

In this study, we recast the ‘military competition’ framework to provide a novel explanation of local development patterns across India today. 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). We construct a new geocoded database of historical interstate conflicts in this context. 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 between the years 1000 and 1757. Consistent with the logic described above, our empirical analysis reveals a positive and significant relationship between pre-colonial conflict exposure and local economic development. This result sheds new light on the historical roots of Indian development patterns.

To test the robustness of this relationship, we perform numerous checks. First, we restrict our analysis to within-state variation by including state fixed effects, to show that time-invariant features specific to Indian states do not drive our results. Second, we control for a wide range of local geographic features, including climate, terrain ruggedness, soil suitability, disease environments, and waterway access. Third, 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 initial state capacity levels, ethnic and religious fractionalization, colonial-era and post-colonial conflict exposure, and a host of other potential confounders. Fourth, we demonstrate that our main results are robust to alternative ways of operationalizing local exposure to pre-colonial interstate military competition that take into account faraway conflicts by exploiting information about

the pre-colonial states that participated in them. Fifth, we perform an instrumental variables analysis that exploits variation in pre-colonial conflict exposure driven by proximity to the Khyber Pass, the main historical route of invaders from Central Asia into India.

We next analyse the channels through which pre-colonial warfare may have influenced long-run development patterns in India. We argue that reductions in local levels of violence and greater investments in physical and human capital were functions – at least in part – of more powerful local government institutions. Drawing on rich data from both archival and secondary sources, we show evidence for a significant relationship between local exposure to pre-colonial conflict and diverse measures of pre-colonial and colonial-era state-making, as well as long-term reductions in political violence. Furthermore, we show that pre-colonial conflict exposure significantly predicts larger investments in physical capital (i.e., irrigation infrastructure) and human capital (i.e., literacy and education) in the long term.

Our study provides new evidence that the ‘military competition’ framework applies outside Western Europe, at least in India. A recent literature has explored the relationship between interstate military competition and long-run state capacity in India (Roy, 2013; Gupta *et al.*, 2016; Foa, 2016). We go further in several ways. First, we analyse the long-run implications of historical warfare in India for *economic* development. Here, we view local state capacity as a means through which pre-colonial conflict exposure may have influenced local development patterns, and not simply as an end in and of itself. In this respect, our theoretical framework and empirical analysis go beyond classic arguments such as those by Tilly (1992), which primarily focus on the relationship between warfare and state-making. Second, we construct a new geocoded database of historical interstate conflicts on the Indian subcontinent. Third, 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.

Of equal importance, our study casts new light on the deep roots of Indian development patterns. The vast majority of the literature analyses the role of British colonialism (Banerjee and Iyer, 2005; Iyer, 2010; Bharadwaj and Ali Mirza, 2019; Castelló-Climent *et al.*, 2018; Lee, 2019; Chaudhary *et al.*, 2020). Our empirical analysis evaluates the importance of colonial factors in several ways. Yet we emphasize the role of pre-colonial events in India, which the

literature tends to overlook. Namely, our argument that pre-colonial conflict exposure promoted early local state-making, followed by less political violence and greater infrastructure investments in the long term, provides a novel explanation for local development patterns in India.

There is a growing literature about the significance of pre-colonial factors such as state capacity for long-run development (Gennaioli and Rainer, 2007; Michalopoulos and Papaioannou, 2013; Dell *et al.*, 2018). Among these papers, ours is one of the first to systematically analyze 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 capacity – which, according to our argument, was actually an outcome of prior military conflict. We show evidence in support of this view. In this manner, our study extends the literature by shedding new light on the historical roots of state capacity in the developing world.

Finally, there is a growing body of quantitatively-oriented research on pre-colonial India (Jha, 2013; Gaikwad, 2014; Iyer *et al.*, 2017), to which our study brings the role of pre-colonial interstate military competition in explaining long-run development patterns. We provide new insights about the pre-colonial military roots of current 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. Section 4 describes our empirical strategy and data. In Section 5, we present the main results along with the robustness checks. Section 6 performs the instrumental variables analysis. In Section 7, we analyse potential channels. In Section 8, we conclude by explaining how our study helps clarify the geopolitical conditions under which warfare can ‘make states’ and promote economic development.

## 2 Theoretical Framework

### 2.1 General Argument

Interstate warfare is a common explanation for long-run state-making. We characterize the general logic of this argument 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 defence, and be willing to pay more in taxes to fund it. In this manner, the threat of foreign attack may drive higher tax revenue, along with a more robust public administration to help organize the government's fiscal and military efforts.<sup>1</sup> If there are recurring threats, then institutional reforms may continue in ratchet-like steps (Rasler and Thompson, 2005, 491-3). Once the government 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 foreign threats dissipate.

In time, a more powerful government may help promote long-run economic development via at least two channels. The first channel is the greater provision of domestic security. A more powerful government should be better at maintaining domestic law and order (Morris, 2014, 3-26). If there is a reduction in levels of internal violence, then individuals will be more willing to make growth-enhancing investments (North, 1981, 24-6). The second channel is the provision of other public goods that depend on a less violent domestic environment (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.

Given that warfare can destroy physical and human capital, greater interstate military competition may actually reduce the long-run prospects for economic development. Here, it makes sense to distinguish between the short-run and long-run economic implications of interstate military competition. Namely, while warfare may be destructive over the short

---

<sup>1</sup> This logic should hold so long as large financial resources matter to battlefield success (Gennaioli and Voth, 2015). As we will describe in Section 3, money played a key role in interstate military competition in the context of pre-colonial India.

run, the economic benefits that may derive from competition-related institutional reforms – as characterized in the two channels above – may eventually outweigh such destruction over the long run.<sup>2</sup>

## 2.2 Application to India

As described in Section 1, the ‘military competition’ framework centres on the historical experience of Western Europe. Yet there is also reason to think that this general logic may apply in India. As in early modern Europe, political fragmentation, instability, and interstate military competition were recurrent features of the landscape of pre-colonial India (Tilly, 1992, 45, de la Garza, 2016, 12). Similarly, large fiscal resources played an important role in military success in both contexts (Gennaioli and Voth, 2015; Roy, 1994). Finally, population density was high enough in both early modern Europe and pre-colonial India to make territorial conquest through battle a worthwhile endeavour (Herbst, 2000, 13-16). Thus, the logic by which greater levels of interstate military competition may have promoted local institutional reforms should have held within pre-colonial India. We will provide further historical evidence in support of the above claims in Section 3.

Local institutional reforms made in response to interstate military competition in pre-colonial India may have endured across the colonial and post-colonial eras, as new regimes took advantage of traditional institutional structures, rather than trying to build new ones from scratch. Between the mid-eighteenth century and the mid-nineteenth century, the British East India Company became the dominant political power on the Indian subcontinent (Dutt, 1950, 1-2; Gommans, 1999, 120). However, the total influx of British settlers to India was relatively small (Iyer, 2010, 697). Thus, British colonialists had incentives to establish indirect forms of rule, under which traditional local 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). Lange (2004, 909) writes 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

<sup>2</sup> According to Centeno and Enriquez (2016, 124), even the short-run destruction of warfare may bring economic benefits, to the extent that it reduces reliance on outdated technology and reallocates public spending toward infrastructure investments.

ruled Africa.’ Princely states (i.e., ‘Native’ states) ruled by hereditary kings spanned 45% of British India (Iyer, 2010, 694), while colonial dependence on ‘customary’ (i.e., traditional) courts was 60% (Lange, 2004, 909).

By relying on existing local institutions, the British could reduce overall governance costs. We may therefore expect local pre-colonial institutions to have endured into the colonial era. For example, the British East India Company was able to quickly extract sizeable tributes following its victories in the British-Mysore Wars in the late eighteenth century (Roy, 2011, 65-6). This example speaks to the fiscal strength of the pre-colonial Mysore Kingdom, as well as to Britain’s practice of exploiting traditional local institutions.

In 1947, India became independent of British rule. A single federal government characterizes post-colonial India. Still, the traditional strength of local governance structures could influence the local effectiveness of national-level institutions, particularly given the decentralized nature of India’s federal government, as well as India’s vast geographic scale. In this manner, pre-colonial institutional structures could continue to influence long-run local development outcomes into the modern era.

### 2.3 Empirical Predictions

When applied to the Indian context, therefore, the ‘military competition’ framework produces several predictions. The main ‘reduced-form’ prediction is that there should be a positive and significant relationship between pre-colonial conflict exposure and current economic development levels in India. The logic is as follows. If a given area in India experienced more pre-colonial warfare, then we would expect more powerful local government institutions to have emerged there, which in turn would have helped promote local long-run economic development.

Our argument produces three further predictions that reflect the channels through which the main prediction may have operated. The first such prediction is that greater pre-colonial conflict exposure should be associated with pre-colonial and colonial-era state-making. A related prediction is that there should be a negative and significant relationship between pre-colonial conflict exposure and (eventual) political violence levels. A final prediction is that there should be a positive and significant relationship between pre-colonial conflict exposure

and subsequent investments in physical and human capital such as irrigation infrastructure that depend on a less violent domestic environment.

We will rely on these four predictions to guide our empirical analysis.

### 3 Historical Background

We now provide historical evidence regarding the relationship between interstate military competition and institutional development in pre-colonial India in support 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).<sup>3</sup> By the early sixteenth century, major rival states included the Delhi Sultanate, the Rajput states, the Deccan Sultanates, and the Vijayanagar Empire (Roy, 1994, 57).

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

Between 1526 and 1707, the Mughal Empire was among the most powerful states on the Indian subcontinent (Richards, 1995, 1, 6-9; de la Garza, 2016, 1). This Empire was established by Babur, who after several attempts 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.

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

---

<sup>3</sup> We take 1000 as the start year for synchronicity with the case of Western Europe, which provides the backdrop for our analysis. There, the turn of the first millennium marked the approximate onset of political fragmentation after the demise of the Carolingian Empire (Strayer, 1970, 15). It was in this context that the logic of interstate military competition became relevant.



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 political 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% of total state expenditure was granted to Mughal military officials called *mansabdars*, while another 9% was devoted to the central military establishment (Richards, 1995, 75-6). By contrast, annual spending on the Mughal imperial household was less than 5%.

To help manage Mughal military affairs, Akbar implemented new bureaucratic and fiscal structures (Richards, 1995, 58-9; de la Garza, 2016, 6). Under the institutional innovation of the *mansabdari-jagirdari* system, Akbar granted land to military officials in order to extract surplus agricultural output (Nath, 2018, 253-5). Data available for the late 1680s indicate that the top 6% of military officials (roughly 450 persons) were in possession of more than 60% of total tax revenue, indicating a high degree of bureaucratic and fiscal centralization under a small military elite (Qaisar, 1998, 255-6). A large portion of these funds were spent on the military.

Furthermore, Mughal government officials developed a 'pyramid' treasury system that linked the central treasury with those in provincial capitals and other towns (Richards, 1995, 69-71). Akbar exploited this bureaucratic innovation to quickly move funds during conflicts. 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 institutional innovation (Richards, 1995, 187-90). In the late sixteenth century, the state began to overhaul the land tax revenue system, increasing bureaucratic centralization and introducing better agricultural data. By enabling the state bureaucracy 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 state to extract agricultural output and finance the military. Moreover, the system may have incentivized farmers to shift production to high-value cash crops, thereby promoting rural economic development.

The Mughal Empire fell into decline at the start of the eighteenth century (Richards, 1995, 253-81). Building off of the institutional legacy that the Mughals left behind, indigenous kingdoms including the Maratha, Mysore, and Travancore began to compete for political control with the British East India Company (Roy, 1994, 37-50; Roy, 2011, 95-102). In this context, states undertook major institutional reforms (Stein, 1985, 391; Roy, 1994, 37-50; Ramusack, 2003, 12; Roy, 2011). 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 greater revenue (Foa, 2016, 93-4). Describing this system, Foa (2016, 94) writes that the 'flow of revenues to the centre 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).

Victory at the Battle of Plassey in 1757 helped establish the British East India Company as a major political entity on the Indian subcontinent (Dutt, 1950, 1-2; de la Garza, 2016, 12). Over the next century, the East India Company systematically defeated its rivals in India, including indigenous states such as the Marathas, Mysore, and Sikhs, along with foreign powers such as the Dutch and French (Dutt, 1950, 1-2; Gommans, 1999, 120).

Overall, this historical evidence links interstate military competition and recurrent warfare in pre-colonial India to bureaucratic and fiscal development and state capacity improvements. As we have explained in Subsection 2.2, there is reason to think that local pre-colonial institutional reforms made in response to interstate military competition endured across the colonial and post-colonial eras. Through the (eventual) greater provision of domestic security and basic infrastructure, such institutions could promote long-run economic development at the local level in India.

## 4 Empirical Strategy and Data

### 4.1 Empirical Strategy

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

$$Y_{i,j} = \beta \text{ConflictExposure}_{i,j} + \lambda \text{PopDensity}_{i,j} + \mu_j + X'_{i,j} \phi + \epsilon_{i,j}, \quad (1)$$

where  $i$  indexes districts and  $j$  indexes states in modern-day India.

We take the district as our main unit of analysis because it is the unit for which the most comprehensive data on outcomes, potential channels, and controls are available. District borders, however, are potentially outcomes of pre-colonial conflict exposure. In the Appendix, we show that the main results are robust if we take grid cells as the unit of analysis rather than districts (Table A.3).<sup>4</sup>

$Y_{i,j}$  measures local economic development levels in terms of luminosity. We justify the use of luminosity as our main economic outcome in Subsection 4.3. Following Michalopoulos and Papaioannou (2013), we take the natural logarithm, adding a small number such that  $Y_{i,j} \equiv \ln(0.01 + \text{Luminosity}_{i,j})$ . 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  $\ln(1 + \text{Luminosity}_{i,j})$  rather than  $\ln(0.01 + \text{Luminosity}_{i,j})$ ; (2) keep  $Y_{i,j}$  in its original linear form; or (3) take the inverse hyperbolic sine function (Table A.4).

$\text{ConflictExposure}_{i,j}$  measures pre-colonial conflict exposure, our variable of interest. We discuss the data and construction of this variable in the next subsection.

$\text{PopDensity}_{i,j}$  controls for log population density in the most recent year available prior to the year in which the dependent variable is measured.<sup>5</sup> While local luminosity levels in India do not simply reflect population density, we follow the guidance in Michalopoulos and

<sup>4</sup> Furthermore, the main results continue to hold if we take the sub-district (i.e., tehsil) as our unit of analysis (Table A.3).

<sup>5</sup> For the main regression analysis, this year is 1990. When the dependent variable is historical (e.g., 1881), then this year is subject to data availability. The 1990 population data are taken from the Center for International Earth Science Information Network (CIESIN) (<http://sedac.ciesin.columbia.edu/gpw>), and the historical population data from Klein Goldewijk *et al.* (2010).

Papaioannou (2018, 391) and include log population density as a control. This approach allows us to interpret our results in terms of local differences in per capita living standards, without the imposition of any prior restrictions on the luminosity-population elasticity.<sup>6</sup> We thus believe that the inclusion of log population density as a control outweighs the potential cost in terms of post-treatment bias, given that dense populations may in part be outcomes of pre-colonial conflict exposure. In the Appendix, however, we show that the main results continue to hold if we exclude log population density (Table A.6). Similarly, they remain robust if we control for historical log population density in 1000, the start year of our analysis (Table A.7). In both cases, the coefficient estimates are larger than the main estimates, which suggests that we are taking a conservative approach by including log population density.

$\mu_j$  is the fixed effect for each of the 36 federal states (more precisely, 29 states and 7 union territories). Modern state borders may in part be outcomes of pre-colonial conflict exposure. Thus, including state fixed effects in our regression analysis may induce post-treatment bias. In our view, however, their inclusion outweighs this potential cost, since fixed effects help control for state-level institutional and cultural features, along with potential measurement error in the quality and coverage of the conflict data across space. Nonetheless, the main results remain robust if we exclude the state fixed effects from the regression analysis (see column 1 of Table 1 ahead). Furthermore, in the Appendix we show that the main results continue to hold if we use ‘exogenous’ grid cell fixed effects rather than state fixed effects (Table A.8).

Local geography may influence patterns of both pre-colonial conflict and economic development alike. Geographic zones with mild climates and high quality soils may promote human settlements (Ashraf and Galor, 2011). Settlements may reduce the cost of collective military action and incite violent conflict (Besley and Reynal-Querol, 2014). Similarly, populated zones may make for attractive targets for attackers (Glaeser and Shapiro, 2002). Agricultural conditions, moreover, can influence the viability of new taxation (Lee, 2019). To account for the possibility that certain zones engender recurring conflict as well as institutional and economic development due to favourable geography, the vector  $X_{i,j}$  controls for a wide range of local geographic features, including latitude, longitude, altitude, ruggedness,

<sup>6</sup>Furthermore, this approach enables us to account for potential non-linearities in log population density, which we do by including polynomial terms as a robustness check (Table A.5). The main results continue to hold.

precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk.<sup>7</sup>

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. Table A.9 reports standard errors that allow for general forms of spatial autocorrelation of the error term (Conley, 1999) for six different cutoff distances between 250 and 1,500 kilometres. Table A.10 reports (1) standard errors robust to clustering at the state level and (2) tests of  $\beta$  using the wild cluster bootstrap at the state level (Cameron *et al.*, 2008) based on 9,999 replications. The main results remain significant for both Conley spatial standard errors and cluster-robust standard errors, and just miss statistical significance for the wild cluster bootstrap procedure.<sup>8</sup> As another way to account for potential spatial correlation, Table A.11 includes polynomial terms for latitude and longitude. The main results continue to hold.

Tables A.1 and A.2 display the summary statistics for the variables in our regression analysis.

## 4.2 Conflict Data

According to the theoretical framework in Section 2, higher local levels of interstate military competition helped incentivize local institutional reforms in pre-colonial India. To proxy for local interstate military competition in this context, we use geocoded data on historical conflicts. The logic here is that there was a meaningful link between the actual prevalence of local conflict and local levels of interstate military competition in pre-colonial India.

<sup>7</sup> We compute latitude and longitude by identifying district centroids using a polygon file of district boundaries from [gadm.org](http://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) (<http://www.fao.org/nr/gaez/en/>). 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). We take the raster index for the stability of malaria transmission from Kiszewski *et al.* (2004).

<sup>8</sup> To further account for the possibility that spatial correlation leads to standard errors that are too small (Kelly, 2019), we 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 geographic controls is 0.044, indicating spatial autocorrelation in the regression residuals (Table A.10). However, the placebo variables nearly always fail to produce treatment effects as large as those of our main coefficient estimates (Figure A.3).

To construct our historical conflict database, we rely primarily on the book by Jaques (2007), the goal of which is to document as many historical conflicts as possible (Jaques, 2007, xi, xiii). For inclusion, a conflict must have been written down and cross-referenced with a minimum of two independent sources. 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.

The conflict information in Jaques' book is organized alphabetically by individual conflict names. For each individual conflict, Jaques provides a paragraph-length description, including the type (e.g., land battle), date, approximate duration (e.g., single-day), approximate location, and major participants. 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 ( $34^{\circ} 1' 0''$  N,  $71^{\circ} 35' 0''$  E).

Our database includes all individual conflicts – for example, land battles, sieges, and naval battles – on the Indian subcontinent between 1000 and 2010 as recorded by Jaques.<sup>9</sup> For our benchmark measure of local exposure to pre-colonial conflicts (to be described ahead), we focus on land battles, 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 capital would be destroyed. For robustness, we control for local exposure to: (1) pre-colonial sieges; and (2) all pre-colonial conflict types; in the Appendix (Table A.12). There continues to be a significant relationship between local exposure to pre-colonial land battles and current development.<sup>10</sup> Figure A.1 maps the locations of the conflicts in our sample, while Figure A.2 breaks them down by historical sub-period. We show that the main results are robust if we restrict the conflict data to the sub-period of 1500 to 1757 in the Appendix (Table A.13).

<sup>9</sup> By 'Indian subcontinent,' we mean that we include conflict events that took place in 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 limited interactions between China and India. There were few if any historical interstate conflicts anywhere near China's border with India (Dincecco and Wang, 2018, 345). In the Appendix, we restrict our benchmark conflict exposure measure to conflict events that took place within modern India only (Table A.12). The main results remain unchanged.

<sup>10</sup> However, there is no such significant relationship for local exposure to pre-colonial sieges. This evidence is consistent with the logic described above that sieges were generally more destructive than land battles.

To verify the accuracy and breadth of our historical conflict coverage, we constructed alternative conflict data according to similar procedures from two other 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 the use of Jaques as our baseline source. Naravane's book focuses on battles in medieval India. While his coverage does expand on Jaques, it lacks details on individual conflicts.<sup>11</sup> Regardless, in the Appendix, we add the non-overlapping pre-colonial data from Clodfelter and Naravane to our benchmark conflict exposure measure (Table A.14). The main results remain significant.<sup>12</sup>

Although we systematically check the breadth of our conflict coverage, there may still be measurement error. First, the available data do not enable us to systematically account for potential differences in the intensity of pre-colonial conflicts. However, our primary focus on *interstate* conflicts – and in particular land battles as our benchmark measure – helps ensure that we are making ‘apples-to-apples’ comparisons between conflict events. Furthermore, interstate conflicts are exactly the sort of conflict that the literature (e.g., Besley and Persson, 2011) indicates should matter most for political and economic development. Nonetheless, in the Appendix, we show that the main results remain robust if we include all pre-colonial conflict types as recorded by Jaques (Table A.12).

Second, the quality and coverage of the historical conflict data may potentially vary by geographic zone. Our regression analysis accounts for potential differences in historical data quality and coverage across space in several ways, including: (1) the use of fixed effects for Indian states (Table 1) or for grid cells (Table A.8); (2) controls for initial state capacity (Table A.19); and (3) the exclusion of individual states (Figure A.4) or colonial provinces (Figure A.5) from our main specification one at a time.

---

<sup>11</sup> 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.

<sup>12</sup> Brecke (1999) is another potential alternative source for historical conflict data. Relative to Jaques, however, there are two main shortcomings of this work: (1) his data do not start until 1400; and (2) similar to Naravane, he does not provide specific information about conflict locations.

As described above, we use geocoded historical conflict data to proxy for local levels of interstate military competition in pre-colonial India, which may have influenced the likelihood of local institutional reforms. To compute our benchmark measure of local exposure to individual conflicts, we define the conflict exposure of Indian district  $i$  as:

$$\sum_{c \in \mathcal{C}} (1 + distance_{i,c})^{-1}. \quad (2)$$

We measure  $distance_{i,c}$  from the centroid of district  $i$  to the location of conflict  $c$ . To reduce the measure's sensitivity to any single conflict, we add one to  $distance_{i,c}$  before taking the inverse.<sup>13</sup> According to this measure, the nearer a district is to a particular conflict, the more exposed that district is. Conflicts occurring at the district centroid receive a weight of one, or full weight; as the distance of conflicts from the centroid increases, they receive lower weights. Importantly, this measure does not rely on any (anachronistic) cutoff occurring at a district's borders. Our benchmark conflict exposure measure includes pre-colonial land battles between 1000 and 1757 within a radius of 250 kilometres. In the Appendix, we use an alternative radius of 5,000 kilometres (Table A.15). 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 (Table A.16).<sup>14</sup> The coefficient estimates are very similar in magnitude and significance to the main estimates across both checks.<sup>15</sup>

We view local exposure to conflicts as the most straightforward way to measure the local extent of interstate military competition in the pre-colonial context. This approach, however, may overlook conflicts that were fought at a large distance from a pre-colonial state's political centre, but nonetheless prompted institutional reforms there. In Section 5, we will describe and test a set of alternative ways of operationalizing local interstate military competition that take into account faraway conflicts by exploiting information about the pre-colonial states that participated in them.

<sup>13</sup> If we did not add one to this measure, then a district in which a conflict took place very near to the centroid would receive a large conflict exposure value, regardless of its proximity to any other conflicts.

<sup>14</sup> Specifically, this measure includes conflicts from 1000 to the (potentially post-1757) year of British annexation, which differs by district.

<sup>15</sup> For further robustness, we exclude 155 districts for which our benchmark conflict exposure measure takes a value of zero. The main results continue to hold (not reported).



### 4.3 Economic Data

To measure local levels of current development across India, we use nighttime luminosity data. Luminosity data provide a systematic proxy of economic activity across locales in relatively poor parts of the world in which official data may not be widely available (Henderson *et al.*, 2012; Michalopoulos and Papaioannou, 2013; Min, 2015, 51-73). We take these data from the Operational Linescan System of the Defence 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 kilometre) resolution. We compute average luminosity across all square kilometre cells within each district for every year between 1992 and 2010, and then take the district averages over the entire 1992-2010 period. In the Appendix, we show the main results if we restrict the luminosity data to each year from 1992 to 2010 (Figure A.6). The coefficients are always significant, with only a small decline in magnitude over time. The 1992-2010 luminosity data are significantly correlated with the more recent Visible Infrared Imaging Radiometer Suite (VIIRS) data from NASA. The main results are similar in magnitude and significance if we take the VIIRS data averaged between 2015 and 2020 as the outcome variable (Table A.17).

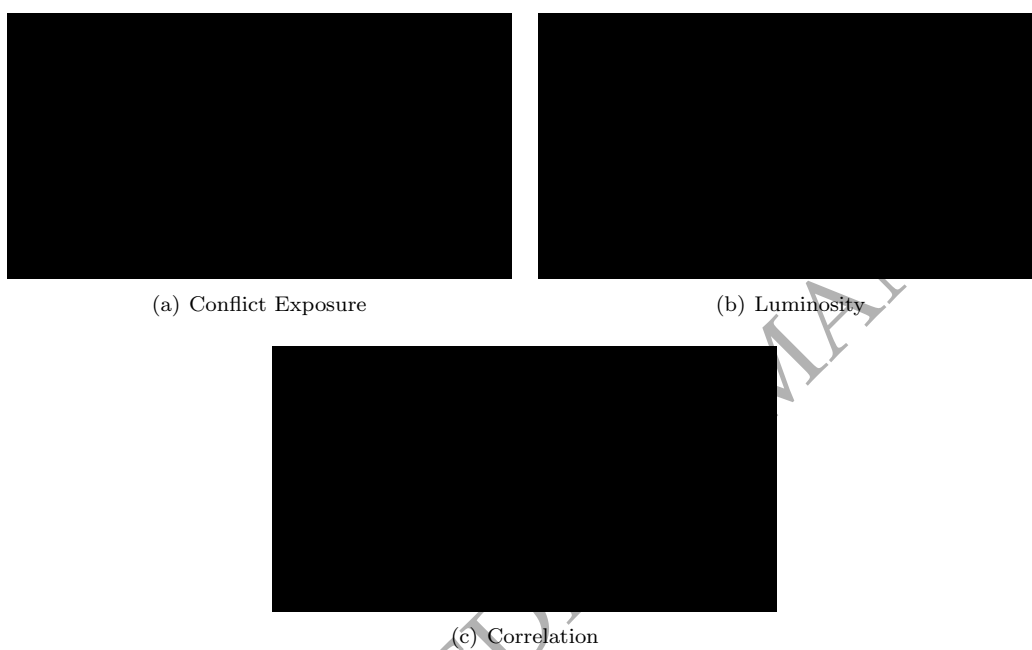
District-level GDP per capita data exist for India, but they have been constructed by a private company, and differ from official sources such as the National Sample Surveys in their rankings of districts on economic development outcomes. They are thus not widely used in the empirical literature (Castelló-Climent *et al.*, 2018, 5). Nonetheless, we use them as an alternative development outcome in the Appendix (Table A.18). The main results continue to hold.

### 4.4 Descriptive Patterns

Panel (a) of Figure 1 maps our benchmark measure of pre-colonial conflict exposure across Indian districts. This panel 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. Panel (b) maps average luminosity. This

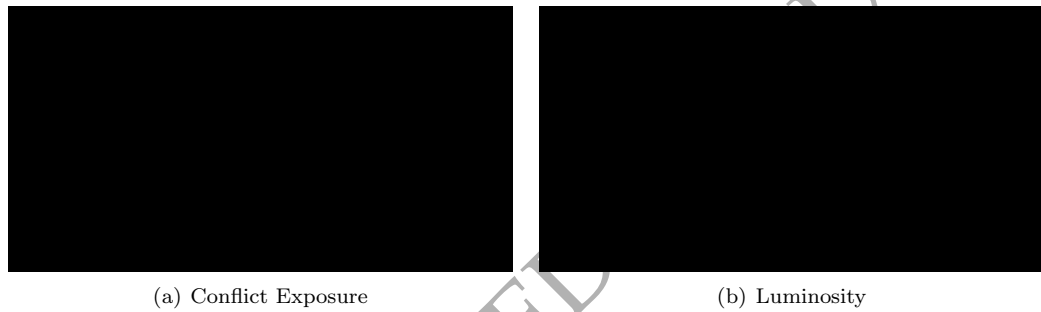
panel suggests that economic development levels tend to be highest in the vicinity of the four main geographic zones of pre-colonial conflict. Panel (c) shows that there is in fact a strong positive correlation between the two variables. For a micro perspective, Figure 2 zooms in on districts within Uttar Pradesh, India's most populated state, which contains more than 70 districts. The correlations between pre-colonial conflict exposure and luminosity within this state are striking.

ORIGINAL UNEDITED MANUSCRIPT

Figure 1. *Pre-Colonial Conflict and Luminosity: India*

*Notes.* Panel (a) shows pre-colonial conflict exposure to land battles between 1000-1757 by district in India, while Panel (b) shows average luminosity between 1992-2010. Districts are shaded by decile: districts in the top decile receive the darkest shade. Panel (c) plots pre-colonial conflict exposure against luminosity. Both variables are residualized by controlling for  $\ln(\text{PopulationDensity})$  in 1990.

Figure 2. *Pre-Colonial Conflict and Luminosity: Uttar Pradesh*



*Notes.* Panel (a) shows pre-colonial conflict exposure to land battles between 1000-1757 by district in Uttar Pradesh, while Panel (b) shows average luminosity between 1992-2010. Districts are shaded by decile: districts in the top decile receive the darkest shade.

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

<i>Dependent variable:</i>	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	3.713*** (0.305) [0.000]	1.601*** (0.380) [0.000]	1.465*** (0.370) [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. 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.

## 5 Empirical Results

### 5.1 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 (unstandardised) coefficient estimate for  $ConflictExposure_{i,j}$  is 3.713, and is significant at the 1% 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.<sup>16</sup>

Overall, the Table 1 results support the main ‘reduced-form’ prediction of our theoretical framework. Namely, there is a positive and significant relationship between local exposure to pre-colonial conflicts and levels of economic development in India today. The coefficient

<sup>16</sup> Local geographic features may influence long-run development patterns in non-linear ways. To help account for this possibility, we include the quadratic term for each such control. The main results remain robust (not reported).

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 magnitude is roughly similar in size to the effect of pre-colonial political centralization on current luminosity levels in Africa found by Michalopoulos and Papaioannou (2013, 130), who report a standardized beta coefficient of 0.12. It is also roughly 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, for which one can compute a standardized beta coefficient of 0.14.

## 5.2 Additional Control Variables

### 5.2.1 Initial Conditions

In Table A.19, we control for initial state capacity by district in multiple 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 or eleventh through twelfth centuries based on Nag (2007, 28, 30), or in 1525 based on Joppen (1907).

In Table A.20, we repeat the main analysis after taking into account several additional local geographic controls beyond those in the baseline specification. They include the natural logarithm of (one plus) the distance to the nearest coast, river presence, irrigation potential, rainfall variation, the natural logarithm of (one plus) the distance to the nearest resource deposits (i.e., diamonds, gems, gold, petroleum), and the percentage of forested area.<sup>17</sup>

<sup>17</sup> We compute the natural logarithm of (one plus) distance from each district to the coast. We report a 'river' dummy that indicates whether a district is intersected by a major rivers according to Natural Earth Data (<https://www.naturalearthdata.com/>). We compute each district's irrigation potential using data from Bentzen *et al.* (2017). To account for drought prevalence, we control for the mean and coefficient of variation in rainfall according to Matsuura and Willmott (2009). We control for the natural logarithm of (one plus) distances from the district centroid to deposits of diamonds, gems, gold, and petroleum using data from Tollefsen *et al.* (2012). We compute the percentage of forested area using data from the India Institute of Forest Management (2015).

The main results continue to hold across both types of robustness checks.

### 5.2.2 Colonialism

One strand of the literature highlights the colonial origins of contemporary economic development in India. In Table A.21, we control for the potential role of colonial institutions. 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. In Table A.22, we account for the establishment of the colonial railroad network according to Fenske *et al.* (2021), who identify the year in which the first colonial railroad connection was made within each district.<sup>18</sup> Finally, in Table A.23, we evaluate whether agricultural investments by the British in the ‘canal colonies’ confound our analysis by excluding all districts in states and union territories that correspond to the historical Punjab. The results of these tests indicate that pre-colonial history – and in particular conflict exposure – continues to significantly predict current local development levels in India above and beyond colonial-era factors.

### 5.2.3 Fractionalization

Another strand of the literature emphasizes the role of inter-ethnic and religious relations in India. In Table A.24, we control for such factors in multiple ways. Column 1 includes a dummy variable for districts that had major medieval ports, which according to Jha (2013) were traditionally zones of ethnic tolerance. Alternatively, we account for the duration of medieval Muslim rule in each district in column 2 (Jha, 2013). Column 3 controls for the current share that is Muslim in each district, while column 4 controls for current religious polarization levels. In columns 5 and 6, we account for current local linguistic and religious fractionalization levels, respectively. In column 7, we control for the current shares of Scheduled Castes and Scheduled Tribes in each district.<sup>19</sup> Finally, in column 8 we account for

<sup>18</sup> In order to include districts that had not received a railroad by 1934, the year in which the edition of *History of Indian Railways Constructed and in Progress* that underlie these data was published, we recode missing years as zero and control for a dummy for missing years.

<sup>19</sup> The religion, Scheduled Castes, and Scheduled Tribes data are from the 2011 Indian Census (<https://censusindia.gov.in/>). We compute religious polarization levels according to the method in Montalvo and Reynal-Querol (2005), and compute fractionalization levels according to Omid’s Peoples of South Asia Database (<https://legacy.joshuaproject.net/data-sources.php>).

whether a district is intersected by the Ganges River, which due to Hindu sacred geography is linked with the proportion of upper castes (Jha, 2013, 815). These robustness checks imply that inter-ethnic and religious relations do not confound our main results.

#### 5.2.4 Post-1757 Conflict

To account for the potential role of post-1757 conflict exposure, 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 cutoff marked by the emergence of British military and political dominance over the Indian subcontinent in the 1840s (Clodfelter, 2002, 244-50).<sup>20</sup> The results in Table A.25 indicate that local exposure to colonial and post-colonial conflicts does not diminish the predictive importance of pre-colonial conflict exposure. Colonial 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 that of pre-colonial warfare.

#### 5.2.5 Other Potential Confounders

In Table A.26, we control for distance from a district's centroid to five major urban centers: Bangalore, Bombay, Chennai, Delhi, and Kolkata. Distance from India's major urban centres does not confound the main results.<sup>21</sup> In Table A.27 we control for whether a district is intersected by Asian Highway 1, the longest overland route in Asia. The main results continue to hold.

<sup>20</sup> 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 the results described in Table A.25 remain similar in terms of sign and significance for this alternative cutoff (not reported).

<sup>21</sup> Additionally, we control for distance from 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 results are robust (not reported).



### 5.3 Alternative Measures of Conflict Exposure

Our benchmark measure may overlook conflicts that were fought at a large distance from a pre-colonial state's political centre, but nonetheless prompted institutional reforms there. To address this possibility, we produce a set of four alternative conflict exposure measures. Here, we code each major state participant in our pre-colonial conflict database, and identify its capital city.

In our first alternative measure, we calculate the number of conflicts in which each pre-colonial state participated, and assign these conflicts to the district in which the state was headquartered (regardless of how far away they were actually fought). As a second way to include a pre-colonial state's participation in faraway conflicts, we compute conflict exposure using Equation 2, but replace the locations of the conflicts with those of the capitals of the pre-colonial states that participated in them. As another way to apportion pre-colonial conflicts across state participants, 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. To illustrate this approach, Figure A.7 plots the convex hull for pre-colonial conflicts involving the seventeenth-century Mughal ruler Shah Jahan. As a final way to include a pre-colonial state's participation in faraway conflicts, regardless of conflict location, we compute the convex hull for each broad cluster of conflicts as categorized by Jaques (e.g., 'Later Mughal-Maratha Wars').

The five conflict exposure measures described above are all significantly correlated (Table A.28), suggesting that our empirical results are not contingent upon the choice of any specific measure. In Table A.29, we re-run the regression analysis for these alternative measures of pre-colonial conflict exposure. Our main results remain robust.

## 6 IV Analysis

To instrument for pre-colonial conflict exposure, we construct a measure of each district's proximity to the Khyber Pass. The South Asian subcontinent is naturally protected from

invasion by several mountain ranges (e.g. the Himalayas). Historically, the Khyber Pass was the main route for invaders coming from Central Asia to India (Docherty, 2008). Thus, proximity to the Khyber Pass can be treated as a forcing variable that affects a district's exposure to pre-colonial conflict.

## 6.1 IV Construction

The proximity of each district in India to the Khyber Pass in terms of simple geodesic distance does not accurately measure how difficult it was to reach it. Despite their proximity to the Khyber Pass, mountainous states such as Jammu and Kashmir, Himachal Pradesh, and Uttaranchal were 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.

We construct our measure of cost distance using raster data on ruggedness (Özak, 2010). We define the ruggedness of a cell as the average difference in absolute elevation between that cell and its eight neighbours (Nunn and Puga, 2012). We assume that the cost of crossing a cell is proportionate to the square of its ruggedness. We compute the least-cost path and associated cost of travel between each grid cell in India and the Khyber Pass. Our benchmark cost-distance measure to the Khyber Pass averages over all cells in each district.

As distance from the Khyber Pass increases, the relationship between our cost-distance measure and pre-colonial conflict exposure becomes nonlinear, driven by conflicts such as the Carnatic Wars that were unrelated to invasions from Central Asia (Figure A.8). Hence, 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 Khyber Pass in terms of cost distance. We operationalize our instrument in this way since proximity to the Khyber Pass should decrease the cost of a treated district's exposure to threats of invasion from Central Asia. This cost decrease is for reasons external to the district itself – namely, the specific geography of the territory that lies between it and the Khyber Pass. Figure A.9 plots this measure, while Figure A.10 shows that most pre-colonial land battles fought by invaders from Central Asia actually took place within the region that our instrument would predict. There is a positive and significant

relationship between the Khyber proximity instrument and current economic development across Indian districts (Table A.30).

## 6.2 IV Results

Table 2 reports the first-stage and second-stage results for the same three specifications as in Table 1. Our first-stage results suggest that being proximate to the Khyber Pass increases exposure to conflict by 0.08 to 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 are larger than 10, indicating instrument strength and a relatively low propensity for bias at the second stage. The second-stage coefficient estimates suggest 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.

## 6.3 IV Robustness

### *Historical Trade*

One potential objection to our instrument is that the Khyber Pass introduced South Asia to phenomena beyond conflict exposure that may have been relevant to both pre-colonial state-making and later economic development. Trade is the most notable such potential factor. In the main IV analysis, we have controlled for local geographic features that may have influenced historical trade patterns. Similarly, we have included state fixed effects to help account for modern differences in trade policy across state governments.

For robustness, we account for historical trade in several additional ways. First, we code a district as having access to a historical trade route if it was intersected by a major trade route or had a major port in the seventeenth century according to the map in Raychaudhuri (1982, 334). Second, we code a dummy equal to one for districts containing Silk Road sites in India according to UNESCO. Third, we code a dummy equal to one for districts with a major medieval port according to Jha (2013).

Table 2. *Pre-Colonial Conflict and Economic Development: IV*

<i>Panel A: First Stage</i>			
<i>Dependent variable:</i>	Pre-Colonial Conflict Exposure		
	(1)	(2)	(3)
Proximity to Khyber Pass	0.204*** (0.018) [0.000]	0.094*** (0.025) [0.000]	0.080*** (0.024) [0.001]
Population density	Yes	Yes	Yes
State FE	No	Yes	Yes
Geographic controls	No	No	Yes
$R^2$	0.415	0.645	0.665
Observations	660	660	660
<i>Panel B: Second Stage</i>			
<i>Dependent variable:</i>	Ln(0.01+Luminosity)		
	(1)	(2)	(3)
Pre-colonial conflict exposure	4.930*** (0.609) [0.000]	4.626*** (1.291) [0.000]	3.482** (1.389) [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, while variable of interest is proximity to Khyber Pass. In Panel B (second stage), dependent variable is  $\ln(0.01 + \text{Luminosity})$  averaged between 1992-2010, while variable of interest is pre-colonial conflict exposure between 1000-1757, as instrumented by proximity 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  $\ln(\text{PopulationDensity})$  in 1990. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

In Table A.31, we regress each measure of historical trade on the Khyber proximity instrument. There are no statistically significant correlations between our instrument and the historical trade measures, which suggests that historical trade is unlikely to confound the IV analysis. Nonetheless, in Table A.32 we exclude districts that contained a major historical trade route, Silk Road site, or major medieval port, respectively. The IV results remain robust. Thus, districts that were historically important to trade do not drive the IV

findings. Alternatively, Table A.33 indicates that the IV results are robust to the inclusion of the historical trade variables as controls. Furthermore, the IV results hold for cost-distance constructions of the historical trade controls (Table A.34).

### *Placebo Entry Points*

For further robustness, we employ as placebos alternative points of entry into South Asia that invaders did not historically rely on. The placebo points of entry are Surat in west-central India and Bombay, Calicut, Goa, and Kodungallur in the southwest. For each of them, we compute an analogous cost-distance measure as 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 in cost-distance terms. Table A.35 shows that these cost distances generally fail to predict conflict exposure. In the IV specifications, the placebo points of entry cannot generally be used to infer a positive effect of conflict exposure on modern economic development. The exception is Bombay, with a weak first stage F-statistic of less than four.

### *Alternative Cost Distance*

To show that the specific construction of our instrument does not drive our IV results, we implement four alternatives to measure the cost of crossing a grid cell: linear slope; squared slope; linear ruggedness; and a human mobility index (HMI).<sup>22</sup> The HMI cost measure is based on the speed that a military infantry unit can maintain while walking over different terrain types (Özak, 2018). This measure helps capture the historical costs of military movements prior to the invention of railroads. Table A.36 shows that these alternative measures of computing cost distance yield results similar to the main IV results.

---

<sup>22</sup> For the HMI cost measure, we expand our IV cutoff to the closest 100 districts. 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.

### *FATA Region of Pakistan*

The federally administered tribal areas (FATA) of Pakistan are near the Khyber Pass and experienced pre-colonial conflicts, yet remain underdeveloped today. To show that their exclusion from the main IV analysis does not artificially inflate the results, Table A.37 includes districts in Pakistan and Bangladesh (i.e., the two other main components of the British Raj apart from modern-day India). The results are robust, with second-stage coefficient estimates that are actually larger in magnitude than in the main IV analysis.<sup>23</sup>

### *Railroad Network*

Military concerns were one motivation for the construction of the colonial-era railroad network (Bogart and Chaudhary, 2016). If the British feared attack by Russia via the Khyber Pass, and so constructed railways in its vicinity, then it could threaten the exclusion restriction. In Table A.39, we control for the year in which the first colonial railroad connection was made within each district (if any). The main IV results continue to hold.

## 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 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, we have argued that reductions in local levels of violence and greater investments in physical and human capital were (at least in part) functions of more powerful local government institutions. In line with this argument, our framework produces three predictions that reflect the channels through which the main ‘reduced-form’ result may have operated. First, greater pre-colonial conflict exposure should be associated with pre-colonial and colonial-era state-making. Second, there should be a negative and significant relationship between pre-colonial conflict exposure and (eventual) political violence levels. Finally,

<sup>23</sup> For completeness, Table A.38 shows that the OLS results remain robust to the inclusion of districts in Pakistan and Bangladesh.

there should be a positive and significant relationship between pre-colonial conflict exposure and subsequent investments in physical and human capital such as irrigation infrastructure and literacy and education that depend on a less violent domestic environment. In this section, we evaluate each channel one at a time.

## 7.1 Pre-Colonial and Colonial-Era State-Making

We begin the channels analysis by testing the link from pre-colonial conflict exposure to early state-making efforts.

Fiscal strength is central to state power (Levi, 1988, 2). Ideally, we might measure the extent of pre-colonial state-making across the Indian subcontinent in terms of fiscal outcomes. However, systematic fiscal data are not available for the pre-colonial era. We thus rely on the best available measures of pre-colonial state-making that we were able to uncover. The first is the number of important Mughal sites reported by Schwartzberg (1978). 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 bridges, forts, and palaces. To the extent that public works depend on the state's ability to extract resources from local populations, then this variable proxies for pre-colonial state capacity. Second, we use maps of the Mughal Empire digitized by Jha (2013) to identify districts incorporated by the rulers Babur, Akbar, and Aurangzeb, respectively. Following a precedent in the literature (e.g., Bockstette *et al.*, 2002; Heldring, 2021) we interpret the longevity of pre-colonial state history as a measure of early state strength.

In Table 3, we regress our 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. 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, Akbar, and Aurangzeb. There is a positive and significant relationship between pre-colonial conflict exposure and early state capacity under both Babur and Akbar.<sup>24</sup>

<sup>24</sup> In Table A.40, we examine the possibility that sieges were less conducive to pre-colonial state-making than land battles. Consistent with the logic described in Subsection 4.2, the coefficients for local exposure to pre-colonial sieges are generally not statistically significant.

Table 3. *Pre-Colonial Conflict and Pre-Colonial-Era State-Making*

<i>Dependent variable:</i>	Important Mughal Sites	State History		
	(1)	Babur (2)	Akbar (3)	Aurangzeb (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 including public works. 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. 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  $\ln(\text{Population Density})$  in 1500. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

Fiscal outcomes – our ‘ideal’ way to measure historical state-making – are in fact available for the colonial era. We rely on the archival source of Baness (1881), from which we construct data on land tax revenue, physical size, and population for several hundred historical political units under direct or indirect British rule during the late nineteenth century.<sup>25</sup> To match historical states to modern districts, we rely on the information on provincial and state

<sup>25</sup> Here, indirect rule refers to major Princely states. 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).



names in Baness.<sup>26</sup> Finally, to supplement the late nineteenth-century fiscal data, we rely on the 1931 land tax revenue data for districts in British India from Lee (2019).

Table 4 regresses our measures of colonial fiscal development on pre-colonial conflict exposure. In columns 1 to 6, we take the available 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.<sup>27</sup>

In columns 7 and 8 of Table 4, we take the available 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, even if historical fiscal differences between them later diminished. These results are in line with Lee (2019), 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.

Overall, this evidence is consistent with the first prediction described above, namely that pre-colonial conflict exposure played a significant role in pre-colonial and colonial-era state-making.

---

<sup>26</sup> 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.

<sup>27</sup> Furthermore, the results in Table A.41 suggest that pre-colonial conflict exposure does not simply proxy for underlying levels of pre-colonial state centralization.

Table 4. Pre-Colonial Conflict and Colonial Fiscal Development

	Dependent variable:							
	1881				1931			
	All		British India		Princely State		British India	
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)	
Pre-colonial conflict exposure	2.245*** (0.550) [0.000]	1.245*** (0.382) [0.001]	2.205*** (0.516) [0.000]	1.157*** (0.354) [0.001]	6.386* (3.524) [0.076]	1.612 (2.228) [0.473]	1.310 (0.911) [0.153]	0.835 (0.709) [0.238]
Population density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.256	0.173	0.281	0.227	0.496	0.210	0.135	0.098
R <sup>2</sup>	0.468	0.518	0.596	0.545	0.606	0.408	0.731	0.696
Observations	270	274	200	200	70	74	145	144

Notes. Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows:  $\ln(Tax/Area)$ , 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 measure 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, and battles between 1800-1757. Geographic controls include latitude, altitude, ruggedness, precipitation, land and quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is  $\ln(PopulationDensity)$  in 1850 in columns 1-6 and 1930 in columns 7-8. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

## 7.2 Political Violence

In line with our theoretical framework, we continue the channels analysis by testing the relationship between pre-colonial conflict exposure and (eventual) political violence levels.

In Table 5, we regress local exposure to colonial and post-colonial conflicts on pre-colonial conflict exposure. There is a positive and 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 significant for the post-colonial era. Districts that experience more pre-colonial conflict exposure, therefore, experienced significantly less conflict between 1947 and 2010.<sup>28 29</sup>

---

<sup>28</sup> The results in Table 5 (and in column 1 of Table 6) continue to hold if we take the sub-district (i.e., tehsil) as our unit of analysis (Table A.42).

<sup>29</sup> The results in Table A.43 suggest, moreover, that pre-colonial conflict exposure is not simply a proxy for underlying levels of pre-colonial state centralization.

Table 5. *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 1947-2010	
	Land Battles (1)	All Conflicts (2)	Land Battles (3)	All Conflicts (4)	Land Battles (5)	All Conflicts (6)
Pre-colonial conflict exposure	0.170*** (0.036) [0.000]	0.448*** (0.090) [0.000]	0.037 (0.039) [0.344]	0.313 (0.306) [0.306]	-0.024*** (0.005) [0.000]	-0.030*** (0.007) [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	0.433	0.041	0.204	-0.126	-0.105
R <sup>2</sup>	0.560	0.561	0.740	0.562	0.813	0.873
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 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. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is  $\ln(\text{Population Density})$  in 1750 in columns 1-2, in 1850 in columns 3-4, and in 1950 in columns 5-6. 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 Post-Colonial Political Violence*

<i>Dependent variable:</i>	Political Violence	Maoist Control	Fractionalization	
	(1)	(2)	Linguistic (3)	Religious (4)
Pre-colonial conflict exposure	-0.241** (0.102) [0.019]	-0.381** (0.163) [0.020]	-0.209* (0.113) [0.065]	0.080 (0.071) [0.260]
Population density	Yes	No	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	No	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 2015-18 (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. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is  $\ln(\text{PopulationDensity})$  in 1990. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

We take two other measures of political violence as outcome variables in Table 6. Column 1 regresses the number of fatalities per district between 2015 and 2018 according to the ACLED Project on pre-colonial conflict exposure. Here, we find a negative and significant relationship between pre-colonial conflict exposure and contemporary political violence in terms of fatalities. Column 2 regresses local Maoist control in 2003 according to Mukherjee (2018) on pre-colonial conflict exposure. In 2006, Prime Minister Manmohan Singh called the Maoist insurgency ‘India’s number one internal security threat’ (Mukherjee, 2018, 236). We show that pre-colonial conflict exposure predicts a significantly lower likelihood of local control by Maoist insurgents.

Overall, the above results are consistent with the second prediction described above, that previous conflict exposure may pave the way for domestic peace in the long term (Morris, 2014, 3-26).<sup>30</sup>

<sup>30</sup> 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 government) was able to establish a

Finally, in columns 3 and 4 of Table 6, we regress linguistic and religious fractionalization in 2001 on pre-colonial conflict exposure. 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 (Alesina *et al.*, 2020).

### 7.3 Physical and Human Capital

We conclude the channels analysis by testing the relationship between pre-colonial conflict exposure and the subsequent provision of other public goods that depend on a less violent domestic environment, and in particular those that promote investments in physical and human capital.

#### 7.3.1 Irrigation Infrastructure

Columns 1 and 2 of Table 7 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 era and the post-colonial one.<sup>31</sup> These results are consistent with our theoretical framework, namely that those districts that were more exposed to pre-colonial conflict – and hence may have developed more powerful local government institutions, and have eventually provided greater domestic security – may have been better placed to make local investments in physical capital.

Column 3 of Table 7 indicates that pre-colonial conflict exposure predicts significantly greater local yields of major crops in post-colonial India. The prior results for irrigation infrastructure suggest that physical capital improvements may help explain this increase in agricultural production. Greater crop yields may enable a larger non-agricultural labour

---

widespread monopoly over violence across India. In the meantime, however, we would expect warfare to persist so long as there was ongoing interstate military competition in India. During the pre-colonial era, in fact, we find evidence for conflict persistence from one century to the next (not reported).

<sup>31</sup> We take colonial-era data on irrigation infrastructure in 1931 from Bharadwaj and Ali Mirza (2019). We rely on irrigation data for post-colonial India from Banerjee and Iyer (2005). They provide district-level data for 13 major Indian states averaged between 1956 and 1987.

Table 7. *Pre-Colonial Conflict and Irrigation Infrastructure*

<i>Dependent variable:</i>	%Irrigated		Ln(Yield)	%Non-Agriculture
	1931	1956-87		
	(1)	(2)	(3)	(4)
Pre-colonial conflict exposure	21.275** (10.357) [0.041]	37.413** (15.758) [0.018]	0.737* (0.381) [0.054]	0.197** (0.085) [0.021]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	0.194	0.173	0.134	0.100
$R^2$	0.391	0.611	0.683	0.566
Observations	257	271	271	660

*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).  $\ln(Yield)$  measures the total yield across 15 major crops averaged between 1956-87 (column 3). *%NonAgriculture* measures the share of non-agricultural workers in 2011. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757. 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 2-3. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

force. Column 4 shows that pre-colonial conflict exposure does in fact predict a significantly greater share of workers employed outside of agriculture.<sup>32</sup>

In column 1 of Table A.44, we show that the extent of irrigation infrastructure in the colonial era significantly predicts its extent in the post-colonial period. Yet pre-colonial conflict exposure still exerts a significant direct influence on the extent of post-colonial irrigation infrastructure above and beyond the colonial-era investments. This evidence further suggests that the influence of pre-colonial conflict exposure on infrastructure investments endured beyond the colonial era.

Table 8. *Pre-Colonial Conflict and Literacy*

<i>Dependent variable:</i>	%Literacy			
	1881 (1)	1921 (2)	1961-91 (3)	2011 (4)
Pre-colonial conflict exposure	-1.933 (3.188) [0.545]	-5.635 (3.772) [0.136]	11.796* (6.888) [0.088]	10.146** (4.119) [0.014]
Population density	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Standardized beta coefficient	-0.047	-0.095	0.112	0.103
$R^2$	0.464	0.556	0.623	0.599
Observations	251	303	271	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, 2011 measures the adult literacy rate across both rural and urban populations for ages 7-plus. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is  $\ln(\text{PopulationDensity})$  in 1850 in column 1, in 1900 in column 2, in 1950 in column 3, and in 2011 in column 4. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

### 7.3.2 Literacy and Education

Table 8 takes the local literacy rate in 1881, in 1921, averaged between 1961 and 1991, and in 2011 as the dependent variables.<sup>33</sup> There is no significant relationship between pre-colonial conflict exposure and district-level literacy rates under British colonial rule. This relationship, however, turns positive and significant across both post-colonial observation years. We interpret these results in terms of our theoretical framework as follows. India began to pursue a state-led industrialization policy after independence in 1947 (Gupta, 2019, 2). In this context, districts that had been more exposed to historical conflict (and were thus more likely to develop more powerful local government institutions and less violent domestic environments) may have been better placed to make basic investments in human capital.

<sup>32</sup> The yield data are from Banerjee and Iyer (2005), while the non-agricultural worker data are from the 2011 Indian Census (<https://censusindia.gov.in/>).

<sup>33</sup> We take the 1881 and 1921 data from Fenske and Kala (2021), the 1961-91 data from Banerjee and Iyer (2005), and the 2011 data from the Indian Census (<https://censusindia.gov.in/>).



Table 9. *Pre-Colonial Conflict and Education*

<i>Dependent variable:</i>	%Primary	%High	%Infant Mortality
	(1)	(2)	(3)
Pre-colonial conflict exposure	18.683* (11.150) [0.096]	-16.094** (6.553) [0.015]	-35.283** (14.405) [0.015]
Population density	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
Standardized beta coefficient	0.099	-0.139	-0.112
$R^2$	0.712	0.840	0.674
Observations	203	187	270

*Notes.* Estimation method is OLS. Unit of analysis is district. Dependent variables are as follows. %*Primary* measures the proportion of villages having a primary school in 1981. %*High* measure the proportion of villages having a high school in 1981. %*InfantMortality* is the infant mortality rate in 1991. Variable of interest is pre-colonial conflict exposure to land battles between 1000-1757. Geographic controls include latitude, longitude, altitude, ruggedness, precipitation, land quality, dry rice suitability, wet rice suitability, wheat suitability, and malaria risk. Population density is  $\ln(\text{PopulationDensity})$  in 1950 in columns 1 and 2, and in 1990 in column 3. Robust standard errors in parentheses, followed by p-values in brackets. \*\*\*, \*\*, and \* indicate statistical significance at 1%, 5%, and 10% level.

In columns 2 and 3 of Table A.44, we show that the extent of literacy in the colonial era significantly predicts its extent in the post-colonial period. Yet pre-colonial conflict exposure still exerts a significant direct influence on the extent of post-colonial literacy rates above and beyond the colonial era. This evidence further suggests that the influence of pre-colonial conflict exposure on literacy has endured beyond the colonial era.

The first two columns of Table 9 employ local education measures in 1981 as outcomes (taken from Banerjee and Iyer (2005)). Pre-colonial conflict exposure predicts significantly greater provision of primary education (column 1). These results make sense in light of the previous findings for literacy rates. By contrast, the coefficient estimate for  $\text{ConflictExposure}_{i,j}$  is negative and significant when the dependent variable is the provision of secondary education in 1981 (column 2).

Overall, this evidence suggests that pre-colonial conflict exposure promoted greater investments in basic human capital in the long term, but ran counter to more advanced human capital investments. In an agriculturally-oriented society such as India, widespread access

to literacy and primary education may bring greater economic benefits than investments in higher (i.e., secondary) education (Kotwal *et al.*, 2011, 1159).

The final column of Table 9 indicates that pre-colonial conflict exposure predicts significantly lower infant mortality in 1991 (taken from Banerjee and Iyer (2005)). Viewed in combination with the results above, this reduction in infant mortality may reflect greater investments in basic human capital such as literacy and primary education.

#### 7.4 Section Summary

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) pre-colonial and colonial-era state-making; (2) greater domestic security in the long term; and (3) greater colonial or post-colonial investments in physical and human capital. In line with our argument, we view reductions in local levels of violence and greater investments in physical and human capital as functions – at least in part – of more powerful local government institutions.

### 8 Conclusion

We have analysed the role of pre-colonial history – and in particular the role of interstate 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 the greater provision of domestic security and other basic public goods.

To evaluate the predictions of this argument, we have exploited a new geocoded database of historical interstate 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, followed by lower political violence and higher

investments in basic infrastructure in the long term, help explain this relationship. This analysis casts new light on the deep roots of Indian development patterns.

Our study shows that the ‘military competition’ 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. Furthermore, historical population density in pre-colonial India was high enough – as in Western Europe – to make territorial acquisition through warfare worthwhile. Pre-colonial Sub-Saharan Africa, by contrast, was land-rich but labour-scarce. This high land-labour ratio meant that a traditional goal of African warfare was to capture slaves (Herbst, 2000, 13-16, 20). Thornton (1999, 16) writes that ‘ownership of slaves in Africa was virtually equivalent to owning land in Western Europe or China.’ Low population density may have weakened the relationship between warfare and state centralization, since in the face of conflict individuals could simply migrate and farm faraway virgin land, rather than stay and participate in the institutional build-up for defense. Herbst (2000, 36) labels this phenomenon the ‘primacy of exit.’ There was no significant correlation between warfare and state centralization in pre-colonial Africa (Osafa-Kwaako and Robinson, 2013). Moreover, the correlation between pre-colonial conflict and long-run development outcomes in Africa is negative (Besley and Reynal-Querol, 2014; Dincecco *et al.*, 2019). In this manner, our study helps clarify the geopolitical conditions – in particular, political fragmentation and high population density – under which warfare ‘makes states’ and promotes economic development.

## 9 Supplementary data

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: <https://doi.org/10.5281/zenodo.5583263>.

**Affiliations**

<sup>1</sup>University of Michigan

<sup>2</sup>University of Warwick

<sup>3</sup>University of Toronto

ORIGINAL UNEDITED MANUSCRIPT

## References

- Alesina, A., Reich, B. and Riboni, A. (2020). 'Nation-Building, Nationalism, and Wars', *Journal of Economic Growth*, vol. 25(4), pp. 381–430.
- Ashraf, Q. and Galor, O. (2011). 'Dynamics and Stagnation in the Malthusian Epoch', *American Economic Review*, vol. 101(5), pp. 2003–41.
- Banerjee, A. and Iyer, L. (2005). 'History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India', *American Economic Review*, vol. 95(4), pp. 1190–1213.
- Baness, J.F. (1881). *Geographicus Indicus*, Edward Stanford.
- Bentzen, J., Kaarsen, N. and Wingender, A. (2017). 'Irrigation and Autocracy', *Journal of the European Economic Association*, vol. 15(1), pp. 1–53.
- Besley, T. and Persson, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton University Press.
- Besley, T. and Reynal-Querol, M. (2014). 'The Legacy of Historical Conflict: Evidence from Africa', *American Political Science Review*, vol. 108(2), p. 319336.
- Bharadwaj, P. and Ali Mirza, R. (2019). 'Displacement and Development: Long-Term Impacts of Population Transfer in India', *Explorations in Economic History*, vol. 73, p. 101273, ISSN 0014-4983.
- Bockstette, V., Chanda, A. and Putterman, L. (2002). 'States and Markets: The Advantage of an Early Start', *Journal of Economic Growth*, vol. 7(4), pp. 347–69.
- Bogart, D. and Chaudhary, L. (2016). 'Railways in Colonial India: An Economic Achievement?', in (L. Chaudhary, B. Gupta, T. Roy and A. Swamy, eds.), *A New Economic History of Colonial India*, pp. 140–60, chap. 9, Routledge.
- Brecke, P. (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, J. (1989). *The Sinews of Power: War, Money, and the English State, 1688-1783*, Harvard University Press.
- Cameron, A.C., Gelbach, J. and Miller, D. (2008). 'Bootstrap-Based Improvements for Inference with Clustered Errors', *Review of Economics and Statistics*, vol. 90(3), pp. 414–427.
- Castelló-Climent, A., Chaudhary, L. and Mukhopadhyay, A. (2018). 'Higher Education and Prosperity: From Catholic Missionaries to Luminosity in India', *Economic Journal*, vol. 128(616), pp. 3039–3075.
- Centeno, M. and Enriquez, E. (2016). *War and Society*, Polity Press.
- Chakrabarti, J. (1896). *The Native States of India*, Luzac.
- Chandler, T. (1987). *Four Thousand Years of Urban Growth: A Historical Census*, St. David's University Press.
- Chaudhary, L., Rubin, J., Iyer, S. and Shrivastava, A. (2020). 'Culture and Colonial Legacy: Evidence from Public Good Games', *Journal of Economic Behavior and Organization*, vol. 173, pp. 107–129.
- Clodfelter, M. (2002). *Warfare and Armed Conflicts: A Statistical Reference to Casualty and Other Figures, 1500-2000*, McFarland.
- Conley, T. (1999). 'GMM Estimation with Cross Sectional Dependence', *Journal of Econometrics*, vol. 92(1), pp. 1–45, ISSN 0304-4076.
- de la Garza, A. (2016). *The Mughal Empire at War*, Routledge.
- Dell, M., Lane, N. and Querubin, P. (2018). 'The Historical State, Local Collective Action, and Economic Development in Vietnam', *Econometrica*, vol. 86(6), pp. 2083–2121.
- Dincecco, M. (2017). *State Capacity and Economic Development: Present and Past*, Cambridge University Press.
- Dincecco, M., Fenske, J. and Onorato, M. (2019). 'Is Africa Different? Historical Conflict and State Development', *Economic History of Developing Regions*, vol. 34(2), pp. 209–250.

- Dincecco, M. and Wang, Y. (2018). 'Violent Conflict and Political Development over the Long Run: China versus Europe', *Annual Review of Political Science*, vol. 21, pp. 341–358.
- Docherty, P. (2008). *The Khyber Pass: A History of Empire and Invasion*, Union Square Press.
- Dutt, R. (1950). *The Economic History of India under Early British Rule*, Routledge.
- Fenske, J. and Kala, N. (2021). 'Linguistic Distance and Market Integration in India', *Journal of Economic History*, vol. 81(1), pp. 1–39.
- Fenske, J., Kala, N. and Wei, J. (2021). 'Railways and Cities in India', *Centre for Competitive Advantage in the Global Economy Working Paper 559*.
- Foa, R.S. (2016). *Ancient Politics, Modern States*, PhD Dissertation, Harvard University.
- Gaikwad, N. (2014). 'East India Companies and Long-Term Economic Change in India', *Working Paper, Columbia University*.
- Gennaioli, N. and Rainer, I. (2007). 'The Modern Impact of Precolonial Centralization in Africa', *Journal of Economic Growth*, vol. 12(3), pp. 185–234, ISSN 1573-7020.
- Gennaioli, N. and Voth, H.J. (2015). 'State Capacity and Military Conflict', *Review of Economic Studies*, vol. 82(4), pp. 1409–1448.
- Gerring, J., Ziblatt, D., Van Gorp, J. and Arévalo, J. (2011). 'An Institutional Theory of Direct and Indirect Rule', *World Politics*, vol. 63(3), pp. 377–433.
- Glaeser, E.L. and Shapiro, J.M. (2002). 'Cities and Warfare: The Impact of Terrorism on Urban Form', *Journal of Urban Economics*, vol. 51(2), pp. 205 – 224.
- Gommans, J. (1999). 'Warhorse and Gunpowder in India, c. 1000-1850', in (J. Black, ed.), *War in the Early Modern World*, pp. 105–28, chap. 5, Routledge.
- Gupta, B. (2019). 'Falling Behind and Catching up: Indias Transition from a Colonial Economy', *Economic History Review*, vol. 72(3), pp. 803–827.

Gupta, B., Ma, D. and Roy, T. (2016). 'States and Development: Early Modern India, China, and the Great Divergence', in (J. Eloranta, E. Golson, A. Markevich and N. Wolf, eds.), *Economic History of Warfare and State Formation*, pp. 51–69, chap. 2, Springer.

Hariri, J. (2012). 'The Autocratic Legacy of Early Statehood', *American Political Science Review*, vol. 106(3), p. 471494.

Heldring, L. (2021). 'The Origins of Violence in Rwanda', *Review of Economic Studies*, vol. 88(2), pp. 730–763.

Henderson, V., Storeygard, A. and Weil, D. (2012). 'Measuring Economic Growth from Outer Space', *American Economic Review*, vol. 102(2), pp. 994–1028.

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

India Institute of Forest Management (2015). *India State of Forest Report*.

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

Iyer, S., Shrivastava, A. and Ticku, R. (2017). 'Holy Wars? Temple Desecrations in Medieval India', *Working Paper, University of Cambridge*.

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

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

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

Joppen, C. (1907). *Historical Atlas of India*, Longmans, Green, and Company.

Kelly, M. (2019). 'The Standard Errors of Persistence', *Centre for Economic Policy Research Discussion Paper 13783*.



- Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S.E. and Sachs, J. (2004). 'A Global Index Representing the Stability of Malaria Transmission', *The American Journal of Tropical Medicine and Hygiene*, vol. 70(5), pp. 486–498.
- Klein Goldewijk, K., Beusen, A., and Janssen, P. (2010). 'Long-Term Dynamic Modeling of Global Population and Built-Up Area in a Spatially Explicit Way: HYDE 3.1', *The Holocene*, vol. 20(4), pp. 565–73.
- König, M., Rohner, D., Thoenig, M. and Zilibotti, F. (2017). 'Networks in Conflict: Theory and Evidence From the Great War of Africa', *Econometrica*, vol. 85(4), pp. 1093–132.
- Kotwal, A., Ramaswami, B. and Wadhwa, W. (2011). 'Economic Liberalization and Indian Economic Growth: What's the Evidence?', *Journal of Economic Literature*, vol. 49(4), pp. 1152–99.
- Lange, M. (2004). 'British Colonial Legacies and Political Development', *World Development*, vol. 32(6), pp. 905–22.
- Lee, A. (2019). 'Land, State Capacity, and Colonialism: Evidence From India', *Comparative Political Studies*, vol. 52(3), pp. 412–444.
- Levi, M. (1988). *Of Rule and Revenue*, University of California Press.
- Mann, M. (1984). 'The Autonomous Power of the State: Its Origins, Mechanisms and Results', *European Journal of Sociology*, vol. 25(2), pp. 185–213.
- Matsuura, K. and Willmott, C. (2009). 'Terrestrial Air Temperature: 1900-2008 Gridded Monthly Time Series', *Center for Climatic Research, University of Delaware*.
- Michalopoulos, S. and Papaioannou, E. (2013). 'Pre-Colonial Ethnic Institutions and Contemporary African Development', *Econometrica*, vol. 81(1), pp. 113–152.
- Michalopoulos, S. and Papaioannou, E. (2018). 'Spatial Patterns of Development: A Meso Approach', *Annual Review of Economics*, vol. 10(1), pp. 383–410.
- Min, B. (2015). *Power and the Vote: Elections and Electricity in the Developing World*, Cambridge University Press.

- Montalvo, J. and Reynal-Querol, M. (2005). 'Ethnic Polarization, Potential Conflict, and Civil Wars', *American Economic Review*, vol. 95(3), pp. 796–816, ISSN 00028282.
- Morris, I. (2014). *War! What Is It Good For?: Conflict and the Progress of Civilization from Primates to Robots*, Princeton University Press.
- Mukherjee, S. (2018). 'Colonial Origins of Maoist Insurgency in India: Historical Institutions and Civil War', *Journal of Conflict Resolution*, vol. 62(10), pp. 2232–2274.
- Nag, P. (2007). *Historical Atlas of India*, NATMO.
- Naravane, M. (1997). *Battles of Medieval India: AD 1295-1850*, APH Publishing.
- Nath, P. (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*, vol. 41(2), pp. 245–258.
- North, D. (1981). *Structure and Change in Economic History*, W.W. Norton.
- Nunn, N. and Puga, D. (2012). 'Ruggedness: The Blessing of Bad Geography in Africa', *The Review of Economics and Statistics*, vol. 94(1), pp. 20–36.
- Osafo-Kwaako, P. and Robinson, J. (2013). 'Political Centralization in Pre-Colonial Africa', *Journal of Comparative Economics*, vol. 41(1), pp. 534–564.
- Özak, Ö. (2010). 'The Voyage of Homo-Economicus: Some Economic Measures of Distance', *Working Paper, Brown University*.
- Özak, Ö. (2018). 'Distance to the Pre-Industrial Technological Frontier and Economic Development', *Journal of Economic Growth*, vol. 23(2), pp. 175–221.
- Qaisar, A. (1998). 'Distribution of the Revenue Resources of the Mughal Empire among the Nobility', in (A. Muzaffar and S. Subrahmanyam, eds.), *The Mughal State, 1526-1750*, pp. 252–258, chap. 7, Oxford University Press.
- Ramankutty, N., Foley, J., Norman, J. and McSweeney, K. (2002). 'The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change', *Global Ecology and Biogeography*, vol. 11(5), pp. 377–392.

- Ramusack, B. (2003). *The Indian Princes and their States*, Cambridge University Press.
- Rasler, K. and Thompson, W. (2005). 'War Making and State Making: Governmental Expenditures, Tax Revenues, and Global Wars', *American Political Science Review*, vol. 79(2), pp. 491–507.
- Raychaudhuri, T. (1982). 'Inland Trade', in (T. Raychaudhuri and I. Habib, eds.), *The Cambridge Economic History of India, Volume 1: c.1200-c.1750*, pp. 325–359, chap. 11, Cambridge: Cambridge University Press.
- Richards, J. (1995). *The Mughal Empire*, Cambridge University Press, second edn.
- Roy, K. (2011). *War, Culture, and Society in Early Modern South Asia, 1740-1849*, Routledge.
- Roy, M. (1994). *Politics, War, and State Formation in Early Modern India*, PhD Dissertation, Harvard University.
- Roy, T. (2013). *An Economic History of Early Modern India*, Routledge.
- Schwartzberg, J. (1978). *A Historical Atlas of South Asia*, University of Chicago Press.
- Stein, B. (1985). 'State Formation and Economy Reconsidered', *Modern Asian Studies*, vol. 19(3), pp. 387–413.
- Strayer, J. (1970). *On the Medieval Origins of the Modern State*, Princeton University Press.
- Thornton, J. (1999). *Warfare in Atlantic Africa, 1500-1800*, Routledge.
- Tilly, C. (1992). *Coercion, Capital and European States, AD 990-1992*, Oxford: Blackwell.
- Tollefsen, A., Strand, H. and Buhaug, H. (2012). 'PRIO-GRID: A Unified Spatial Data Structure', *Journal of Peace Research*, vol. 49(2), pp. 363–74.