Ancient Polities, Modern States

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Foa, Roberto. 2016. Ancient Polities, Modern States. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:26718768">http://nrs.harvard.edu/urn-3:HUL.InstRepos:26718768</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dashboard.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dashboard.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>

Ancient Polities, Modern States

A dissertation presented
by
Roberto Stefan Foa
to
The Committee on Degrees in Government
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Government

Harvard University
Cambridge, Massachusetts
January 2016
Ancient Polities, Modern States

Abstract

Political science is concerned with the study of polities. However, remarkably few scholars are familiar with the polities of the premodern era, such as Vijayanagara, Siam, Abyssinia, the Kingdoms of Kongo or Mutapa, or the Mysore or Maratha empires. This dissertation examines the legacies of precolonial polities in India, during the period from 1707 to 1857. I argue that, contrary to the widespread perception that the Indian subcontinent was a pre-state society, the late eighteenth and early nineteenth centuries were a time of rapid defensive modernization across the subcontinent, driven by the requirements of gunpowder weaponry and interstate warfare among South Asian regimes and against European colonial powers. These changes included the broadening and deepening of the tax base, consolidation of territorial control, reorganization of domestic militaries to use infantry and gunpowder weapons, rationalization of the administration through use of accounts and printed records, and the professionalization and functional differentiation of the executive branch.

I then trace the boundaries of precolonial eighteenth-century South Asian polities, in order to show that districts of India that lie narrowly within the boundary lines of historically centralized states perform significantly better today on a wide variety of district-level indicators of state effectiveness than those narrowly outside these boundaries, despite the fact that these borders largely ceased to exist in the early nineteenth century. These estimated effects are robust to a wide variety of controls, placebo tests for border displacement, the exclusion of individual polities, and controls for the boundaries of India’s contemporary federal states. I verify the persistent legacy of precolonial states using a combination of archival research, district-level colonial data on taxation and public goods from 1853 to 1901, and a field test of bureaucratic responsiveness conducted in the state of Karnataka. Using extensive archival research on the fiscal and bureaucratic structure of Indian states in the eighteenth century, I show that following the decline of the Mughal Empire, warfare
between “challenger states” prompted an accumulation of bureaucratic and fiscal capacity at the local level, and that this capacity has persisted through the colonial era to the present day. In contrast to “bottom-up” theories of state capacity which root institutional strength in societal characteristics such as ethnic homogeneity, social capital, or land equality, it is argued that government effectiveness is cumulatively built through long-term historical investments in state capacity, and that, in India, an important phase of investment occurred during the warring states period of the eighteenth century.

Finally, I show that this relationship exists beyond the South Asian context, both in cross-country regressions of the effect of state antiquity on contemporary state capacity, and by conducting a subnational historical analysis within districts of the Former Soviet Union. I conclude that augmenting the state’s power to tax, regulate, or conscript is, in Weber’s phrase, “a long and slow boring of hard boards”, and the resources required in order to attain a functioning state - bureaucratic infrastructure, norms of compliance, and affective loyalty - are accumulated only very gradually. Yet where long-extant political regimes were successful in monitoring, coercing, and mobilizing citizens towards state goals they generate a reservoir of legitimacy and compliance, that is essential for making states work in the world today.
Figure 0.1: India in the Late Eighteenth Century.

Figure 0.2: Land Tax per Capita, Districts of India in 1900-1.

Entries in the District Gazetteer: India Archives, London. Green indicates higher levels of taxation (4-12 Rs./capita), red lower levels (<1 Rs./capita).
Contents

1 Why Ancient States Matter 1
2 The Culture of Compliance 38
3 Precolonial States in South Asia 80
4 Testing for Border Discontinuities in Modern India 107
5 Trajectories of State Formation 157
6 The Puzzle of Persistent States Legacies Under Colonial Rule 207
7 The Global Perspective 239
8 State History and State Capacity in Russian Regions 279
9 Conclusion: Ancient Polities, Modern States 305

Bibliography 315
Acknowledgments

“Thus by the beginning of the fourteenth century, Italy had produced, not one, but two innovative patterns of governance with their associated social and cultural features” - Putnam, Making Democracy Work, 1993

“Hindostan is an Italy of Asiatic dimensions, the Himalayas for the Alps, the Plains of Bengal for the Plains of Lombardy, the Deccan for the Apennines, and the Isle of Ceylon for the Island of Sicily. The same rich variety in the products of the soil, and the same dismemberment in the political configuration.”

- Marx, ‘The British Rule in India’, 1853

This dissertation explores fundamental questions about institutional persistence and state capacity by studying district-level variation in India. It is written primarily for those who care about historical institutions and their effects on contemporary quality of government, but will also, I hope, prove informative and enjoyable for who share my interest in South Asia and its myriad differences in language, culture and religion.

The inspiration for this research lies in a place very far from South Asia, however, and also very far from Europe, where much of the historical literature, and my own prior knowledge, is based. Working in Central Africa in the year before my graduate studies, I had the opportunity to cross by land from Rwanda, in the African interior, to the North Kivu area of the Democratic Republic of Congo. Crossing the border from Gisenyi to Goma, I was struck by the contrast between the immaculate and orderly nature of the Rwandan side, where the guards are polite and efficacious, the streets, if rudimentary, free of litter, and the roundabouts sport planted gardens, and the Congo, where waste is ubiquitous, poverty conspicuous, and the police predatory. What could explain these sharp differences in public order? Rwanda’s strong “culture of compliance” was commonly-accepted fact among the international development community in Kigali, who reported positive impressions of the country’s government and civil service, and yet, morbid fascination at how so many ordinary citizens had been inducted into the recent genocide. Rwanda’s culture of compliance is rooted in longstanding social and political institutions, as a tradition of collective
service, for example, termed *Umuganda*, has existed since precolonial times, by which all residents of a neighborhood are required to engage a day a week in local tasks, such as paving roads, planting trees, and cleaning streets. Moreover, reading the diaries of the first European explorers to visit the kingdom in the late nineteenth century (collected in the *Études Rwandaises*), I found similar descriptions and anecdotes of the degree of public order in the realm, and the extent to which the royal court exercised surveillance over individual villages. The long history of Rwanda as a centralized kingdom, seemed linked to the unusual degree of compliance offered to its king and public authorities, which in modern times, despite several changes of regime, have transferred to the country’s civilian leaders. This led me to entertain a more general hypothesis of how stable, long-term state formation leads to the accumulation of state capacity and persistent cultures of compliance in the world today.

For a scholar interested in the legacy of precolonial polities upon contemporary social and political institutions, India, in many ways, presents a better grounds for hypothesis testing. The country has lived under the same constitutional order since 1947, but contains an exceptional diversity in precolonial institutions, exhibiting the full range from centralized bureaucratic polities, to patrimonial dynasties, to petty chieftaincies and tribal societies. If variation in historical state formation can explain differences in contemporary governance, we would observe it there. The idea seemed unique, moreover, as South Asia has also been largely invisible to comparative historical analysis: most studies of state formation are focused on the experience of early modern Europe, and it seemed a great neglect that no-one had yet conducted a rigorous comparative study of how the various historical polities of the Indian subcontinent - as rich and diverse in their own ways, as the shifting polities of Machiavelli’s Italy or Metternich’s Europe - led to the institutional variation that can be found in India today. As I make clear in the relevant parts of this dissertation, we find at work in precolonial South Asia many of the dynamics of state formation that authors such as Charles Tilly, Michael Mann, and Thomas Ertman have uncovered in Europe.

In pursuit of this project, in late 2012 I moved to Delhi in order to begin fieldwork. This was a long process of designing and piloting surveys, drawing up book lists, and coding India’s vast and intangible history into (perhaps equally vast though hopefully more tangible) datasets, the purpose of which must no doubt have seemed at times oblique to
those who assisted me along the way. As the project advanced through its various stages, there are innumerable people who have played a role, and deserve my thanks. I am grateful to David Clarance for matching the 2,601 ethnic groups of the Joshua Project database to the categories used in Murdock’s Ethnographic Atlas, and to Anandaroop and Sukhalata Sen for drawing up scores of books in search of any minor detail that might reveal anything about such esoteric topics as the tax arrangements of the Deccan sultanates, or the court structure of the Vijayanagara kings. I’m also grateful to Anuradha Singh and Shubh Sharma for advice on how to go about testing bureaucratic responsiveness through the Right to Information (RTI) act, and Nilanjana Goswami for help in sending hundreds of information requests to far-flung corners of India. Most of all, I’m grateful to Oeendrila Lahiri for dealing with my innumerable requests for help and who never lost patience with what must have seemed a never-ending stream of data gathering.

Innumerable people have also provided academic advice, mentorship, and suggestions throughout this project, and space is too short to list every person here. Among the scholars who I consulted at varying stages of this work, I would like to thank my advisors James Robinson, Steven Wilkinson, Ronald Inglehart and Daniel Ziblatt, as well as Robert Chase at the World Bank office in Delhi, Tirthankar Roy at the LSE, Philip Oldenburg at Columbia University, Christian Welzel, Pippa Norris, Bo Rothstein, Benedikt Herrmann and Francisco Herreros for suggestions on the theory of compliance, Alexander De Juan, Didier Queralt and Martin Lodge for comments at the meetings of the American Political Science Association, comments and feedback from Anjali Bohlken, Ajay Verghese, John Marshall, Shivaji Mukherjee, Jacob Gerner, Ijaz Gilani, and Rengin Firat, the support of my peers in graduate life, Aditya Dasgupta, Yascha Mounk, Emily Clough, Camille Sindhu, Daniel Koss, Tristan Reed, and Jamie Loxton, as well as the staff at the India Archives in New Delhi and in at the British Library in London, and for feedback from participants in seminars at the Quality of Government Institute in Gothenburg, the Higher School of Economics in Moscow, at Yale University, at the European Union Joint Research Centre, and participants in the South Asia and Comparative Politics seminars at Harvard.

For financial and logistical support in this project I thank the South Asia Initiative at Harvard University, the Himalaya Foundation, the Laboratory for Comparative Social Research (LCSR) and Higher School of Economics (HSE), the Centre for the Study of De-
veloping Societies (CSDS) in Delhi, the Jawaharlal Nehru University, as well as Harvard University. I am also grateful to the many friends and colleagues who hosted me and tolerated my rather laconic presence while I was engaged in writing up the various sections of this work: especially to Miguel and Tatiana Basáñez in Cambridge, Nadine Feyder in Addis Ababa, Alec Wildenstein in Kenya, Hannah and Shashank Joshi in London, and to my brother Daniel in China. Their support has made it possible to complete this manuscript, while responsibility for the errors is mine alone.
Chapter 1

Why Ancient States Matter

No great thing is created suddenly, any more than a bunch of grapes or a fig. If you tell me that you desire a fig, I answer you that there must be time. Let it first blossom, then bear fruit, then ripen.

Epictetus - Discourses Chap. xv.

The science of government being therefore so practical... it is with infinite caution that any man ought to venture upon pulling down an edifice, which has answered in any tolerable degree for ages the common purposes of society.

Burke - Reflections on the Revolution in France

The Himalayan kingdom of Bhutan is an isolated, agrarian economy that introduced television for the first time in 1999, yet has a corruption perceptions rating superior to South Korea, and a homicide rate lower than Finland\(^1\). Situated 900 miles to the west, Pakistan was among the most economically developed regions of South Asia following decolonization, yet has partitioned twice, experienced multiple coups, and is plagued by

\(^{1}\)The homicide rate for Bhutan was registered as 2.78 per 100,000 by the UN statistics division in 1999; for Finland, in 2000 the rate was marginally higher, at 2.83 (UNCJIN 2008).
terror and corruption\textsuperscript{2}. Turning to Africa, the kingdom of Lesotho is a small, landlocked country with a nominal income per capita of just $1,075 that raises over forty per cent of national income in general taxation\textsuperscript{3}, while the Republic of Cote d’Ivoire, among the most developed regions of Africa in the 1970s, was subsequently to descend into civil war, pervasive criminal violence, and state failure\textsuperscript{4}.

How is it that small, poor, and isolated societies can succeed in building strong and cohesive states, while more economically advanced polities frequently experience state stagnation and collapse? The question of the origins of political order is one of the fundamental concerns of political science, and has attracted many responses since it was posed by Samuel Huntington in the 1960s. In the early post-war years, the dominant paradigm in social and political thought was modernization theory, according to which, post-independence economic development would serve as a “motor” to produce beneficial social and political changes such as greater literacy, political mobilization, and programmatic political ties (Deutsch 1961; Rostow 1960; Kothari 1970). As societies developed economically, individuals would become more materially secure, educational levels would increase, patrimonial ties to family, ethnicity or caste would erode (Lerner 1958; Inkeles and Smith 1974; Inglehart 1977). This would in turn produce political changes, as mass education would facilitate the recruitment of a modern bureaucracy, and improvements in transportation and communications facilitate the functioning of large, impersonal organizations (Parsons 1972; Bell 1973).

Already by the time that Samuel Huntington was writing in the late 1960s to mid-1970s, the experience of the post-colonial era had shown initial levels of economic development as a poor predictor of the subsequent impartiality, transparency and efficacy of public administration. Countries like South Korea, which had started in the 1950s as among the poorest in the world, were forming cohesive, bureaucratic, and developmental states, while others that began with a significant economic advantage, such as Algeria, Cote d’Ivoire,
Development of Bureaucratic Quality and Preceding 20-Year Growth (r = -0.17)

Development of Bureaucratic Quality and Initial GDP per Capita (r = -0.20)

Figure 1.1: Economic Development in 1984 (World Bank 2014) and Subsequent Changes in Ratings of Bureaucratic Capacity, 1984-2005 (International Country Risk Guide). Developing countries (GDP per capita in 1984 of $10,000 or less) only.

Algerian GDP per capita (at PPP) was $4,843 in 1975, well ahead of South Korea at $3,693, while Papua New Guinea had GDP per capita of $2,273 and Cote d’Ivoire of $2,432.

or Papua New Guinea, were failing to do so⁵ (Kohli 2004; Evans 1995). By the time that international financial organizations were abandoning broad development interventions in the 1980s to focus on providing incentives for policies and institutional reforms, the non-correlation of economic and political development were clear. From this point forward, time-series data shows no relationship between initial economic conditions and subsequent improvement or deterioration in ratings of bureaucratic capacity. Countries that by the 1980s were at a medium-income level of development, such as Venezuela, South Africa, and Thailand, were to subsequently deteriorate on such assessments, while very poor countries such as Bangladesh and Mali were to improve; nor was there to be any association between economic growth since independence and subsequent changes in public administration (Figure 1.1). While modernization theories have been validated as theories of societal change, their predictive power in the domains of bureaucratic development or state formation are less clear (Inglehart and Welzel 2005; Inglehart and Norris 2004; Welzel 2013).

⁵Algerian GDP per capita (at PPP) was $4,843 in 1975, well ahead of South Korea at $3,693, while Papua New Guinea had GDP per capita of $2,273 and Cote d’Ivoire of $2,432.
By the 1990s, as modernization theory had lost its appeal and structural adjustment reforms were being frustrated by absence of political commitment and capacity, another answer common among western foreign policymakers, international organizations, and media commentators was that absence of reform initiative stems from a deficit of political legitimacy, that is best enhanced through a democratic mandate (Rodrik 2000; Diamond 1999). Whereas modernization theories had seen representative government as the crowning event of a development sequence that begins with economic growth and passes through the spread of literacy, civic activism, and the formation of a stable middle class, instead policymakers came to advocate competitive, multiparty elections as a solution for countries facing diverse problems such as endemic corruption, state fragility, and the absence of infrastructure or welfare provision. In contexts as diverse as Afghanistan, East Timor, Kosovo, and the Democratic Republic of Congo, the international community has seen the organization of competitive, multiparty elections as the central objective of its intervention, and the prime condition for its eventual withdrawal.

Yet in the last decade, the disappointing experience of state-building in post-conflict democracies, as well as with the low quality of governance in many transitional regimes, has led to a reassessment of the relationship between democratic legitimation and acceptance of the state’s governing prerogative (Fukuyama 2005). Convergence upon the liberal democratic model has not seen a corresponding convergence in the state’s ability to secure compliance of administrative officers, or broader social compliance through the payment of taxes, obedience to laws, or the administration of territory (Figure 1.2). Across developing countries, there is no clear correlation between democratic institutions and measures of state capacity, nor any discernable increase in state capacity over the course of countries’ democratic consolidation (Figure 1.3). And many authoritarian and hybrid regimes continue to exhibit relatively high levels of state capacity, both in comparison with democratic regimes, where governments should a priori command greater legitimacy and authority through their electoral mandate.

This Project

This study follows the “institutional turn” that has led scholars towards historical ex-
Figure 1.2: Non-Correlation of Bureaucratic Quality and Combined Freedom House Democracy Ratings (r-square = 0.019). Sample: developing countries only, with population above 10m. Bureaucratic quality index composed from: assessment of meritocratic recruitment; civil service career opportunities; and bureaucratic salary and compensation, taken from the Evans and Rauch (1999) dataset; a measure of “professional public administration” taken from Teorell, Dahlstroem and Dahlberg (2011), which combines meritocratic recruitment, nepotism, and bureaucratic autonomy; and World Bank ratings on the quality of public administration and the quality of project budgetary management from the Country Policy and Institutional Assessment (CPIA) exercise (World Bank 2013).

planations of persistent state weakness (North 1993; Fukuyama 2014). However, departing from the general view that the “legitimacy deficit” of developing countries is a legacy of western colonialism, I argue that this deficit instead stems from the strength or weakness of precolonial institutions. The general view, represented by scholars as diverse as Wallerstein (1974, 1982), Acemoglu et al. (2002), Lange et al. (2006), Mahoney (2010), Kohli (2004) and Mamdani (1996), is that European colonial rule left the developing countries of Africa, the Americas, and Southeast Asia with illegitimate institutions, predatory elites, and an inability to enforce the state’s monopoly of coercion beyond a narrow metropolitan
base, and that, by extension, these institutions have persisted despite changes in the formal regime. Wallerstein (1974) and world-systems theorists, for example, have argued that colonial regimes were designed to serve as providers of natural resources to countries at the core of the world economy, leading to their narrow economic specialization, absence of governing legitimacy, narrow tax base due to the underinvestment in public goods, and dependency upon governing and administrative resources provided by western companies, governments, and multilateral agencies (Wallerstein 1980, 2004). As a result, post-colonial elites faced few incentives to invest in state capacity, instead relying upon the administrative, military, and technical support of western governments and organizations in order to manage their economies, face security challenges, and suppress the threat of domestic revolt (Verschave 1998). A similar line of argument has been pursued by Acemoglu, Johnson and Robinson (2002), who maintain that areas ruled but not settled by Europeans were subject to “institutional reversal” as European powers established predatory regimes to exploit labor and natural resources. Because these colonial regimes sought only to extract and faced few external threats, they did not invest in building the fiscal and regulatory capacity required to provide universal rule of law, deliver public goods, or engage in interstate

Figure 1.3: Declining Bureaucratic Quality (ICRG Index) in New Democracies: South Africa, Brazil, Albania. Year of Democratic Transition Marked.
warfare. The deep path dependencies created by such patterns of institutional development cannot easily be corrected; though post-colonial states could be forced to reduce their dependency on western support, or aid could be made conditional upon the implementation of specific policy and institutional reforms (Easterly 2001, 2006; Burnside and Dollar 2000).

Instead, I argue that patterns of state formation found in non-western countries today stem from persistent legacies that predate the colonial era. In the world as a whole, states at the highest level of political development at the point of contact with European powers, such as Japan, China or Turkey, were strengthened by this interaction as they responded through defensive modernization: these states “persisted,” or remained independent, and continue to exhibit the highest levels of state capacity today, as measured by such conventional metrics as fiscal depth, the rule of law, or the quality of the bureaucracy. However polities that were sufficiently developed by the early modern era to offer opportunities for labour-extractive practices such taxation and mining, but not sufficiently developed in their fiscal or military infrastructures to resist foreign rule, such as Mesoamerica or Bengal, were by contrast the first to be targeted by European colonialists in the sixteenth and seventeenth centuries, while a third group of polities were created entirely under colonial administration in areas in which there was little or no prior indigenous state capacity, and while these territories ultimately benefited from colonial state-building, they remain weak states today primarily as a result of their “late start” to the state formation process.

I show that this pattern can be found, not only among post-colonial regimes in the world as a whole, but also within a major subregion subject to colonial influence: the South Asian subcontinent. In the Indian subcontinent, precolonial regimes ranged from pre-state hunter-gatherer societies in the interior northeast to centralizing, proto-bureaucratic states in the Deccan. Where historically centralized regimes had ruled in centuries past, I find that citizens today exhibit greater willingness to comply with the law, cede revenue in taxes, and fulfill public service duties, and thereby enable the attainment of a higher degree of state capacity. I test for such enduring historical legacies by looking for regression discontinuities along precolonial borders, conducting archival research to find evidence of historical persistence, and implementing a bureaucratic responsiveness test in the field to detect evidence of border differences in bureaucratic capacity. Differences in state effectiveness can be identified at the borders of precolonial polities, despite the fact that such boundaries
have not existed, in many cases, since at least the colonial era.

I find that such discontinuities cut through not only the arbitrary borders imposed by colonial administrators, but also the ethnolinguistic states constructed after Indian independence. For example, in the Indian state of Tamil Nadu (Figure 1.4), the districts of Kanyakumari and Tirunelveli are both predominantly ethnic Tamil, belong to the same federal entity, and have achieved a similar level of economic development, with average household incomes of $809 and $735, respectively\(^6\). Yet since precolonial times, Kanyakumari was situated within the domain of Travancore, an indigenous state that ruled much of southern Kerala, while Tirunelveli was under independent Tamil nayaks. A generation after independence, Kanyakumari is overwhelmingly more similar to former Travancore domains across state lines in Kerala than to its Tamil counterparts in Tamil Nadu. There is greater public order, as in Kanyakumari, there were just 5 reported robberies per 100,000 (compared to 23 per 100,000 in Tirunelveli) and 4 cases of rape (compared to 36 per 100,000)\(^7\). And local government seems better able to ‘deliver the goods’, as in Kanyakumari, 41 per cent of villages in the 1980s had access to a hospital, compared to just 4.5 per cent in Tirunelveli, and 82 per cent access to high schools, compared to just 19 per cent in the neighboring Tamil district\(^8\).

Tellingly, the same “Travancore effect” cuts through the state of Kerala itself, as former Travancore districts in the south have a 30 per cent rate of villages with access to hospitals and 91 per cent rate with access to high schools, against 12 and 73 per cent, respectively, in

---

\(^6\)Data on household income from the Maryland Human Development Survey (2005), and converted into dollars at a nominal exchange rate of 45 to one dollar (2005 average).

\(^7\)Furthermore, these differences existed also in 1900. Average health spending per 1,000 in 1900 was Rs. 30 in Tirunelveli district, compared to Rs. 92 per 1,000 across the border in Travancore.

\(^8\)Other examples are widespread. For example, Kurnool and Mahbubnagar are two districts in the federal state of Andhra Pradesh, yet in precolonial times Kurnool was a longstanding subject of the rising Mysore Empire, while Mahbubnagar was part of the declining Mughal realm. Today, Mahbubnagar has become more prosperous, with household income of $1,280, against $1,026 in Kurnool. Yet in common with other former Mysore provinces in Karnataka, Kurnool exhibits substantially greater evidence of state capacity. Not only is Kurnool a great deal safer, with half the rate of highway robberies, and a homicide rate of just 1 per 100,000, against 14.5 in Mahbubnagar. Additionally, Kurnool also shows a much stronger record of public goods delivery, with 75 per cent of villages having paved roads, against 45 per cent, and 16 per cent with primary health centers, versus 4 per cent in the non-Mysore Mahbubnagar.
the north\(^9\). Such differences are evident not only in contemporary data, but also in historical measures, taken over a century ago (Figure 1.5).

How can we explain why differences in state effectiveness persist along the boundaries of precolonial polities? Theories based on ethnolinguistic fractionalization such as Alesina et al. (2003) or Keefer and Khemani (2004) are of limited utility in explaining

\(^9\)These differences were even more pronounced in the 1900 data, in which districts of Travancore were well ahead of Malabar in most categories of public goods.
differences between districts that contain identical ethnic and linguistic groups. Studies of colonial legacies (e.g. Iyer 2010) cannot explain differences between regions that shared administrative boundaries under colonial rule, while arguments that focus on the performance of post-independence federal states (e.g. Kohli 1989, Singh 2011) cannot explain divergences among districts within federal subunits. Classical theories based on the role of democratic legitimation (Bendix 1978; Diamond 1999) may be of limited use in explaining differences between districts belonging to a single electoral democracy, while modernization approaches such as Inglehart and Welzel (2005), that attribute democratic quality to values arising from economic development, are of limited applicability where districts are similar in respect of income and urbanization.

What is needed is a theory that, in accordance with the arguments of Skocpol (1979), Mann (1986) and Tilly (1992), “puts states first” - by explaining how legacies of state formation from the historical past restructured citizen-state ties at a local level, generated norms of compliance with public authority, and developed the state’s infrastructural power. Rather than attribute the state’s capacities to features of the social environment, such as ethnic homogeneity (Alesina et al. 2004), social capital (Putnam 1993), or patterns of historical inequality (Engerman and Sokoloff 2005), this dissertation therefore proposes a theory of endogenous state formation, whereby rachets in state capacity are seen as the driving process, which then in turn reshape social, economic, and broader political institutions in ways that are evident generations later (Pierson 1992; Skocpol 1979; Tilly 1992; Ertman 1997).

Contemporary India is the perfect grounds for testing such an argument. The gold standard of quantitative methodology is the randomized controlled trial, in which experimenters choose an intervention and apply it at random, and then, after some period of time, compare the “treatment” group that received the intervention to the ‘control’ group that did not (Diamond and Robinson 2010). By imposing an interstice between the period of treatment and the period of observation, the colonial experience in India offers a “natural” experiment. In the centuries preceding colonial rule India was host a remarkable diversity of states and polities, from pre-state hunter-gatherer societies in the interior uplands and the mountainous northeast, to the centralized, proto-bureaucratic states of the southern Deccan. Yet many of India’s major precolonial states, such as the Maratha Empire, Awadh,
or the Sikh Empire ceased to exist entirely following the East India Company’s consolidation of power from 1757 to 1856. Those polities such as Mysore or Sikkim which survived as rump “princely states” were shorn of territory, and nine years after independence, their remaining territorial boundaries scrambled into a new system of ethnolinguistic states. The result is that the political boundaries of the post-independence era have scarce resemblance to those of the colonial or precolonial past, and spatial identification of historical regime effects is relatively untainted of post-treatment bias.\(^{10}\)

I argue that we can explain such contemporary district level variation by examining the legacies of precolonial states in India, and specifically how centralizing states in the eighteenth century sought to enhance their local infrastructural control by raising taxes, soldiers, and surveillance of subject populations. As maintained by authors such as Elias (1982) or Scott (2005, 2009), states are agencies that seek to render society legible so that they can more easily conscript manpower, tax resources, and enforce behavior in pursuit of military or extractive goals. India’s precolonial regimes conducted land cadastres, raised revenue demands, replaced local intermediaries with appointed officials\(^{11}\), and extended the state’s surveillance and enforcement powers deeper by assuming policing and judicial responsibilities directly. The upshot of these socially penetrative measures was that public order, if it did not trump alternative sources of legitimacy, such as caste, family and village, at least became a viable alternative, and generated a compliance “equilibrium,” in which cooperation with, rather than resistance to, the state became the accepted behavioural default (Levi 1998; Rothstein 2005; Herreros 2004). By establishing sovereign authority as legitimate, that is, having the quality of being viewed as being entitled to be obeyed (Tyler 1990), precolonial regimes thus laid the foundation of contemporary state capacity. Where officers of state perceive public authority as legitimate then they will comply with its orders and enforce its laws, and, where the public do so, they obey these voluntarily.

In making the argument that contemporary patterns of state capacity reflect the lega-

\(^{10}\)Taking districts within a single nation-state also holds constant many of the variables that would confound cross-country analysis: these regions having been subject to a common first-past-the-post electoral system, democratic regime, judicial authority, and sixty years of nationwide policies and legislation (Snyder 2001).

\(^{11}\)In this way, the Mysore Empire for example “rid Karnataka of many warring minor principalities” such as “Doddaballapur, Chikballapur, Sira, Harapanaballi, Ballery, Ramdurga and others”; Karnataka State Gazetteer.
cies of historical states I build upon a large and growing scholarship in political science. By examining local variation in political legacies, this study contributes to the historical institutionalism of Acemoglu and Robinson (2001, 2002, 2011), and specifically, attempts at identifying spatial variation in historical regime legacies such as Becker et al. (2014), Dell (2010), Iyer (2010), Acemoglu et al. (2011) and Michaloupoulos and Pappaiannou (2013). More generally, by choosing to focus on the origins of state capacity, I build on Huntington’s (1968) distinction between the “degree” versus the “type” of governance, and more recently, analyses of the historical origins of states worldwide, such as Fukuyama (2011, 2013) and the works of Victor Liebermann (2008, 2010). In arguing for the merits of indigenous institutions, I concur with the large literature on the negative effects of colonialism, including Iyer’s (2010) arguments concerning the effects of direct and indirect rule in India, Mamdani’s (1996, 2007) arguments regarding colonial rule in Africa. By focusing on historical states, this study echoes Lieberman’s (2010) arguments concerning the importance of cohesive elites for building state capacity, and Kohli’s (2004) thesis regarding the importance of elite cohesion for state-directed development. Finally, in seeking to root contemporary state capacity in historical junctures such as wars and revolutions, rather than in aspects of economic or social structure, this study echoes the “state-centered” approach of Skocpol (1979, 1992), Tilly (1992), and Mann (1986). The South Asian case provides excellent external validation of the relationship between historical war-making and the accumulation of state capacity made in the European context by Tilly (1992), Hintze (1975) and Roberts (1956), as in precolonial India as in Europe, centralizing eighteenth-century regimes invested in building gunpowder armies, conducting land cadastres, and displacing local powerholders with centrally appointed bureaucrats (Roy 2013b; Stein 1996).

**Defining State Capacity**

What does it mean for a state to exercise “capacity”? State capacity is often understood in general terms as the ability of a government to “administer its territory effectively” (Skocpol 1985), or the extent to which “the government governs” (Huntington 1968). More concretely, state capacities can be further divided into specific functional prerogatives, such as extractive or fiscal capacity (the ability to raise taxes), coercive capacity (the ability to
enforce laws), and enlistment capacity (the ability to raise armies and defend territorial boundaries) (Mann 1986; Levi 1988; 1998). Because these abilities are “capacities” they reflect a power to act, rather than a specific outcome: state capacity is a “counterfactual conditional” such that if authorities were to propose a new tax or law, they would have the means to raise such revenues or enforce such conduct, regardless of whether they do, in fact, do so (Fukuyama 2013).

States are able to exercise these capacities because there is compliance, or the willingness of citizens and agents of the state to engage in voluntary cooperation with the demands of sovereign authority. Where officers of state comply with higher authorities, they implement official edicts rather than use official time and resources for private ends, and where citizens comply with agents of state, the state is able to raise taxes and enforce laws. Compliance can therefore be categorized into its component forms, such as fiscal compliance, or civic acceptance of the state’s revenue demands, bureaucratic compliance, which refers to the willingness of civil servants to follow official protocol and superior orders, and regulatory compliance, which refers to the acceptance of citizens that they follow public laws and regulations. Pervasive non-compliance across these areas entails that states are unable to mobilize revenue, illegality is widespread, and officials will use their office to pursue private gains.

The Importance of State Capacity

The recent rise of interest in the study of state capacity has been driven by both academic and policy considerations. The persistence of state attributes, such as fiscal depth or the bureaucratic capacity, across changes of regime necessitates a distinct approach to the study of the state as an organization, and in particular a distinction between the state as a set of institutions surrounding the bureaucracy and government, and the regime, as a set of institutions for selecting political leaders. Jensen (2013) and Bagenholm (2014), respectively, show how norms of non-corruption in twentieth-century Denmark and Sweden have their origins in reforms initiated under nineteenth-century absolutist monarchies, while Lieberman (2005) for example has shown how fiscal structures developed under the apartheid regime in South Africa have persisted through to the post-apartheid era. Numerous studies, meanwhile, have shown how institutions developed under colonial administration have
continued under post-colonial regimes, such as Kohli’s (2004) argument concerning bureaucratic development in the former Japanese colonies in Asia, or Herbst’s (2000) attempt to trace the origins of the limited, “metropolitan” state of many contemporary African polities.

There are also strong practical reasons for studying state capacity. While state capacity can be used to wage wars, acts of politicide or genocide, reinforce dictatorship, and implement wasteful schemes (Gurr and Harff 2003; Gerner 2012; Bauman 1989; Scott 1999), it is also a precondition for providing public order, delivering public goods, redistributing income, and enforcing property rights (Acemoglu and Robinson 2012; Boix 2004). By the opening of the twenty-first century, however, policymakers have largely neglected state-building, in order to focus on areas such as civil society development and democracy promotion (Fukuyama 2005). In contexts as diverse as East Timor, Kosovo, and the Democratic Republic of Congo, the international community has seen the organization of elections, rather than the organization of states, as the central objective of its intervention, and the prime condition for its eventual withdrawal. However while democratic institutions may increase the incentive for politicians to engage in redistributive policies and service delivery, they do not necessarily provide the capacity to deliver on such promises. The “third wave” of democratization of the late twentieth century has seen scores of new democracies, from South Africa, to Brazil, to Indonesia or Mexico, in which leaders have entered office with ambitious plans for social reform only to confront vested societal opposition, absence of bureaucratic capacity, and the widespread abuse of program funds. Time-series indicators suggest that the experience of democratic transition in former communist countries of Central and Eastern Europe has led to marginal if any improvement over time in control of corruption, rule of law, and public service delivery (and in some cases, declines), while emerging democracies such as South Korea and Taiwan, that currently perform well on such measures, had already performed well before the transition. The success of autocratic states in the Arab Gulf, Singapore, and Hong Kong in delivering welfare, public order, and infrastructure has called into question whether democracy is a necessary condition for such outputs, or has simply appeared covariant due to the tendency for well-governed states to develop economically and later experience pressures for democratic transition (Inglehart and Welzel 2005; Ross 2006).
The state-building scholarship encourages policymakers to lend greater attention to the capacity of elites to deliver, in addition to the incentives or disincentives they face as a result of electoral institutions (Fukuyama 2013). Within the scholarly literature there is now a growing acceptance that components of government effectiveness, such as the control of corruption, provision of the rule of law, or political stability, correlate only imperfectly with measures of democracy, in particular in developing countries (Norris 2012). Many recent studies show state capacity metrics to be more useful than measures of democracy in explaining outcomes such as public goods provision, law and order, or civil conflict.\footnote{Raby and Teorell (2010) for example show that measures of good governance are stronger in predicting the absence of violent interstate conflict than are measures of democracy, while Gilley (2006, 2009) finds that state capacity dimensions of governance (rule of law, control of corruption and government effectiveness) have a stronger impact on political legitimacy than democratic rights or welfare gains.}

**State Capacity in India**

As a large, developing democracy, India is a perfect example of a country in which absence of local capacity has, from the start, hampered the ability of political elites to deliver upon policy objectives. Shashi Tharoor (2007: 214) describes the Indian model as one of “centralised policy-making and decentralised policy administration,” insofar as ambitious social or economic objectives are determined centrally, but implementation is left to individual states and local governments, according to their capacity and commitment. Thus already in the 1960s Myrdal (1968) had observed how in India laws are left “intentionally permissive” and that while “social legislation points the direction in which society should travel” it leaves “the pace indeterminate.” The result is that while the Indian constitution guarantees wide-ranging rights to education, nutrition, and healthcare, today there is widespread variation in the extent of actual delivery. Banerjee, Deaton and Duflo (2004) show that in a sample of local health facilities (sub-centers and primary health clinics) in Rajasthan, only around half of salaried medical staff were attending work during stated hours of operation, and even after intervening with incentives, this rate was increased only to two-thirds. Audits of social programs, such as the recently introduced National Rural Employment Guarantee Act (NREGA), show rates of local wage corruption range as high as 30 per cent (Imbert 2013; Dreze 2014). More direct indicators of state capacity, such
as bureaucratic responsiveness or fiscal mobilization, show similar variation and shortfalls. Across India as a whole for example, revenue from income tax equates to just 2.8 per cent of gross domestic product (RBI 2010). Audits of compliance with bureaucratic requirements, such as the 2005 Right to Information Act, show widely varying rates of compliance, with officers in some states ignoring the majority of information requests. The inability of central government to exercise local authority in large areas of India has led Lant Pritchett (2009) to describe India as a “flailing state,” in that “the head, that is the elite institutions at the national (and in some states) level remain sound and functional” but “this head is no longer reliably connected via nerves and sinews to its own limbs.” A map of state capacity in today’s India would be marked by large gaps, reflecting zones in which central state authority is simply unable to exert effective control over societal actors or its own salaried officers.

![Colonial-era Extractive Capacity in India – Land Taxation per Capita, 1900-1. Source: District Gazetteer.](image1)

![Contemporary Public Order in India – Homicide, 1980-90. Source: Iyer (2010).](image2)

Figure 1.6: District-Level Measures of State Capacity, 1900 and 1990.

Figures 1.6 offer such a map using two summary measures of state capacity across India’s 640 districts: a historical measure of fiscal capacity (land taxation per capita) in 1900, and a contemporary measure of public order, the district-level homicide rate in the period 1980-90. What is perhaps immediately apparent is an enduring divide between the
North of India, where state authority is persistently weak, and the South, where it is much stronger. The northern, Hindi-speaking states that comprise Rajasthan, Madhya Pradesh, Uttar Pradesh, Bihar and Jharkhand - together, a group uniting some 475m individuals, or 6.6 per cent of the world’s population - are disproportionately India’s weakest governed and most unsafe areas. They form a sharp contrast to both the southern, “Dravidian” states of Kerala, Tamil Nadu, and Karnataka, and the intermediary belt of non-Hindi yet Sanscrit-language based states that run across the center of the country, from Gujarat through Maharashtra and Odisha. Moreover, these patterns are deeply historically rooted, and other indicators of historical state capacity in 1900 - such as the density of police, local military recruitment, provision paved roads, or per capita health or education spending - also exhibit a clear north-south divide\textsuperscript{13}.

\textit{Compliance}

Just as there are sharp differences in output-based measures of state capacity between the North and the South of India, so too are there sharp differences in survey measures that reflect the prevalence of norms of \textit{compliance} with public authority. Table 1.1 shows survey responses of individuals across three regions of India to a range of questions, designed to reflect deeper differences in norms and patterns of compliance. Respondents in North India are far more likely to condone “anti-social” acts such as cheating on taxes, avoiding fares on public transport, or acts of interpersonal theft. Whereas just in 1 in 6 respondents in southern India thinks it acceptable to “claim benefits to which one is not entitled,” almost half of northern Indians do. Moreover, almost a third of North Indians report having carried a weapon out in public, compared to just 5 to 10 per cent of respondents in other parts of the country (World Values Survey 2014). The divide between the North and the South of India goes beyond political institutions, and reaches into the basic norms and expectations of personal conduct; in terms of the relationship of the citizen to public authority and the state.

Such differences in patterns of authority compliance between the North and the South of India can be seen in other outcome-based measures of state capacity, such as levels of

\textsuperscript{13}District-level measures for each of these variables were extracted from the records of the India Archives, and mapped to contemporary district boundaries.
<table>
<thead>
<tr>
<th>Item</th>
<th>North India</th>
<th>South India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent feels there are circumstances in which it can be ‘justifiable’ to:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘claim government benefits to which you are not entitled’</td>
<td>46.0%</td>
<td>15.1%</td>
</tr>
<tr>
<td>‘steal other people’s property’</td>
<td>28.1%</td>
<td>11.5%</td>
</tr>
<tr>
<td>‘cheat a fare on public transport’</td>
<td>35.4%</td>
<td>15.5%</td>
</tr>
<tr>
<td>‘cheat on taxes if you have a chance’</td>
<td>32.8%</td>
<td>15.6%</td>
</tr>
<tr>
<td>‘take a bribe in the course of one’s duties’</td>
<td>24.9%</td>
<td>13.3%</td>
</tr>
<tr>
<td>Respondent has ‘carried a knife or other weapon in public’</td>
<td>31.4%</td>
<td>9.0%</td>
</tr>
<tr>
<td>Respondent: is not ‘like me’ to ‘always behave properly;’ to avoid doing anything people would say is wrong’</td>
<td>22.2%</td>
<td>15.3%</td>
</tr>
</tbody>
</table>

Table 1.1: Survey Indicators of Compliance Norms, by Region of India. Notes: World Values Survey, wave 6 (conducted 2012-3). “North India” includes Uttar Pradesh, Madhya Pradesh, Rajasthan, Bihar and Jharkhand. “South India” includes Kerala, Karnataka, Tamil Nadu, and Andhra Pradesh.

crime, the degree to which public officials are perceived to engage in graft, measures of judicial effectiveness, or the numbers of pending criminal charges among elected representatives. These indicators are shown, aggregated at the state level, in Figure 1.7, which shows the large, deep, and persistent differences in state capacity and compliance between broad regions of the country.
Legacies of Precolonial States

What explains patterns of persistent variation in contemporary state capacity across India? Having laid out the puzzle of uneven and persistent gaps in public compliance, in this section I will set out how understanding precolonial patterns of state formation can help to account for a divergence in state capacity by the onset of the colonial era.

State Legacies of Precolonial States

In the early historical and sociological literature, there is a tendency to view India as stateless until foreign arrival. Marx (1972: 35) for example wrote of how “Hindustan, when not under the pressure of the Mohammedan, or the Mogul, or the Briton” was “dissolved into as many independent and conflicting States as it numbered towns, or even villages” and colonial administrators and authors like Henry Maine described India as a sea of village communities, each a separate “little republic” (Dumont 1980: 158-9; Maine 1871: lecture IV). This view of India as recurring from village to empire to village, is also found in
twentieth century political sociology, with Barrington Moore describing Indian society as “like the starfish whom fishermen used to shred angrily into bits, after which each fragment would grow into a new starfish” (Moore 1966: 458).

Yet the British East India Company did not encounter India as a sea of villages, but rather as a set of hardening states which it had to defeat sequentially during more than one century of warfare. The New Cambridge History of India classifies Indian regimes of the eighteenth century into three categories: (i) successor provinces of the Mughal Empire, such as Awadh, Bengal, and Hyderabad, in which a Mughal regional governor and his officials adopted the territory as their personal landed domains; (ii) “antique” micro-polities such as the Rajputs and the Tamil poligars, i.e. local kingdoms and principalities with limited fiscal or regulatory authority, and (iii) new and expanding “warrior states” on the Mughal periphery - notably Mysore, the Marathas, and Travancore - which were centralizing, proto-bureaucratic regimes with extensive military apparatuses, and which challenged both Mughal and European claims to supremacy over the subcontinent (Ramusack 2004). Such polities were not static but undergoing very substantial processes of political rationalization: the increasing centralization of political authority, professionalization of the military and civil services, and the deepening and widening of the tax base (Stein 1984).

Countering the colonial narrative of India’s past as one of “stagnation” or decline, recent histories of the subcontinent in the eighteenth century have drawn attention to the rapid defensive modernization of these regional states, leading Burton Stein to conclude that “centralization was the signature of regimes of eighteenth century India” (Stein 1984: 391; Bayly 1983; Gordon 2007). Challenger regimes in particular adopted European military tactics, developed salaried bureaucracies, and engaged in territorial expansion. While former Mughal provinces such as Awadh and Bengal were quick to cede to European suzerainty, challenger regimes did not. The Mysore Empire for example fought four wars against the East India Company between 1767 and 1799, the Marathas fought three wars

---

14The implication was that from these isolated communities, no durable progress could be made. As Marx had continued, these “idyllic village communities, inoffensive though they may appear, had always been the solid foundation of Oriental despotism,” for “they restrained the human mind within the smallest possible compass, making it the unresisting tool of superstition, enslaving it beneath traditional rules, depriving it of all grandeur and historical energies” (Marx 1972: 40).
from 1775 to 1818, and Travancore fought, and defeated, the Dutch East India Company in 1741.

How did these precolonial regimes durably reshape local institutions such that their legacy persisted through the colonial and post-independence era? First such regimes established a stock of “infrastructural power” at the local level in order to mobilize revenue and manpower. As the fiscal demands upon Indian precolonial states expanded - Bayly (1983) estimates that between 1600 and 1800 Mysore “upped its nominal tax revenue from under 10 per cent of the gross produce to about 40 per cent under Tipu Sultan in 1790s” – states encountered greater local resistance, eliminating by force local chieftaincies and appointing their own tax collectors to collect revenues, as well as police officers to maintain order. This necessitated the expansion of bureaucratic capacity, in the form of a vast army of civil servants and a documentary system to keep track of land cadastres and rights. As Gordon notes of the precolonial Maratha state:
Based on the kamavisdar’s on-the-ground surveys, the body of information flowing from the
district to the Peshwa’s clerks at Pune dramatically increased in volume and specialization.
The once-a-year accounting gave way to shorter and shorter periods. Six-months accounts
were replaced by one-month accounts for the larger and more stable places (like Sironj and
Bhilsa), reaching a peak intensity of daily accounts and treasury receipts from several hundred
villages.” (Gordon 1993: 141)

If at the height of the British Empire, the pattern of land taxation across the subcontinent
(shown in Figure 1.6) still mirrored the territorial reach of India’s extant eighteenth-century
powers, the Maratha Confederacy and the Mysore Empire, it is because the East India Com-
pany inherited the state apparatus that they built. As Robb (2002: 114-5) remarks, under
the East India Company the precolonial “pattern of administration, and many other fea-
tures of the state, were continued.” Outside of Bengal, “Indian institutions, practices and
personnel were to form the basis upon which administration was founded” and “distinc-
tive provincial patterns in their administrative, revenue and judicial practices resulted in a
variability which reduced the degree of centralized administrative authority that could be
achieved over all of British India” (Stein 1996: 207, 242). It could hardly have been oth-
ewise: a trading company with so few administrative staff could not possibly develop its
own fiscal and administrative institutions rather than manage the structures in place. Even
at the height of the British Empire in 1872, a mere one thousand members of the Indian
Civil Service nominally ruled a population of 240m. Thus, as Robb (2002) continues:

“Many of the same titles and offices persisted, as did many of the alliances with local mer-
cantile power, much of the emphasis upon armed force, and much of the reliance on personal
ties and interests among the ruling elites, both European and Indian. The high officers of the
Company, and its historians, presented the establishment of its rule increasingly as a decisive
break. In its everyday affairs, however, it was necessarily and often deliberately gradual and
partial in its innovations. Both Mughal and British empires intruded on and weakened local
and regional entities, but did not replace them.” (Robb 2002: 115).

Second, precolonial states established the legitimacy of public authority, that is, an “ac-
ceptance by people of the need to bring their behavior into line with the dictates of an
external authority” (Tyler 1990: 25). In part this occurred through coercion, or the threat
of it, as states not only taxed and enlisted, but also surveilled and enforced; Rao (1944:
291) reports how the Mysore Kingdom developed a system of “domestic intelligence” that
“penetrated into the inmost recesses of every private dwelling throughout the kingdom,” and Fullarton similarly remarks on how the state operated a network of “spies and intelligencers in every corner.” Yet in part it was generated through genuine commitment to shared political order, over and above the loyalties of the village or the tribe, by making the state the natural locus of political authority.

Such “traditional” legitimation (Weber 1968) is very different from the legitimacy studied in much of the literature in political theory, law, and political psychology, which focuses for example on procedural legitimacy, or the process by which decisions are reached. The procedural justice scholarship for example argues that citizens will comply with the outcomes of decision-making processes that are seen to be fair, even if they disagree with or face material losses from the specific policy recommendations that it generates (Tyler 1990; Sunshine and Tyler 2003; Esaiasson and Ottervik 2014). Deliberative democracy theorists argue that “results are made legitimate by being the results of the procedure” and that “democratic discussion, deliberation, and decision-making under certain conditions are what make the outcomes legitimate for each person” (Christiano 1996; 35). Yet people comply with the law not simply because they are stakeholders in a formal democratic decision-making process, but because historical institutions have established a compliance consensus through monitoring and sanction, and because affective solidarities have been generated by shared experience and collective memory. Such “second-order” judgments regarding “beliefs regarding the beliefs of others” or pre-existing understandings of how others are likely to behave, are just as important as first-order preferences in determining group behavior (Aumann and Dreze 2005, Herrmann and Theni 2009).

Effects of Precolonial States

A consequence is that evidence of past historical legacies is widespread. In addition to the case of Travancore, cited earlier, important precolonial boundary effects are found within the states of Karnataka and Maharashtra. The federal state of Karnataka, for example, is not only known as the home of Bangalore and the Indian IT industry, but also as one of the better governed states in the country, with “a civil service that is more compact,
efficient and responsive than in most other states” (Rao, 2005). In a rating of implementation of sanitation provision across the 423 largest Indian towns and cities, the Karnataka sample (24 cities) came highest outside of the Delhi national union territory and the states of the northeast, and a survey of public goods delivery across “health, education, bus transport, and subsidised distribution” at the turn of the millennium placed “Karnataka in the top decile of India’s states” (Paul et al. 2004). It is plausible that Karnataka, known as ‘New Mysore’ until 1973, has drawn on the loyalties, organization, and bureaucratic infrastructure generated by the Mysore Empire, the key challenger to British hegemony in eighteenth-century southern India. Yet what makes this argument perhaps more compelling is that Karnataka’s impressive record of state-level governance also masks a deep

---

15 In a country with weak fiscal compliance, the state of Karnataka raises 9.8 per cent of gross state product in taxation, the highest level in India. Annual interest payments on debt are just 1.8 per cent of state GDP, the lowest among Indian states (Raju 2012). Karnataka also leads in bureaucratic innovation and in particular, given its status as the home of the Indian IT sector, in e-governance. In a bid to reach the bulk of Indians whose only electronic device is a cellular phone, in 2013 the Karnataka state government piloted in three cities (Bangalore, Hubli and Mysore) a new “m-governance” program: combining services such as utility bill payment, traffic and information services into a single custom mobile phone application (CIOL, 2013).

16 Mysore City, former capital of the erstwhile Mysore Kingdom, was second in the country overall (National Urban Sanitation Policy, 2010).
internal divide, between the state’s northern and its southern districts, that perfectly maps the northern limit of the eighteenth-century Mysore Empire, even though this border has not existed since the Anglo-Mysore wars of 1781-1799, over two centuries ago (Figure 1.9). In a survey conducted in 2005 by the University of Maryland, while respondents in southern Karnataka expressed above-average confidence in local government to ‘deliver public projects’, respondents in northern Karnataka were among the lowest decile of all India (Desai et al. 2006). Excluding the city of Bangalore, crime data from the 1980s shows twice as many reported kidnappings in northern Karnataka than in the south\(^{17}\), and the same University of Maryland survey revealed 24 per cent of households in northern Karnataka report being victim of a theft, while the average among southern Karnataka districts was just 2 per cent. Both northern and southern Karnataka are inhabited by the same Kannadiga peoples, and most districts along the border separating the two were subject to direct rule under the British Empire, as part of the Bombay Presidency\(^{18}\).

**Alternative Theories**

In the previous section I have established the fact of diverging patterns of state formation in precolonial India, and outlined some of mechanisms linking eighteenth-century states to contemporary institutional strength. In this section I will examine the argument for historical state legacies against common alternative arguments for why patterns of state capacity are so divergent across the map today. In particular, I assess alternative theories that focus upon colonial-era policies, economic conditions, intergroup heterogeneity, subnational identity, caste structures, and religious tensions. In doing so, I emphasize how many of these factors are endogenous to the precolonial institutional inheritance, and that differences in precolonial trajectories are a deep determinant of subsequent endowments of state capacity.

---

\(^{17}\)The rate was 25.8 per 100,000 in the north, against 14.2 per 100,000 in the south. Including Bangalore City, the southern rate reaches 18.9.

\(^{18}\)This is a generalization, as Bellary district was assigned to Madras, again under direct British rule, while Raichur was part of Hyderabad. The southern portion of Karnataka was part of the rump princely state of Mysore, the remaining territories left by the defeat of the Mysore Empire in 1799.
Colonial Era Policies

British rule left a number of distinctive legacies on regions of India, notably those areas, such as Bengal or Bihar, which were under direct colonial rule for almost two centuries. Under the policy begun by Cornwallis in 1793, in these and other areas of India tax collectors (“zamindars”) were transformed into hereditary landlords, whereas in other regions tax collectors remained in their traditional role and assessed the tax annually as a percentage of output (Banerjee and Iyer 2005). In addition, after 1857, India was permanently divided between areas that would remain under indirect rule, and areas that would be governed directly as British presidencies. Discussions of the British legacy dominate historical and institutional studies in South Asia, which have generally focused on these aspects of the colonial era, rather than precolonial institutional development. Mukherjee (2013) and Iyer (2010) for example attribute legacies of political order, public goods delivery and economic development to patterns of direct versus indirect rule under the British Raj, while Banerjee and Iyer (2005) attribute differences in long-run economic development across regions of India to decisions by colonial administrators regarding land tenure patterns.

Yet neglect of precolonial legacies is problematic, as the decisions and policies of colonial administrators were heavily endogenous to pre-existing institutions. As Robb (2002: 114-5) remarks, under the East India Company the precolonial “pattern of administration, and many other features of the state, were continued.” With respect to direct and indirect rule, as Gerring et. al. (2010) observe, areas with strong precolonial institutions were somewhat more likely to be governed indirectly, due to both the cost of subjugation and the fact that pre-existing polities had already established effective mechanisms of taxation and political order. Polities that continued as indirectly ruled “princely states” were those which had managed to resist British incursion until the official end of Company Raj in 1857. Likewise, with respect to land-tenure arrangements, choices regarding the system of tax collection were similarly dependent upon preexisting level of fiscal centralization. The East India Company in 1793 experimented with a landowner-based system of tax collection due to the inefficiency and imprecision of the zamindari method; in areas of India that were governed by more centralized states, however, the assessment method achieved much higher yields and did not need reform. As Kulke and Rothermund remark,
During these years of economic change [the first half of the nineteenth century] the British had penetrated deeply into the interior of India and had organized the revenue settlement of many provinces. The ‘permanent settlement’ [of zamindars as landlords] remained restricted to Bengal and Bihar. In the South, Tipu Sultan and the Marathas had established a rather rigorous direct assessment of the peasants; in the North the assessment of zamindars or village communities prevailed. The British continued nearly everywhere the type of assessment set up by their immediate predecessors. (Kulke and Rothermund 1986: 233, emphasis added).

British direct rule may have had negative effects upon institutional development by interrupting preexisting trajectories of state formation, yet the distribution of the effects of colonial policy cannot be considered as exogenous from earlier patterns of institutional development. Indeed given the exceptionally thin administrative apparatus with which Britain ruled India, and the consequent need to govern with the cooperation of local elites, colonial administrative structures inevitably borrowed heavily from the administrative structures they inherited19.

**Economic Origins**

A conventional explanation in political science would be to attribute patterns of institutional diversity and governance to prevailing economic conditions. In areas with fewer economic resources available, local governments may have less resources to invest in higher quality personnel and training, judges and civil servants would face stronger incentives to augment their salaries with bribes, and voter behavior may be more likely to remain clientelistic rather than programmatic, rewarding politicians that form rent-seeking coalitions rather than promise impartial, accountable norms of government. An important “modernization” tradition in political science has advanced the view that after a generational lag, rising incomes lead to greater norms of transparency, accountability, and rising demands for good governance through civic mobilization and activism (Deutsch 1961; Dalton 1989; Inglehart and Welzel 2005).

A problem with such an explanation is that, while low capacity districts of northern

19As Bayly remarks, “where the state was strong as in the Maratha territories around Poona between 1751 and 1818, or in Mysore where Haider Ali and Tipu Sultan tried to eliminate all intermediaries between themselves and the peasant farmers, subjecting them to new demands for the punctual payment of revenue. Elsewhere the absence of powerful authorities outside the village enhanced the authority of the elites as protectors and petty rulers of the villages” (Bayly 1988: 28).
India may be poorer today, historically this has not consistently been the case - including during the early twentieth-century when the current north-south state capacity gap was already in evidence. As Figure 1.10 shows, income divergences between the Hindi heartland and the southern districts are a relatively recent phenomenon, that became pronounced only since the 1980s. Yet data from colonial records shows that deep north-south differences in state capacity, ranging from taxes, to policing, to public goods provision, already in 1900. Regional estimates of income per capita in the *Cambridge Economic History of India* for the late nineteenth century place a number of southern regions, such as Madras, behind northern regions such as Bengal (Desai 1993). Enduring differences in economic starting points cannot explain differences in institutional outcomes: if anything, it seems more likely that differences in quality of institutions would explain India’s diverging economic trajectories since independence.

**Figure 1.10: Income Divergences between Northern and Southern States, 1960-2006.**

*Intergroup Heterogeneity*

An alternative theory that is dominant in contemporary political science is that the quality of institutions reflects the prevailing degree of social diversity. Where many different groups - whether they are ethnic, religious, or linguistic - coexist under a single political
arrangement, they are less likely to form a consensus over desirable public goals, have larger transaction costs when engaging in collective action, and require larger selective incentives to form governing coalitions (Easterly and Levine 1997; Alesina et al. 2003). Intergroup diversity undermines state capacity directly by increasing the incentive to engage in clientelistic behavior rather than follow impartial bureaucratic norms, and reducing the commitment of public servants to the state over and above sub-state identity and loyalties. This theory is especially widely cited with respect to India, whose social fabric is extremely diverse, with its many divisions over language, caste, religion, and tribal status. Social heterogeneity is frequently cited as a reason for India’s fragmented political landscape, clientelism, and patrimonialism (Keefer and Khemani 2004).

However, whether or not social heterogeneity explains weak state capacity for India in cross-country perspective, it is problematic as an explanation for variation in state capacity subnationally. First there is the simple problem of causal identification, for one of the key legacies of historical state formation is that – as Gellner (1997) argued – states create social homogeneity. Eugen Weber’s classic (1978) study showed how the policies of the French Republic turned “peasants into Frenchmen,” by creating a French-speaking population out of disparate provincial identities, and regional states in precolonial India similarly served to create or to preserve Malayali, Marathi, or Kannada speaking enclaves. Moreover, just as centralizing states can create encompassing linguistic or ethnic identities, similarly a collapsing imperial state can leave in its wake disparate groups that have not yet been weld into a shared society. The collapse of the Mughal Empire in the early eighteenth century left behind large enclaves of urban, Urdu-speaking Muslims across northern India, while the nested ownership structures and unequal landlordism of precolonial north India enforced the economic disparities of caste identity (Alam 1986). In this way, again, state formation and state collapse had the tendency of reinforcing or weakening the salience of group identities.

For these reasons measures of intergroup heterogeneity are only imperfectly and noisily associated with indicators of state capacity. The standard social science measure of intergroup diversity, an indicator of fractionalization based on the Herfindahl index, provides a summary statistic of the likelihood that any randomly chosen other person in that society will belong to a different group. Figures 1.11 show a district-level fractionalization index
based on the likelihood that another random individual in that district will belong to a different caste level (*varna*), religious denomination, or membership of a scheduled tribe, and its correlation to local-level state capacity.

![Figure 1.11: Ethno-Caste Fractionalization and State Capacity Indicators.](image)

A somewhat positive relationship does exist between intergroup diversity and public order, reflected in the homicide rate. Yet as will be explored in greater detail in Chapter 5, the prevalence of crime across gangetic northern India has its origins in the decline of the Mughal imperium in the eighteenth century, in particular in the breakdown of organized policing and rise of “dacoity” or highway robbery, across the former Mughal heartland, and because such districts also have larger Muslim minorities may explain the covariance between homicide and intergroup fractionalization.

*Variation in the caste system*

Caste is a Hindu social institution according to which every individual is born into an occupation-based marriage group (*jati*) that is ranked by a grade (*varna*), from “upper” caste roles such as religious and political authorities to middle roles, such as tradesmen, merchants, and artisans, to “lower” caste or “non-caste” roles such as mendicants and menial workers (Jaffrelot 2000). While the caste system exists across India, it is sometimes
noted that caste hierarchy is more pronounced in the North than in the South, and that this in turn might explain differential patterns of clientelism and the absence of rule of law. Weiner (1991) argues that caste is a cause of weak state capacity as it justifies the exclusion of lower caste communities from state services, while Jauregui (2013) argues that caste hierarchies undermine the state’s ability to enforce laws, by providing an alternative source of legitimacy that justifies upper castes to habitually violate the lives and property of lower caste groups, and to even defy attempts by officers of state to enforce justice. In Northern states lower-caste parties have achieved prominence, notably in the state of Uttar Pradesh, and this has also been accompanied by accusations of clientelism and financial mismanagement.

There are two problems with theories that posit caste divisions as a prime explanatory variable. The first is that differences in caste empowerment between the north and the south are comparatively recent; they are largely a consequence of the relative success of southern states in implementing caste quotas in the 1950s and 1960s, while differences in state capacity are more deep rooted (Yadav 2000; Jaffrelot 2003; Manor 2010). The second problem is that the relative salience of caste structures cannot be treated as exogenous from broader social and political institutions, including state capacity and the legacies of precolonial states. After all, formally, the caste system exists in all of India, from the north to the south of the country. Arguments relating to caste therefore rotate around the “relative importance” of caste identities in some regions with respect to others. Stated thus, it seems more clear that the “salience” of caste cannot be treated as an exogenous social fact, but is endogenous to the factors that might make caste identification more or less useful to social actors: such as the existence (or absence) of impartial state structures. Similar arguments have been made regarding the social construction of ethnicity in Africa, where ethnic identities have been generated as mobilizing devices in weak states lacking other means of group identification (Posner 2000; Bates 1991).

**Regional subnationalism**

Another argument for variation in patterns of public goods provision across India is that in those regions where subnational political mobilization occurred during the era of the independence struggle and the period preceding states reorganization, stronger, more inclusive
political institutions emerged (Singh 2011). There is some overlap between the subnational identity and precolonial state arguments. One of the longstanding debates in the nationalism literature is whether states precede nations or vice versa (Gellner 1997; Smith 1983), and India’s medieval and early modern states set the preconditions for subnationally-oriented movements due to their political autonomy, print media, and shared administrative languages, as well as the shared identities generated by collective mobilization through warfare. States that trace their origins to a dominant precolonial polity are both better governed and those which form the most cohesive subnational units today, such as Karnataka (Mysore), Kerala (Travancore), Maharashtra (the Maratha Empire) and the Punjab (the Sikh Empire). A strong regionalist party that mobilizes based upon a regional identity is also found in Maharashtra (the Shiv Sena) and Punjab (the Akali Dal).

A limitation of the subnationalist argument however is that it cannot explain why there are divergences in state capacity within post-1956 federal states, if the relevant unit of variation is the federal subunit, and in particular why precolonial state boundaries should continue to influence variation in state capacity measures. The case of Kanyakumari, a former Travancore district now located in Tamil Nadu, was cited at the start of this chapter. Also notable is the eastern boundary of the Maratha peshwa’s dominions, which divides the better-governed western from the more backward regions of eastern Maharashtra, and the northern limit of the Mysore Empire, which separates the more functional districts of Karnataka’s south from those of its north (Figure 1.9).

The indigenous population
States in central and northeastern India have larger numbers of indigenous peoples, or “adivasis,” also known as “original inhabitants,” the term “adivasi” deriving from the words “adi” (earliest time) and “vasi” (“resident of”). At the time of writing there are 533 officially recognized adivasi communities in India, and their percentage of the population ranges from less than 0.1 per cent in Goa or Uttar Pradesh, to between fifty and ninety per cent in the states of the Northeast. While there are important differences between groups, adivasi areas are generally in rural communities, have low levels of human development, and are politically and economically marginalized, a fact that has become salient due to the success of Maoist insurgents in mobilizing certain adivasi areas.
There are two key problems with a theory that regional differences in institutional quality can be attributed to the distribution of indigenous peoples. The first is that the size of the adivasi population cannot be considered separately from processes of historical state formation. As Scott (2005) has argued with respect to neighboring upland southeast Asia, the degree to which indigenous peoples have become assimilated into a broader state society is a result of the extent to which states have existed that have sought to tax, mobilize, and otherwise coerce the local population into its fold. At one point, all of northern India was “indigenous,” in the sense of living within a pre-Aryan culture before the adoption of Sanskrit languages and the development of the Hindu religion; now there are degrees of “sanskritization,” ranging from tribal groups who retain their identity and religion, but have adopted Sanskrit languages, to those which have essentially become another Hindu group or jat designation. In areas where Hindu kingdoms developed, indigenous peoples underwent substantial Sanskritization; in other areas (notably, the northeast) they did not.

The other problem is that the size of the tribal population bears little to no relationship to indicators of historical or contemporary state capacity, as is shown in figures 1.12.

![Figure 1.12: Tribal Population and State Capacity Indicators](image)

Between districts there is no significant relationship between the proportion of adivasis and either historical fiscal capacity or the degree of contemporary public order. The reason why tribal population proves a poor overall predictor of north-south variation in state ca-
pacity is simply that prevalence of adivasis is not codeterminant with regions of northern India, but more specifically the broader northeast, ranging from Jharkhand through to Assam and the northeastern states proper, while in many states of the Hindi heartland (such as Uttar Pradesh) the adivasi population is very low. Thus the presence of adivasis does explain the institutional underperformance of the “Hindi belt” from Haryana to Uttar and Madhya Pradesh, and nor can the relatively low number of adivasis in many southern regions explain their relatively higher levels of fiscal mobilization or legal compliance.

Religious conflict
Another salient difference between the north and south of India is the salience of religious conflict. In the North, major instances of Hindu-Muslim violence have occurred in Uttar Pradesh (such as the Babri mosque incident) and in Gujarat (the Gujarat violence of 2002), while significant Hindu-Sikh violence broke out in in Delhi and the Punjab in 1984. Across southern states, instances of intergroup violence are far fewer in number and intensity. Even if caste and ethnic diversity do not predict poor institutional performance, perhaps religious diversity does.

However, while relations between religious denominations may be more fractured in the North, with the exception of significant Muslim minorities within certain Gangetic districts, there is little difference between North and South in the actual number or diversity of religious groups, or in the degree of religious adherence. Indeed, India’s most religiously fractionalized state, Kerala, is also one with the lowest religious tensions. Comparative surveys find that in Kerala, respondents are least likely to object to having neighbors of a different religious denomination, and most likely to agree that “people who belong to different religions are probably just as moral as those who belong to mine” (World Values Survey, 2014). 87 per cent of Keralites disagree with the statement that “the only acceptable religion is my religion”: yet in the more religiously homogeneous states of the North, similar majorities take the opposite view, that only their own religion is valid. Therefore it is true that religious tensions are greater in northern India, yet religious diversity does not seem to be the explanatory cause. If anything, as Varshney (2003) and Wilkinson (2004) have argued, social and political institutions may well be the mediating variable that explain why in some regions, religious diversity leads to conflict, whereas in others,
it results in stability. In this case, however, what needs to be explained are the patterns of institutional variation, and not patterns of religious diversity.

Overview

This dissertation combines econometric analysis along precolonial boundaries, bureaucratic responsiveness tests, new archival data on local state capacity during the colonial era, and case studies of precolonial polities using archival records in order to provide a general theory of how early processes of state formation lead to the accumulation of the norms and resources allowing states to exercise fiscal, bureaucratic, and policy delivery capacity. In this chapter I begin by establishing the puzzle of institutional persistence across India since the earliest local-level data of the late nineteenth century, and the entrenched north-south divide that has its origins in at least the early colonial period. I have also highlighted examples of discontinuities in state capacity within contemporary Indian federal states along precolonial boundaries, pointing to an enduring legacy of precolonial patterns of state formation for the strength of government in India today.

In the next chapter, I outline a theoretical basis for explaining this persistence, proposing a theory of compliance, or norms of obedience to public authority, that is generated by the incentives imposed by historical states. The following chapter then conducts an audit of India’s precolonial regimes, showing how they restructured state-society relations at a local level. In the late seventeenth and early eighteenth century a range of “challenger states” arose in the Indian subcontinent in opposition to both Mughal and European attempts at suzeraignty, achieving a higher degree of state centralization that that which existed either among the traditional micro-polities of the rajputs or poligars, or the declining provinces of the Mughal imperium. I further show how the territorial extent of India’s eighteenth-century regimes can be measured and mapped on to contemporary district boundaries.

Next, in chapter 4 I proceed to formal econometric tests, by conducting full sample regressions and border sample regressions to identify discontinuities along precolonial borders. In order to test for the effects of early political development, I use a border specification to identify the enduring effect of these earlier regimes by comparing districts either side of their long-extant precolonial boundaries. Because the precolonial borders of the major
precolonial polities cut through several of India’s linguistic states, including Maharashtra, Karnataka, and Tamil Nadu, this allows for identification of discontinuities along either side of long-extant boundaries, holding constant state-level policies. I find that peoples who today narrowly live in territories that were part of more centralized, warfaring states such as Mysore, Travancore, and the Sikh Empire have higher confidence in local government, higher confidence in police, and lower rates of reported kidnappings and robberies than those narrowly on the other side. I validate this interpretation by null-testing “placebo” borders outwards of the actual precolonial boundaries, and also show that district-level effects are robust to controls for state-level variation, suggesting that precolonial treatment has persisted over the effect of post-independence state-level policies. While the estimated effects of the full-sample regressions are robust to a wide range of controls, limiting the sample to precolonial borders also removes important sources of unobserved heterogeneity.

Chapter 5 turns to the question of how such differences in state capacity had been able to accumulate by the colonial era. I present results of archival research to analyze the origins of diverging trajectories of state formation in precolonial India, and in particular the divergent paths taken by the former Mughal provinces of the north and the challenger regimes of the Mughal periphery. In particular, I assess the policies and institutions of the precolonial Mysore, Oudh, and Maratha polities, in order to show how defensive modernization led to an increase in fiscal and military mobilization in the case of Mysore and the Maratha dominion, and how increasing autocratic centralism led to the formation of compliance equilibrium between the Mysore sovereign and the agents of the state, as well as among society as a whole. These differences are analyzed further in Chapter 6, which addresses the question of persistence by tracing how the state structures and norms built in the precolonial era survived under the aegis of the colonial state.

Having established precolonial state formation as an important source of variation in contemporary state capacity in India today, in Part B I move to a generalized test of the hypothesis. I begin by conducting a cross-country analysis of early states and contemporary capacity in chapter 7, which examines the relationship between historical states and state capacity globally. Consistent with existing cross-country studies by Gerner (2012), Putterman (2011), and Gennaoli and Rainer (2007), I show that countries which inherit a
legacy of deep historical state formation exhibit greater state capacity in the present, and that such effects outweigh the importance of other explanations for contemporary variation in government effectiveness. I then continue in Chapter 8 by conducting an analysis of state capacity and service delivery in the context of post-soviet Russia, a large and diverse transitional regime that is historically, culturally, and socially distinct from India, and show that provinces with a legacy of historical state formation in the early modern era were more successful in handling the new revenue mobilization and service provision responsibilities that arose from federalization.

Finally, in chapter 9 I bring together the findings of this research. I highlight the ways in which the enduring legacy of early state formation builds on earlier work in studying the long-term effects of historical states and regimes, as well as how it suggests new pathways of research. By arguing for the importance of indigenous institutional development, I draw conclusions for policymakers in terms of working within historical patterns of governance and the importance of ensuring local policy ownership, as well as considering the ways in which political intervention may at once overestimate short-term benefits, while underappreciating long-run negative effects.
Chapter 2

The Culture of Compliance

*In the absence of trust it would become very costly to arrange for alternative sanctions and guarantees.*
- Kenneth Arrow (1969)

*Trust is good, but control is better.*
- V.I. Lenin

If we are to understand differences in patterns of state capacity between countries and within a large country such as India, I argue that we need to start by looking at different notions and patterns of *compliance*. Compliance refers to obedience to public authority, as exhibited by paying taxes, obeying laws, or participating in mandatory public service duties. States demand such compliance across a range of activities, as they secure the manpower, resources, and personnel that are essential to their functioning. Fiscal compliance, for example, refers to the degree to which individuals and organizations pay the taxes,

---

1This phrase is generally attributed to Lenin without appearing in his published works. It is likely either a derivation of his use of the phrase “trust, but verify,” or an abridgement of one or several similar quotations from his writings: “believe not in words but in verification,” or “we do not indulge in ‘dreams’ of dispensing at once with all administration, with all subordination... we want the socialist revolution with human nature as it is now, with human nature that cannot dispense with subordination, control and ‘managers.’ ” (Lenin 1966: 307-8).
tolls, or fees that are required of them, and pay these honestly and promptly. Regulatory compliance refers to the level of adherence that citizens pay to the state’s laws\(^2\), and bureaucratic compliance refers to the extent to which public officials, such as teachers, police officers, public doctors, and civil servants, follow the command of their superiors and the policies they enforce. Where these forms of compliance exist then states, by definition, exert capacity: decisions taken by the government are certain to be implemented. But where compliance does not exist, whether at the central, regional, or local level, states will lack capacity in important ways, because it means that revenues are misappropriated, orders are ignored, or laws are broken. In the words of Stanley Milgram - whose work exposed the more negative dimensions of social compliance - “compliance to authority ... flows from the logical necessities of social organization.” As he continued, “if we are to have social life in any organized form – that is to say, if we are to have society – then we must have members of society amenable to organizational imperatives” (quoted in Blass, 2002).

Yet levels of compliance with the state vary greatly, both between individuals, and among regions, and entire countries. In some countries, tax evasion and regulatory malfeasance are so widespread that the shadow economy accounts for two-thirds of national income, while in others, less than ten per cent of economic activity goes undeclared (Schneider 2006)\(^3\). In some states of India, absenteeism among public employees is so rife that up to two-fifths of staff do not attend their official jobs, while in others, this rate is less than one per cent (Government of India, 2012)\(^4\). Experimental studies have also shown wide differences in compliance when subjects are faced with comparable legal obligations in different countries, such as returning a missing wallet, refraining from jaywalking, or when diplomats from different cultural backgrounds are faced with the identical obligation

---

\(^2\)Strictly speaking, insofar as the state has the monopoly of legitimate public authority, both fiscal bureaucratic compliance are subcategories of legal compliance. Yet we can use the term legal compliance to refer to the specific demands that the state makes of its citizens, beyond the payment of taxes or any role they have as public servants.

\(^3\)In Schneider’s estimates, only the United States and Switzerland have a shadow economy of less than a tenth of national income. Estimates for the size of the shadow economy however exceed sixty per cent in the cases of Zimbabwe, Tanzania, Panama, Peru, Georgia and Azerbaijan.

\(^4\)The Ministry of Labour and Employment’s (2012) *Report on Absenteeism, Labour Turnover, Employment and Labour Cost* shows absenteeism among public employees to have been 38 per cent in Rajasthan in 2011, compared to less than 1 per cent in Sikkim or Tripura.
to pay parking fines (Wiseman 2009, Moysa et al. 2013, Fisman and Miguel 2007)\(^5\).

Such wide patterns of variation call for an explanation. Why do individuals in some circumstances choose to comply with public authority, and in other circumstances not? Why do authorities invest resources in securing societal compliance in the first place? And why do such norms of compliance, once established, tend to persist - long after the initial conditions from which they emerged? In order to answer these questions, this chapter discusses the mechanisms underlying compliance decision-making. I propose a theory that legacies of state coercion are the most important factor in explaining why compliance norms exist, by generating norms in response to the institutional incentive environment, and that this is consistent with both empirical data and behavioral theory. I open with a brief examination of empirical data, showing that the best predictor of worldwide variation on such norms is a country’s history of state formation. Then, in the first section, I consider the microfoundations of compliance, in order to establish a general basis for why and when individuals opt to comply with authority. Addressing limitations with three of the leading theories of compliance norms - those of voluntary compliance, of collective action, and of simple rational choice - I propose a modified rational choice approach, augmented using insights from behavioral economics, arguing that only sustained patterns of coercion will lead to the formation of compliance norms. In the second part of this chapter, I examine the theoretical basis for persistence: why it is that compliance norms, or cultures of compliance, tend to remain durable among communities over time despite changes in policies, political regimes, and other alterations in the overt incentive structure. I argue that collective action dynamics are applicable in explaining how states maintain a compliance equilibrium, and secondarily, that long-term cultures of compliance generate shared mental models, including collective beliefs, identities, and values, as a buffer to such collective action equilibria. Next, I address the issue of why it is that historical political authorities

\(^5\)Though these differences are sometimes interpreted as reflecting norms of ‘social capital,’ they are fundamentally measures of compliance because they primarily measure accordance with a legal obligation. For example, in many countries, the return of a lost wallet is a legal obligation, albeit one that can almost never in practice be enforced. Meanwhile Fisman and Miguel’s study of international differences in adherence to New York parking laws, based on data on parking violations by United Nations staff, also concerns a legal requirement, albeit one without likelihood of prosecution: and diplomats from Kuwait, Egypt, Chad, Sudan and Bulgaria were found to have the highest numbers of incidents, while at the other extreme, there were no recorded violations by delegates from the United Kingdom or Japan.
have sought to generate compliance from their subjects through means of bureaucratic control and enforcement in the first place. I consider three arguments that are common in the literature on state formation, namely the extractive hypothesis, the threat of revolution, and interstate warfare. I suggest that the two latter hypotheses - which can simply be referred to as the ‘elite threat’ hypothesis - provide the most theoretically satisfying explanation for why authorities have sought to establish centralized political technologies for behavioral regulation.

**Milgram’s Question - Variation in Cultures of Compliance**

Do norms of compliance vary systematically between cultures? In the 1950s, Yale psychologist Stanley Milgram was motivated by the observation that, during the Second World War, the Nazi regime had occupied large portions of Europe, including the Netherlands, northern France, and northern Italy, requiring the complicity of occupied peoples in rounding up hundreds of thousands of their Jewish neighbors for deportation to the death camp. As he notes in his autobiography, Milgram wanted to know why it was that in cases such as the Netherlands, a country with a history of religious pluralism and tolerance, the majority of the Jewish population, counting one-hundred and fifty thousand individuals, was sent to such a fate; however in a country such as Italy, which had an official state religion and where ghettos were abolished barely a century prior, only 7,500 persons were murdered, allowing over four-fifths of the country’s Jewish population to evade persecution through the aid of their neighbors and friends. Milgram suspected that the answer lay in differential norms of compliance, and based his dissertation upon a cross-country replication of the Asch conformity test, finding that his northern European (Scandinavian) sample exhibited greater peer conformity than those in his southern (French) test group. However it was Milgram’s eponymous experiment, conducted at Yale University two years later, that has become the most famous test of compliance norms in modern social science (Milgram 1974). Milgram’s original experiment saw 40 male par-

---

680-90 per cent of the Jewish population in Italy survived, as against just 30 per cent in the Netherlands. Indeed, the Netherlands had one of the worst survival rates for its Jewish population of any country under occupation, whereas in occupied Italy, France, and Belgium, a majority of the Jewish population survived.

7Milgram’s original research design was to compare German with French subjects; on the advice of his committee, Germans were substituted with Norwegians in the conduct of the experiments.
participants assigned to the role of a ‘teacher’ and ordered to administer progressively larger electric shocks to a “learner,” played by an actor, under the authority of an experimenter. 63 per cent of Milgram’s American subjects complied with the experimenter to deliver progressively larger shocks until they had delivered the maximum voltage, despite being made aware that this was potentially mortally hazardous to human health, and viewing the visible (albeit, feigned) suffering of the other participant. Yet in parallel experiments conducted in Italy, Germany, Australia, the United Kingdom, Jordan, Spain, Austria and the Netherlands, often higher rates of compliance were found. The highest rate of compliance was found in the Netherlands, where 90 per cent of subjects completed the demands of the experimenters, followed by West Germany, where 85 per cent did so. Slightly lower rates were found in Italy and Austria, where 80 per cent completed the experiment, and much lower rates of compliance in Anglo-Saxon countries, including not only the 63 per cent completion rate of Milgram’s original experiment in the United States, but just 50 per cent in the United Kingdom, and only 40 per cent of men and 16 per cent of women in Australia, the lowest rate of compliance in the experiment before replication elsewhere was discontinued (Smith, Bond and Kagitcibasi 2006; Meeus and Raaijmakers 1986; Mantell 1971; Kilham and Mann 1974).

Why do differences in compliance norms exist between countries? The cross-country replication of the Milgram Experiment offers only a limited number of cases. Yet one variable that relates to differences in experimental compliance are perceptions of individual versus state responsibility, as shown in Figure 2.1, which shows the correlation between Milgram Experiment compliance and a survey item for whether “the state should take greater responsibility to see that people are provided for” or whether “individuals should take responsibility for themselves,” taken from the World Value Surveys (significant in a chi-square test; World Values Survey 1981-2014).8

8In addition, Edwards et al. (1963) report a 87.5 per cent rate of compliance in South Africa, albeit based on a limited test sample of 10 students.

9Of course, the willingness of individuals in Anglo-Saxon societies to defy authority in the application of lethal force is only indicative of compliance in other areas. Moreover, in the case of the Milgram experiments, individuals may have defied the experimenters precisely as an act of compliance - only with a broader framework of laws establishing the participant in the experiment as a legal party to the outcome of their actions. Compliance, similar to trust, is a conditional rather than a general relationship: it is always the compliance of
Figure 2.1: Compliance with Milgram Experiment, and Surveyed Belief in State vs. Individual Responsibility \((r = +0.52)\). Milgram experiment rates of compliance as reported in Milgram et al. (1971), Smith, Bond and Kagitcibasi (2006), Meeus and Raaijmakers (1986), Mantell (1971), Kilham and Mann (1974). Survey data from the World Values Survey: average agreement with the view that “The government should take more responsibility to ensure that everyone is provided for” versus “people should take more responsibility to provide for themselves.”

While experimental data of the kind collected by Milgram and his colleagues is not available for a large number of countries, we can examine differences in the “culture of compliance” more systematically for a larger number of countries using a simple compliance construct from comparative survey data. Using the most recent available data from the World Values Survey, which covers 100 countries on all six inhabited continents, and contains more than 90 per cent of the world’s population, Table 2.1 lists four basic constructs that tap into the a priori attributes of cultural compliance, ranging from a) views of the justifiability of such violations of state imperatives as tax evasion, benefit fraud, and bribery (*internalization of law*), b) the Schwartz value item for a personal sense of propriety in avoiding actions “people would judge as wrong” (*social esteem*), c) emphasis on individual \(x\) with authority \(y\), concerning some matter \(z\).
socialization into disciplinary values such as self-denial and obedience (discipline), and d) perceptions that others in society exercise restraint against non-compliant behaviors (self-restraint). While these are only measures of compliance attitudes rather than behaviors, they at least give an indication of the extent to which such norms exist.

Is a legacy of long-term, historical state formation associated with a culture of compliance? Results from a sample of 64 countries are shown in Figure 2.2. The societies with the highest levels of expressed compliance norms are those of northeast Asia (Japan, China, Korea and Vietnam), the Nordic countries (Denmark, Sweden, Norway, Iceland, Finland and also Estonia), and Switzerland. Those with the lowest degree of cultural compliance are primarily found in Sub-Saharan Africa and Latin America. Also shown on the x-axis is a measure of historical state formation, the “state history” scores of the State Antiquity database assembled by Bockstette, Chanda, and Putterman (2002). The index is constructed by taking each period from 1 to 1950 AD, and allocating points to countries if there was i) a government above the tribal level; ii) if that government was locally based rather than that of a foreign empire; and iii) a fractional point to represent the extent of the country’s modern territory that was under the control of this earlier government. The data from these fifty periods is combined, thereby offering an index of state history for a large sample of countries, and a discount rate applied to weight more recent history over the distant past. This index is shown to correlate broadly with the index of cultural compliance, indicating that countries such as those of northeast Asia and northern Europe that have a history of centralized, bureaucratic state control exhibit stronger compliance norms, while countries in Sub-Saharan Africa or Oceania, with absent state legacies, exhibit weaker commitment to social compliance.

We can investigate this relationship between historical state formation and norms of compliance more thoroughly in the framework of basic OLS. Table 2.2 shows the results of OLS regression analyses of cross-national differences in the compliance index, as measured in 81 societies, using the latest available survey for each country and the Chanda and Putterman (2002) State Antiquity Index (the reduced n accounting for missing data on some independent variables). Also included are a series of control variables for socioeconomic and political factors which might account for differences in such norms such as income per capita, the size of the industrial sector, and the level of democracy, as well as
<table>
<thead>
<tr>
<th>Concept</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) Internalization of law</td>
<td>Please tell me for each of the following statements whether you think it can always be justified, never be justified or something in between: i) Claiming state benefits to which you not entitled; ii) Avoiding a fare on public transport; iii) Cheating on tax if you have the chance; iv) Someone accepting a bribe in the course of their duties.</td>
</tr>
<tr>
<td>ii) Social Esteem</td>
<td>Now I will briefly describe some people. Using this card, would you please indicate for each description whether that person is very much like you, like you, somewhat like you, not like you, or not at all like you? Relative mentions: It is important to this person to always behave properly; to avoid doing anything people would say is wrong.</td>
</tr>
<tr>
<td>iii) Discipline</td>
<td>Here is a list of qualities which children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Relative mentions of: discipline responsibility selflessness</td>
</tr>
<tr>
<td>iv) Restraint</td>
<td>Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair? Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?</td>
</tr>
</tbody>
</table>

Table 2.1: Measurable Aspects of Compliance.
Figure 2.2: State History and Cultural Compliance Index. Index of cultural compliance measures: views on justifiability of tax evasion, benefit fraud, bribery and fare dodging; importance of discipline, responsibility and selflessness; reported sense of shame in doing acts “people would say are wrong”; judgments of social trustworthiness. Settler colonies excluded.
dummy variables for a series of cultural zones used in Inglehart and Baker (2000), namely, a western Protestant, Catholic, Communist or Islamic heritage. Model (1) shows that using the full country sample, state antiquity is the strongest overall predictor of compliance norms, whereas socioeconomic factors are not significant. Civilizational zone variables are not significant, with the exception of the Confucian and Protestant dummy variables, which have a weakly robust association. Removing settler colonies from the sample (Models 2 and 3) produces similar results. In settler colonies, the relationship between state history and social norms does not apply, presumably because settlers bring norms of their own cultures, which may in turn be the product of a long state history. When state antiquity is removed from the model (Model 5), an important difference emerges: the dummy variable for the Confucian cultural zone becomes strongly significant, while the coefficient for the Protestant variable remains positive but non significant at the .05 level. The Confucian dummy variable is thus only weakly robust to the inclusion of a variable for state antiquity.

The non-robustness of “cultural” variables shown in Models (1) and (5) has important implications. Differences in compliance norms between countries may not so much reflect...
<table>
<thead>
<tr>
<th>Model</th>
<th>(1) Full Sample</th>
<th>(2) Europe, Asia, and Africa only</th>
<th>(3) Europe and Asia only</th>
<th>(4) Settler Colonies Only</th>
<th>(5) Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>State History Index</td>
<td>0.979** (0.301)</td>
<td>0.985** (0.359)</td>
<td>1.153* (0.506)</td>
<td>0.133 (0.501)</td>
<td>-</td>
</tr>
<tr>
<td>Freedom House, 2011</td>
<td>0.008 (0.008)</td>
<td>0.004 (0.007)</td>
<td>0.005 (0.011)</td>
<td>0.008 (0.014)</td>
<td>-0.002 (0.006)</td>
</tr>
<tr>
<td>GDP per capita (.000), 2000</td>
<td>0.001 (0.008)</td>
<td>-0.005 (0.008)</td>
<td>-0.006 (0.010)</td>
<td>0.003 (0.016)</td>
<td>0.013 (0.007)</td>
</tr>
<tr>
<td>Industry as a percentage of GDP</td>
<td>-0.003 (0.01)</td>
<td>0.003 (0.008)</td>
<td>0.002 (0.012)</td>
<td>-0.042 (0.037)</td>
<td>0.003 (0.007)</td>
</tr>
<tr>
<td>South Asia (= 1)</td>
<td>0.363 (0.274)</td>
<td>0.38 (0.268)</td>
<td>0.385 (0.319)</td>
<td>-</td>
<td>0.482 (0.255)</td>
</tr>
<tr>
<td>Historically Islamic (= 1)</td>
<td>-0.223 (0.185)</td>
<td>-0.21 (0.167)</td>
<td>-0.187 (0.233)</td>
<td>-</td>
<td>-0.027 (0.148)</td>
</tr>
<tr>
<td>Historically Confucian (= 1)</td>
<td>0.702* (0.32)</td>
<td>0.62* (0.243)</td>
<td>0.688 (0.38)</td>
<td>-0.106 (0.485)</td>
<td>1.211*** (0.272)</td>
</tr>
<tr>
<td>Historically Catholic (= 1)</td>
<td>-0.274 (0.16)</td>
<td>-0.127 (0.181)</td>
<td>-0.153 (0.213)</td>
<td>-0.494 (0.308)</td>
<td>-0.082 (0.169)</td>
</tr>
<tr>
<td>Historically Protestant (= 1)</td>
<td>0.449* (0.194)</td>
<td>0.489* (0.199)</td>
<td>0.497* (0.236)</td>
<td>0.429 (0.344)</td>
<td>0.38 (0.191)</td>
</tr>
<tr>
<td>Ex-Communist (= 1)</td>
<td>0.25 (0.156)</td>
<td>0.197 (0.162)</td>
<td>0.243 (0.22)</td>
<td>-</td>
<td>0.213 (0.141)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.958** (0.329)</td>
<td>-0.968** (0.332)</td>
<td>-1.109 (0.551)</td>
<td>0.523 (1.553)</td>
<td>-0.333 (0.241)</td>
</tr>
</tbody>
</table>

| N | 75 | 61 | 46 | 17 | 69 |
| r² | 0.49 | 0.44 | 0.39 | 0.69 | 0.42 |

Table 2.2: Coefficients from Determinants of the Cultural Compliance Index. * significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level.
inherent civilizational differences, but rather, may simply be reflective of long-run legacies of the institutional environment. In societies in which states have existed over a durable period, and have provided incentives against non-compliance with political authority, individuals express greater sensitivity to political and moral conformity than those in which formation of political order is a relatively recent historical phenomenon. In particular, after taking account of the effects of state formation, the impact of the Confucian variable is diminished, suggesting that much of the effect that is attributed to “Asian values” can instead be attributed to the legacy of a long process of state formation in northeastern Asia. But a similar general relationship between state formation, or state failure, and cultural change is likely to exist across all cultural zones.

**Evidence of Variation in Compliance Norms in India**

The regressions in the previous section show that macro-level differences in compliance behavior and norms exist, and can be studied at the country level. But what about variation within a large country such as India, accounting for a sixth of the world’s population, and in which subnational units are often as diverse and as demographically significant as entire countries in the world? Evidence of such variation can be found from a range of sources, including outcome-based measures such as statewide differences in taxation, public goods, or absenteeism. It is also possible to observe differences in compliance norms more directly, through direct tests in the field. In 2005, for example, the Government of India introduced an ambitious policy to encourage transparency among public officials and agen-
cies, known as the Right to Information Act (RTI). Henceforth, any of the country’s 1.2bn citizens could write an official letter to any of the state’s offices, requesting information on their activities, budget, or other information not restricted by legislation concerning official secrets, while public servants have a legal obligation to write a letter of response. While similar initiatives had been piloted in western democracies such as Sweden and the United Kingdom, no such attempt had been carried through in a vast developing democracy such as India, where widespread problems of public service corruption, cronyism and employee absenteeism would be likely to affect the implementation of the initiative. A decade after its passage, compliance with the RTI varies massively; for example, in states such as Karnataka, civil servants respond to thousands of requests each year, while across vast tracts of the country, the act remains a “dead letter law,” with requests for information from the public being routinely ignored. From October 2013 to April 2014, I worked with research assistants to send out letters to each of India’s state and union territory governments, with a simple information request, in order to test this for myself. Consistent with expectations, we found wide variation in the timeliness, completeness, and quality of government responses (Figure 2.4).

i) Why Individuals Comply

The examples given thus far show that macro-level differences in compliance behavior and norms exist at the group level. However, what are the microfoundations of compliance? Why do citizens in some countries choose to comply with regulatory demands made by the state? And what determines why officials in some regions of India respond to information requests, where officials in other place do not?

The historical, sociological, political and psychological literature on compliance offers two approaches for explaining why citizens and bureaucrats might choose to comply or not comply with official requirements and regulations. According to theories of voluntary or “willing” compliance, individuals comply with the demands of authority because they either agree with the normative principles that motivate policies, such as support for the collective provision of public services or a belief in government transparency, or because
Figure 2.4: Compliance with the RTI Among States and Union Territories. RTI compliance responsiveness index by state. RTI requests were made from October 2013 to April 2014. Governments were assigned a first point if a letter reply of any kind was received to the initial RTI request; a second point if the letter reply contained the relevant information; a third point if state governments implemented the RTI Act requirement of providing information officers by department for which the letter is meant to be addressed; and a fourth point if both mail and electronic submission systems are offered. Partial point were assigned for partial fulfillment of each criteria.
they hold affective ties to the community that the political authority represents, such as a sense of national identity (Levi et al. 2009, Levi 2015, Norris 1999, Dalton 2004). Either of these factors provides states with the legitimacy that grants elites the right to govern, and this in turn, “bolsters willing obedience” (Levi, Sacks and Tyler 2009). On the other hand, according to rational choice theories, what matter are the incentives and disincentives that individuals face, notably the sanctions, threats, fines and other punishments that are delivered for disobedience with its laws or insubordinance in office, as well as the credibility of its means of enforcement (Becker 1969). On this account, individuals comply because they face a credible threat of sanction.

The simple rational choice approach is perhaps the easiest to understand. Gary Becker’s (1969) original formulation proposed that the decision to comply is a trade-off between the cost of fulfilling a demand (for example, the time to write a response to a request for information) and the expected cost of not doing so (a function of the likelihood of being detected and the resultant sanction). Because compliance is costly, this gives individuals a rational default option to not comply; so compliance will only ever occur, therefore, when authorities impose costs for non-compliance, and are able to monitor non-compliant behaviors (Becker 1969). Yet behavioral political scientists, concerned with explaining variation in levels of civic behavior, have generally eschewed this model, largely because it greatly under-estimates the levels of actual, observed compliance that citizens routinely make with state demands, and also fails to explain why patterns of variation would exist at all in contexts where the formal institutional structure is held constant (Almond and Verba 1963, Putnam 1993). Scholars of tax compliance, for example, calculate that no wealth-maximizing individual would ever choose to pay income tax, as the fines for non-compliance are in most countries far less than the expected year-on-year windfall from evading payment, at least in most developed democracies (Braithewaite 2003, 2009). A simple calculation can show that in almost all major cities, it is never economically rational to pay fares on public transportation, as the penalty fine for evasion spread over the probability of inspection consistently undercuts the per-trip cost of a ticket. The surprising fact that people continue to pay taxes, and continue to purchase tickets on public transport, is one of the arguments that theorists of voluntary compliance cite in support of their view that “non-coercive” factors explain civic willingness to comply with the law (Levi 2005, Braithewaite 2013). Mean-
Legitimacy Type | Attitudinal Dimension | Examples
---|---|---
Government | Support for policies | favor taxes for service delivery
 | Support for ruling party | voted for incumbent party
Regime | Acceptance of democratic mandate | belief in democracy
State | Patriotism | Sense of national identity

Table 2.3: Easton’s Three Forms of Legitimacy and the Bases of Compliance.

while, it is unlikely that civil servants in India would reply to RTI requests either, given the extremely complex process of prosecuting cases of non-compliance\(^\text{11}\).

Scholars who emphasize voluntaristic reasons for compliance instead start from the premise that government rules and decisions are followed because they are viewed with moral content, that is as being legitimate, or possessing the right to be obeyed. This right, following Easton’s (1957) classic distinction, is seen as pertaining either to the government, the regime, or the state: i.e., in the citizen’s support for a particular administration’s policies or their affective support arising from a certain model of governance, or their sense of membership of the political community in general (Table 2.3). Any of these three forms of legitimacy is seen as encouraging individuals to lend their support to the decisions that emerge from the political decision-making process, such as obligations to fulfill the financial and physical resource demands of political authorities, or to refrain from acts that authorities declare illegal. If political authority is not seen as legitimate, then individuals may feel less compelled to fulfill such demands (Esaiasson and Ottervik 2014). A number of empirical papers have attempted to show that legitimating beliefs are significantly associated with compliance behaviors: for example, Tyler (2006) and Sunshine and Tyler (2003) report that legitimacy shapes self-reported compliance with the law, while Levi and

\(^{11}\)The RTI law in fact gives officers of state two opportunities to reply, as following a month from the original letter the citizen has to send another, and only a month after this point, is allowed to initiate legal action. Yet legal actions in India also take months or even years to clear through the court system, such that prosecutions regarding RTI requests are an exceptional occurrence.
Sacks (2009) report a link between acceptance of the tax department’s right to make people pay taxes, and attitudes towards the government, implying that citizens will be more likely to practice fiscal compliance when a government is trusted.

Yet if simple rational choice theory consistently under-predicts the amount of civic compliance in the world, then behavioral approaches based on normative beliefs consistently over-predict. If payment of taxes, for example, were genuinely voluntary in the same fashion as charitable giving, then it would be difficult to explain why charitable donation in most countries never exceeds one to five per cent of gross national income, while taxation can rise to 50 per cent and beyond\(^{12}\). Nor could we be greatly confident that taxation would for long remain at present levels, were citizens presented with a discrete and anonymous opt-out from contribution\(^{13}\). Second, if individual compliance was simply a matter of fulfilling individually-held policy preferences, then we would expect qualitative variation in compliance decisions to map the quantitative variation in legitimating beliefs. For example, an individual who felt especially strongly regarding the need for redistribution of income might opt to pay an additional increment of her income in taxes, while an individual who felt less convinced in this view, might opt to pay the officially mandated amount\(^{14}\). In reality, individual compliance actions do not map the intensity of preferences at all, but only the structure of coercive demands made by the state, with the implication that the credibility of state coercion matters more than the pattern of voluntaristic attitudes and

\(^{12}\)While it might be argued that charitable giving would be higher if the state did not engage in public healthcare, income redistribution, or education, the two are clearly not effective substitutes as a reduction in taxation of 10 or 20 percentage points of GDP would likely only produce a very small corresponding increase in voluntary donations.

\(^{13}\)The limitation of the voluntaristic argument becomes clearer when we consider the counterexample of a context in which legitimating beliefs are held constant, but state coercive sanction is allowed to vary. In Brussels, Belgian nationals employed by the European Commission are exempt from the legal obligation to pay taxes on their official income, in common with European Union employees from other member states. How many in fact do so? If voluntaristic theories were correct, we would expect at least some of these individuals to opt voluntarily to pay taxes alongside their co-nationals, who do not work for international organizations. The fact that none, in fact, make this choice, would strongly indicate that feelings of national pride, identity, and belonging, as well as normative attachment to the output of public agency such as public transport, social insurance, or policing, are at best weakly predictive of compliance behaviors, compared to the existence and credibility of legal sanctions for non-contribution to the public purse.

\(^{14}\)Another example: if taxation were voluntaristic, we might expect that when a government abolishes a tax, a significant proportion of civic-minded individuals would continue to pay it anyway.
beliefs\footnote{There are a number of further problems to the voluntary compliance viewpoint. For example, if collective identities determine willingness to comply, then we would expect ‘cultures of compliance’ to be well bounded by the scope of the political community. Yet this is not consistent with observations of individual behavior. Japanese citizens may feel a strong sense of solidarity within the Japanese nation-state, for example, yet this cannot be the proximate cause of compliance as Japanese travelers continue to obey the laws of London or Rio de Janeiro. Meanwhile Japanese diplomats were one of the few nationalities in Fisman and Miguel’s (2007) study to have no parking violations in New York City, for example, despite possessing diplomatic immunity. That the strongest predictor of diplomats’ compliance in this study was the degree of the “rule of law” in the home country, implies that domestic institutions generate cultures of compliance that are transitive to other contexts (Fisman and Miguel 2007). This applies temporally as well as spatially: changes in the identity and structure of political authority, such as occurred during the transition from precolonial, to colonial, to postcolonial rule, can be marked by persistent cultures of compliance (Dell 2010).}. Finally, there often wide discrepancies between individually expressed beliefs and actual compliance behaviors. Surveys in many countries show widespread support for public good provision and high levels of collective identity, yet fiscal and regulatory non-compliance remain widespread\footnote{We might expect a macro-level relationship between strength of collective identity and willingness to fulfill the demands of the state. Yet in countries like Germany or the Netherlands, fiscal and regulatory compliance are high yet surveyed patriotism low; in Greece or Turkey, patriotism is by any measure high, yet fiscal and regulatory non-compliance are widespread.}.

A number of scholars in recent years have attempted to resolve these observational inconsistencies by proposing an alternative explanation in terms of collective action dynamics (De Cremer 1999, Rothstein 2005, Herreros 2004, Parks et al. 2013). According to this viewpoint, it is taken as given that individuals feel they would be made better off by working collectively in pursuit of certain goals, such as the provision of security through mechanisms of justice and policing, or the supply of public goods such as roads and vaccination programs - yet they are also aware that this provision will only succeed if every person commits. The trust paradigm therefore asks us to see compliance as a question of whether the individual citizen “trusts” her peers to equally co-contribute to public funds and enable authorities to deliver public goods and services (Glaeser et al. 2000; Anderson et al. 2004; Herrmann and Thoeni 2011). Whenever citizens are confident that their actions will be coordinated, collective action and the provision of public goods is possible, and citizens will comply with collective rules. Where individuals are unable to trust one another, underprovision, shirking, and ‘tragedies of the commons’ are likely, in which shared resources are depleted by overuse (Ostrom et al. 1999). In countries such as Sweden or the Canada, a trust equilibrium exists whereby all members of the political community know...
that their contributions to the public purse will be rewarded through efficient government and the provision of public goods. In countries such as Greece and Argentina, a non-trust equilibrium has emerged in which members of the political community expect that public funds will be misallocated or squandered. This paradigm has become the dominant mode of thinking about public goods provision, whether in the form of experimental games (e.g. Glaeser et al. 2000; Herrmann and Thoeni 2011), or in terms of the explanation given for why social cohesion, intra-group homogeneity, and shared identities affect civic cooperation (Alesina et al. 2004; Habyarimana et al. 2004).

The collective action perspective on compliance has yielded a number of important insights, and helps to explain several observable puzzles such as the formation of persistent compliance equilibria over time, the stickiness of compliance norms, and the observed inconsistencies between public preferences and general behavior, at least in many societies (Rothstein 2005). However, a major problem for the collective action model is that it only works if we assume from the outset that citizens share altruistic social preferences. That is because citizen compliance in a modern society is very different from the small-scale collective action scenarios typically modeled in experimental games, or in the prisoners’ dilemma, in which peer-group enforcement is possible. In such a small-scale reciprocal exchange, the parties are limited in number, and, at the end of each round, can know whether other actors had chosen to cooperate or defect, allowing cooperative or non-cooperative equilibria to emerge (Axelrod 1986). But the individual citizen’s decision to comply or not with millions of other citizens living in a polity, for example by paying taxes, has almost no marginal effect on the payoffs to decisions taken by other citizens, and therefore no bearing on whether she individually will continue to enjoy collective provision in future rounds. Nor (unlike in a community context) could they expect any selective benefit or reward for doing so, as the non-rival and non-excludable nature of public goods makes it impossible for individual honesty to be rewarded. In the absence of third-party enforce-

---

17 As such, in the absence of state sanctions the compliant citizen is consistently a ‘dupe’, regardless of whether other actors comply or not. In a high trust society, the individual citizen could always simply choose not to contribute and then continue to receive public goods anyway, as their individual decision to defect would not be enough to alter the incentives of other actors in future rounds.

18 The only exception is in clientelist or patrimonial states, which is discussed below.
ment, therefore, even starting from a position of universal compliance, the rational choice of every individual will always be to defect in future. For example, if a person moves to a generous welfare state such as Sweden, the fact that the majority of citizens pay their taxes and fund the ample provision of public goods, from which she herself benefits, still gives her no personal incentive to contribute. She can simply not do so – allowing other actors to continue paying taxes and funding public goods, and acting as a free rider on their generosity. Nor does the fact that the state is non-corrupt, efficient, or fair provide any additional incentive to change this decision. If every actor behaves rationally in this way, the compliance equilibrium will rapidly erode. Her only incentive to co-contribute will arise if pooled resources are also used to fund coercive (tax and judicial) agencies, with the ability to monitor funds flows, suspend payments, initiate investigations, and in this way, bring prosecution against defectors for shirking legal fiscal obligations. Yet this leads us to a very different kind of argument to that presented by the social trust and collective action framework. As I will argue in the next section, this is more akin to a principal-agent scenario, and collective action dynamics in the maintenance of compliance only relate to reducing the state’s costs of enforcement, and not to the incentives of individual citizens to contribute to the public good.

*The Role of States – A Coercive Theory of Compliance*

Instead of seeing compliance with the state as a public goods game, I argue that we should see compliance as a principal-agent dilemma between two structurally different kinds of actors: one, the state, with the means of coercion, surveillance and sanction; the other, the citizenry, in possession of human and economic resources that the state requires (Reinganum and Wilde 1984). The state, with assumed authority for collective decision-making, devolves responsibility for implementing laws and mobilization of resources to its agents and subjects, in conditions of imperfect information and uncertainty regarding implementation. It is not, in this view, simply a “third-party” mechanism for solving the collective action dilemma among citizens, but an organization governed by a political elite which possesses independent objectives of its own. The state may choose to provide public goods, but it may equally seek to leverage resources for its own gain, or rather, for the gain of the elites who control it. It may be a democratic state, in which case these elites
are periodically accountable to the voter through elections; however it is not a voluntary collaborative enterprise, but only an organization with the means to coerce and extract compliance from the members of other organizations, primarily firms and households.

Framing the question of compliance in principal-agent terms - with the state as principal, and the individual citizen, the agent - offers a model closer to classical rational choice accounts for thinking about the question of citizen cooperation with authority. The decision whether or not to comply \( Y \) is a function of the sanction \( S \) facing deviant individuals, the probability of detection \( p(d) \), and the size of the demand placed upon the agent by the state \( C \):

\[
(1) \quad Y = p(d) \cdot S - C
\]

In this framework, for the individual citizen, the act of compliance always incurs a cost. In some cases, such as fiscal compliance, that cost - the sum of the tax - is easy to quantify, while in other circumstances, less tangible costs of time, effort and forgone opportunity are involved. Yet because compliance is costly, this gives individuals a rational default option of not to comply. This imposes a need upon the state, as “principal,” to impose costs for non-compliance, and engage in the monitoring and punishment of non-compliant behavior, for example as administered by its police and its judiciary. The interaction of the probability of detection \( p(d) \) and the size of the sanction \( S \), provides an expected cost to any given act of deviance \( E(C_d) \), and individuals would rationally non-comply whenever this falls below their opportunity cost, reflected by the scale of what is demanded \( C \) of the citizen. This brings us back to the framework of Becker (1969), whereby in some circumstances, and perhaps most of the time, a rational self-interested individual will find it prudent to obey the law. Yet in many instances they do not, and examining different influences upon probabilities of detection \( p(d) \) and the scale of sanctions \( S \) allows us to examine different configurations and levels of compliance.

How then do we deal with the key problem faced by classical rational choice theories of compliance – that compliance behavior is far more widespread than would be predicted.
on the basis of coercive sanctions alone? Rather than resort to the addition of legitimat-
ifying beliefs, a better solution is to add parameters from behavioral economics such as loss
aversion, the non-monotonicity of preferences, and imperfect information, in order to al-
ter the decision-making calculus, in a way that predicts greater levels of compliance among
citizens (Kahneman and Tversky 1979). Loss aversion means that individuals are more sen-
sitive to losses (such as fines or punishments) than to gains, and is now a well-established
feature in experimental psychology (Gächter et al. 2009). Non-monotonicity means that
individuals are not linear in the shape of their preferences, and may be disproportionately
more sensitive to certain kinds of losses (such a period in prison) than to others. Finally,
imperfect information means simply that individuals most of time lack good knowledge of
the likelihood of being detected for wrongdoing, or the exact punishment they will receive,
and must therefore rely on some variable estimation of these parameters in their decision-
making, with some informal margin of error (Aliev and Huseynov 2014).

The combination of loss aversion, non-monotonicity, and imperfect information means
that individuals may weigh the worst-case scenario more heavily than mean expected out-
comes, because they now must factor in the possibility of incurring a large loss as a result of
any decision to defy state authority. For example, in the case of cheating of a fare on public
transport, because individuals are loss averse, on any given ride they are highly sensitive to
the possibility of incurring a fine that may run into hundreds of dollars or euros; and even
though the possibility of inspection is remote, due to the non-monotonicity of preferences,
this “worst case” outcome is weighed far more heavily in the decisional calculus than could
be justified by its low likelihood (and therefore, the low resultant expected cost). Finally,
due to imperfect information, some riders may also be unable to know their odds of de-
tection, making it even harder for them to dismiss this worst-case outcome as improbable.
Most commuters, therefore, “play it safe” by buying their tickets; though a small minority
will eventually learn when and how to engage in fare evasion, by estimating correctly the
odds of detection (including the most likely times and routes of inspection), and realizing
that size of the fine is after all both certain and small (limiting the extent of the worst-case
outcome from doing so)\textsuperscript{19}. We may call this decisional calculus the precautionary prin-

\textsuperscript{19}The triad of loss aversion, non-monotonicity, and imperfect information also explains why psychopaths
principle, according to which any defiance of authority is to be avoided, unless there is near certainty of avoiding sanction, or there is an expected cost of (close to) zero\textsuperscript{20} (Epstein 1980; Sunstein 2003; Sachs 2011). The prevalence of individuals following the precautionary principle is essential for understanding both why observed levels of compliance in many countries is surprisingly high, and also why the institutional environment provided by historical states is so fundamental in generating persistent cultures of compliance.

**How Do States Generate Cultures of Compliance?**

The discussion thus far has not yet allowed us to explain variation in patterns and norms of compliance. If we start from the perspective of citizens deploying the precautionary principle, then there are two reasons why individuals choose to comply: either because the penalties for non-compliance, once detected, are very high, or because any act of defiance has a high likelihood of detection. This gives us a basis for understanding why patterns of political authority are so critical in the generation of cultures of compliance. In weak and despotic states, the only way that authorities can maintain a norm of compliance is by ensuring that the costs of compliance are steady and small, while ensuring that that the cost of non-compliance are uncertain and potentially high (e.g. imposing arbitrary executions), as infrastructural power is too underdeveloped to ensure a high probability of detection, and acting too predictably would risk letting subjects calculate their odds of evasion. Maintaining an uncertain prospect of punishment explains why an extremely harsh sanction applied to a few individuals is often more effective than a small, universal sanction that is fairly applied\textsuperscript{21}, and why despotic rulers have long heeded Machiavelli’s advice that are more likely to engage in criminal acts: because absence of shame entails monotonicity between good and bad outcomes, and therefore, absence of loss aversion. Rational neutrality between risk and reward scenarios means that psychopaths are by consequence more prone to risk-taking activity.

\textsuperscript{20}In many instances the precautionary principle is not utility maximising, however, as it entails the steady cost of compliance, such as payment of taxes, when there was a good chance of “getting away” with non-compliance.

\textsuperscript{21}For example the Roman tactic of “decimation” or the “removal of one-tenth” was admiring cited by Machiavelli (Discourses on Livy; 3:49) as an instance in this regard; rather than punish an entire military cohort for capital offences, such as mutiny or desertion, the Roman armies would execute one-tenth of the men, to serve as an example to the others.
it is “a very wise thing to simulate madness”\textsuperscript{22} (Machiavelli 2009).

As states develop, however, they also develop what Mann (1986) refers to as their “infrastructural power” in terms of agencies of enforcement, records of communication, or their monopoly of certification over work, residency, and service access, and this makes detection for illegal behaviour, tax evasion, or absenteeism significantly more likely than before. This allows states to exert greater power over individuals without the visible appearance of such power, or, in the terms of equation (1), to substitute $p(d)$ for $(S)$, holding the multiplicative term, $E(C_d)$, at least constant, if not a great deal higher. This is one interpretation of the argument developed by Foucault (1976), for whom the permissiveness of the modern regulatory state hides a subtle magnification in its powers of monitoring and control, and is largely a simple descriptive fact. As states have improved at surveillance, their capacity to produce compliance has historically increased, and the scale of demands made of the average citizen, in terms of taxation, enlistment, or regulations concerning personal conduct, has also risen\textsuperscript{23}.

A further dimension to the rising infrastructural power of the state is rising civic uncertainty surrounding the state’s forensic capabilities. Whereas in polities at a very low level of political technology, the monitoring of individuals is exerted by means of personal observation and hearsay, the monitoring capabilities of the modern state greatly surpass the knowledge of the average citizen. This generates further uncertainty regarding the expected cost of defiance ($E(C_d)$), not as a result of varying the size of punishments, but through lack of awareness of the precise mechanisms of the state’s surveillance, audit, and enforcement capabilities, making it difficult for all except specialists to assess whether evasion is in fact probable in any given instance. It is this constant possibility of inspection that makes modern systems of “voluntary” compliance, such as self-assessment tax reporting, function. Bentham’s panopticon, cited by Foucault (1976) - a form of prison design in

\textsuperscript{22}I.e. that appearing erratic and unpredictable generates constant fear of disproportionate retaliation (Machiavelli, Discourses on Livy; 3:2).

\textsuperscript{23}This is obviously the case in the example of totalitarian regimes, but also true in the manner that twentieth-century democratic states have imposed increasing restrictions on individual freedom with respect to use of recreational drugs, personal responsibility in situations of potential self-harm, speech acts considered hateful or anti-social, exchange of sexual services, possession of weaponry and firearms, or the use of and provision of misleading information.
which the prison guard is able to stare directly into the cells of a large number of prison inmates without the latter being able to look back - in some ways provides an analogy for the functioning of the modern state. Because the inmates have no conception of whether or not they are under observation at any given time, they must act constantly as if they were; or rather, those who follow the precautionary principle (to only defy authority in situations where the expected cost is close to zero) will never attain a threshold at which it is rational to engage in a defiant act. Likewise under the modern, bureaucratic state, citizens are kept in check by the background awareness of the state’s infrastructural power, and that even minor deviations can be potentially uncovered. Due to loss aversion, few individuals will incur the potentially large costs that are involved in revealing this information (i.e. in personally testing the boundaries of enforcement agencies), which leads such activities to therefore increasingly become the domain of specialized organizations – either organized criminal groups, corrupt branches of the state itself, or corrupt branches of a foreign state – with the capability of estimating reasonable odds of detection, establishing economies of scale, and thwarting the most commonly deployed surveillance measures.

ii) Why Compliance Norms Persist: Compliance Equilibrium

The question of why compliance norms exist is complemented by the puzzle of why compliance behavior persists. While most scholars agree that states are capable of generating compliance norms through the application of sanctions, a longstanding controversy concerns what happens once the coercive apparatus is removed or relaxed. According to the theory of willed or voluntary compliance that is dominant in behavioral political science, the effect of a coercive state is to habituate individuals into rational calculation of the costs and benefits of compliance, eroding the “good,” voluntary and non-rational bases of cooperation, such as shared identity or communal ties, which are seen as more durable in the long run. Once these incentives no longer exist, individuals rationally, and rapidly, revert to non-compliance. Putnam (1993) famously made such an argument with respect to the Norman states of southern Italy, arguing that the hierarchical authority relations of the feudal state undermined local reciprocal ties, and while this created a strong and powerful regime for some time, it ultimately resulted in a rapid decline into criminality and amoral
familism following Italian unification (Banfield 1958). Rose (1994), Rose et al. (1997) and Howard (2003) make a similar argument with respect to the communist regimes of Central and Eastern Europe, in that the pervasive monitoring of the authoritarian state led individuals to see this sanction as the only basis for compliance, resulting in widespread non-compliance once the regime had collapsed. Boix and Posner (1998) and Huysseune (2003) similarly argue that the state is at best a potential destroyer of social capital and trust, while Ullman-Margalit (2004: 75), for example, argues that states do not generate social trust, but simply render such trust unnecessary, by socializing individuals into the expectation of sanctions against untrustworthy behaviour. This constitutes what Levi and colleagues (1998) term “cooperation without trust,” which, under the aegis of a coercive state, we may also simply call compliance. According to such a view, the public order of the despotic state is merely a chimera, an illusory order, that undermines the long-term basis for voluntary adhesion to the public good.

I believe that this view is greatly mistaken, and that compliance, once generated, tends to persist. While it may be true that authoritarian states undermine or supplant horizontal and civic ties (cf. Bernhard and Karakoç 2007), the hypothesis that strong and centralized states undermine their own long-term ability to govern stands quite poorly as a generalized hypothesis. The historical evidence provides many more examples of the persistence of compliance, than of high compliance societies which collapsed into state failure. As argued by Ertman (1997) the distinction between bureaucratic and non-bureaucratic state structures may be more important than the degree of political openness, as bureaucratic states, even when authoritarian, build infrastructural capacity. Therefore absolutist regimes in Europe often gave rise to bureaucratic states, as in the cases of Denmark or Sweden. Historical studies by Jensen (2014) and Bagenholm (2014) show that absolutist rulers in Denmark and Sweden were able to deliver a large reduction in corruption, that were durable following transition to a less authoritarian regime. Fukuyama (2014) has made a similar argument regarding the sequencing of state formation and democratization more generally: in countries such as Greece, which democratized prior to state centralization, clientistic patterns of rule persisted into the modern era, whereas in cases where absolutism first established a central state, such as Denmark, democratic institutions emerged without clientelism or patronage. A similar argument was made by Huntington (1968), who reprised the
argument of Elton (1953, 1960) regarding the legacy of Tudor absolutism for norm compliance in England, and added its transferral to the founding settlers of the United States. Outside of Europe, Gerner (2012) shows that non-European countries with a longer history of state formation were significantly more likely to become stable autocracies, implying again a persistence rather than a reversal of norms of public compliance with authority.

This can be shown considering both longitudinal and cross-sectional data. Gurr (1981) for example shows that a large decline in homicide occurred under English state formation during the late medieval era, while Eisner (2004) replicates these results for a large sample of countries in Europe. Finally, cross-country evidence also supports such a hypothesis: there is a striking correlation between measures of historical state formation, such as the State Antiquity Index developed by Putterman (2014), and measures of crime such as cross-country homicide rates. Further to this, authors have challenged the explanations of Putnam (1993) or Rose (1994) regarding the breakdown in legal, fiscal, and bureaucratic compliance in southern Italy and eastern Europe. The few examples favored by Tocquevillian social capital theorists often have alternative explanations for subsequent state weakness. Ziblatt (2006) shows that regions of southern Italy had weak state capacity, as measured by numbers of bureaucrats, soldiers, and taxation, even in the period before Italian unification, with the implication that the low compliance of the Italian south may be a persistent legacy of state weakness. Putnam’s narrative of regional variation in Italy stands the facts on their head: in the period preceding Italian unification, the northern polities of Piedmont and Tuscany were the best examples of centralizing, bureaucratic, and absolutist regimes, while the southern Kingdom of Naples was a decentralized, feudal regime, whose very absence of infrastructural power allowed the development of organizations such as the mafia.

\textit{The Formation of Compliance Norms}

The persistence of compliance norms in spite of changes of institutions and incentives requires additional mechanisms to be added to the simple rational choice framework outlined in the previous section. Thinking about compliance in terms of monitoring and en-

\footnote{It might be added that, while Putnam (1993) had cited Machiavelli in support of northern Italian republican government, Machiavelli’s own contrast was not with Naples, but rather with France, which no-one today would cite as an example of a weak state.}
forcement fails to account for why such behaviors would persist over time, or explain how patterns of compliance can vary within a single formal institutional environment. Therefore, to understand variation in compliance, it is necessary to explain the formation of compliance “cultures.” In the first instance, exposure to the state’s regulatory control does not simply generate a stream of successive compliance decisions, but also leads to the formation of compliance “norms”: default modes of action, or behavioural heuristics, in response to an assessment (conscious or otherwise) of the relative costs and benefits under some recurring situation. If a person arrives in a new city in the United States, for example, they may initially avoid jaywalking because they do not know whether such laws are enforced, and continue to do so for a long while (leading to a large cumulative time loss) because updating their prior assumption is too potentially costly at any given moment, relative to the cost of waiting a few extra seconds for a cross signal. Such a norm may originate as an assessment of relative costs and benefits, but persist long after the balance of incentives has altered, as long as information remains imperfect and there remain search costs to updating this knowledge. Behavioral economics has also revealed another feature of human psychology which makes such norms stubborn, which is that of time inconsistent preferences - in particular, weighting present returns over future returns, or preferring an immediate, small loss or gain, for example, over a larger upfront cost for a large net future gain (Loewenstein et al. 2003). Because updating priors involves an immediate cost in the present while payoffs to behavioral change lie cumulatively over an indefinite future, the sub-optimal equilibrium for many individuals is simply to continue complying with the law. Regarding compliance with public authority in general, assessments are made throughout adolescence and adult life whether to comply, based on limited or imperfect information regarding enforcement. As long as the demands of public authority remain low,

25The classic demonstration in psychology is the phenomenon of ‘learned helplessness’, whereby animal subjects ‘give up’ on engaging in actions that would avert a negative stimuli, after a series of trials in which efforts to do so are frustrated (Seligman 1967, 1975). In this case a norm develops (not to make further efforts) and is sustained based on the experience of a few failed attempts, when continued trial and assessment would have allowed for a better assessment of the situation.

26The irrationality of human time preference can be used to explain the persistence of ‘bad habits’ that harm individuals in the long run, yet, as in this instance, also explain the persistence of ‘good habits’ that benefit society. Most individuals are ‘addicted’ to the law - continually opting for the smaller, short-term pain of state compliance, rather than incur the larger upfront cost of discovering individual freedom of action.
and information regarding methods and strategies of enforcement remain uncertain, norms of compliance are likely to persist.

However, norms alone are unable to explain persistent cultures of compliance because as soon as a state begins to make increased demands, for example by introducing compulsory service (e.g. corvée labor or conscription), or individuals become more familiar with the enforcement strategies of public authorities, a threshold may be crossed at which rational choices regarding the decision to comply become more frequent, and instinctive norms of compliance begin to erode. Additional theories are needed to explain why individuals continue to comply.

One such theory is that once compliance has been generated, it constitutes an equilibrium, because the ability of the state to enforce compliance depends upon collective action dynamics among the population as a whole. The expected cost of tax evasion is quite low when every other citizen is avoiding to pay their taxes - the state, evidently, cannot prosecute everyone at once - but the expected cost becomes rather high when one is the first person to break rank. Similarly, a single individual throwing objects at the police is at high risk of arrest: an entire crowd is not. In this way, there are collective action dynamics to public compliance, as the social trust paradigm has proposed, but regarding enforcement rather than cooperation. If a citizen believes that other citizens are not going to pay a new tax, they are indeed less likely to do so themselves. But this is not because they fear being a ‘dope’, as the social trust argument asserts (by the logic of rational choice, they will always be a ‘dope’ if they had the option of being a free rider, regardless of whether others comply or not). Rather, it is because once others defect from payment, they are also more likely to get away with it, whereas were others to keep paying, their chances of discovery, shame, and sanction are significantly higher.

Second, and by extension, this equilibrium dynamic is extended further by the “multi-level” nature of the compliance game (Putnam 1988). An individual’s incentive to comply with the law depends on the investigative and enforcement powers of the state, but the state is engaged in a compliance dilemma with its enforcers: in their capacity as its agents, they may also choose not to comply, and in this way alter the incentives for those who are under their jurisdiction. The decision of the police to enforce depends both on the sanctions placed upon them by their authorities, and the level of sanction or resistance from society.
If social actors are resistant, this increases the incentive of the police to forebear rather than enforce. This is another of the attributes of the non-compliance equilibrium: the degree of non-compliance determines not only the likelihood of individual sanction, but also the absolute level of enforcement. Intermediate authorities are likely to give up on attempting to enforce laws that are pervasively violated, and realise that social resistance may be incurred from attempts to shift to a compliance equilibrium. But once social actors have been pacified, the confidence of authorities to enforce laws will increase.

Third, there are individuals in every population prepared to engage in what Herrmann et al. (2011) term “pro-social punishment” - i.e., the willingness to undertake a short-term personal cost in order to sanction those who refuse to comply with the rules of the political community - and such individuals are likely to be more widespread under compliance equilibrium. The increased prevalence of pro-social monitoring and enforcement in a functioning state changes the incentive for individual compliance. In a compliance equilibrium, there are likely to be more individuals who benefit from the state’s provision of public goods, and whose moral and material investment in the status quo will lead them to act as pro-social enforcers. This may be as simple as engaging in minor, low-cost actions such as reporting to the authorities a neighbour who is known to avoid paying her municipal tax or ‘shaming’ associates into obeying local regulations, or as extensive as engaging in direct retaliatory action, that the existence of third-party enforcement by the state typically vitiates the need for such behaviour. In any circumstance, the presence of pro-social enforcement greatly increases the expected cost of non-compliance, and the provision of public goods and services by the state increases the numbers of those prepared to act as social enforcers. As Herrmann et al. (2011) report in their cross-national data on compliance games - there is a high contribution to public welfare and pro-social punishment in countries such as Switzerland and Australia, and low contributions in anti-social punishment is more common, such as Saudi Arabia, Oman and Greece.

Finally, it would not be possible to account for the exceptional persistence of compliance norms without acknowledging the role of belief-systems, as well as the role of states in sustaining and propagating such beliefs, as long emphasized by the Marxist literature on ideology (e.g. Therborn 1979, Althusser 1975). It is easy to comprehend that normative belief-systems, or beliefs regarding ethical or political values, affect willingness to com-
ply in a very direct way, and collective memories, identities, loyalties, and ethical views concerning compliance are fostered by political rulers and their institutions, and these are the arguments made by “voluntaristic” theories of compliance. However positive belief-systems, or beliefs about the world, are also a mechanism for affecting compliance, by altering individual assessments of the likelihood of detection ($p(d)$) and the severity of the sanction ($S$), and have been less an object of study for political science. The paradigmatic example is that provided by religious belief systems, which achieve compliance with their rules entirely through belief in the monitoring and eventual sanction of a deity – without apparent need for actual divine intervention – and for this reason many early states have solved the problem of order on a theological basis. Having officers of state motivated by belief in otherworldly rewards alleviates the sovereign, at least partially, from having to provide such rewards himself, or impose penalties through constant surveillance and discipline of his auxillaries. Yet even where state power is not supported by a complex metaphysical system, ideology has an important role in maintaining compliance. This can be evidenced from surveys which show individuals to systematically overestimate their likelihood of being caught when engaged in a range of non-legal activities, as well as the typical sanction they would receive if caught, suggesting that belief in state power, rather than the reality of such power, remains a basis for adherence to the law (Scholz and Pinney 1995, Alm et al. 1992).

Another example of how belief-systems raise compliance is the example given by terror under totalitarian regimes, which does so not through enforcement alone, but also by changing the perception of the degree of enforcement, for example through show-trials or

---

27Regarding the effect of religious ideology on compliance, Shariff and Rhemtulla (2012) find that, while there is no correlation between religiosity and crime rate, as measured by the level of homicide, there is a strong association between lawfulness and belief in divine punishment (belief in “hell”). The authors conclude that belief in sanctioning explains compliant behaviors, an intuition borne out in an earlier experimental in which Shariff and Norenzayan (2011) found that students who believed in a more ‘forgiving’ God were more likely to cheat on a test. Finally Barro and McCleary (2004) find that while there is no relationship between religiosity and economic growth, nonetheless cultures that believe in hell have significantly higher growth rates, which they attribute to the effect of norm compliance on social cohesion and the rule of law. With regard to early political order, early states have often united political and theological power, either because in Samuel Huntington’s (1996) phrase “Caesar is God” (as in the case of ancient Egypt or China) or because “God is Caesar” (as in the early Islamic caliphates). Belief in divine sanction is a classic means of solving the Platonic dilemma of who ‘guards the guardians’. 
propaganda, which create an impression of the state’s enhanced reach. Thus even if the surveillance capacities of the state are not objectively greater, it is enough for individuals to believe that public authorities have the capacity to detect or expose their wrongdoing, in order for individuals to rationally refrain from such acts. In a fascinating discussion of compliance in Soviet Russia, for example, Herreros (2004, 2006) argues that the effort of the Soviet authorities to engage in “show trials” was done with a conscious goal – to shift expectations among the rest of the population, as the “plots” of which people were accused and confessed, though fabricated, were perceived as real conspiracies that were detected by the secret police. Thus an impression was generated of the NKVD’s omniscience, such that any plot against the regime would invite comparable detection and punishment. Through social history, we know that such regimes contained substantially greater norm defiance than either external or internal actors understood at the time (Peukert 1988, Shulman 2012). But if the total control of the Soviet state was largely fictitious, still it became a little more true by virtue of people’s belief in it, and this is especially so given that a majority of the public, behaving according to the precautionary principle, was inclined to avoid any action that might lead to trouble. Regimes of terror are very effective in raising compliance, and reliance on perceived, rather than actual, capacity only becomes problematic once this belief evaporates.28

Because of the ability of states to alter not only individual incentives, but also norms and beliefs, sociologists such as Norbert Elias (1982) or Michel Foucault (1976) have long studied the ways in which the state, through disciplinary institutions such as schools, prisons, and the bureaucracy, is capable of instilling in individuals a set of behavioral heuristics that guide individuals into compliant behaviors, often against an objective calculation of rational self-interest. As mechanisms of socialization become more extensive and elaborate, the corpus of rules expands; in this way Foucault (1961, 1963, 1975) argued that the rise of the early modern state was accompanied by profound discursive shifts in medicine, demography, and the law towards an expanded range of mental illnesses to encompass the growing

28The swift collapse of communist regimes, once open expressions of resistance were made without consequence, is a case in point. Similar, if less dramatic instances of shifting equilibrium have been seen in societies undergoing a sudden exogenous shock, such as a financial crisis or natural disaster, which not only creates an incentive for a sufficient number of individuals to engage in defiance but also reveals to other actors that such defiance can be entertained without consequence.
categories of social deviance. The increasingly internalized forms of discipline and control he termed “governmentality.” Elias (1982) had made a similar argument regarding social norms, etiquette, and public order (also discussed in Wouters 1981, and Pinker 2011). The state, therefore, exerts a constant disciplinary influence in shaping individual behavior, and subtly augments its range of demands over time, starting from simple injunctions regarding the use of violence or respect for property rights, and extending well beyond this remit to the myriad of demands made by twentieth and twenty-first century states.

Arguments regarding the ability of the state to maintain individuals within such sub-optimal equilibria are also common in classical psychology and sociology. Freud, for example, proposed in his *Civilization and Its Discontents*, that the state sets “up an institution within [citizens’] minds to keep watch over it, like a garrison in a conquered city” (Freud 1929: 30), arguing that the rational self is only one component of the mind (the ego) that is potentially overruled by the socially conditioned portion (the superego). In this way, states ensure compliance by “the heightening of the sense of guilt” that results from norm deviation (Freud 1929: 35). A similar theory of human action was proposed by Emile Durkheim (1893), who described social norms or “social facts” as having an “external” quality that the rational self often simply cannot overrule. In rational choice terms, we can see these as cases where repeat experience has led to the formation of strong compliance norms, where there is constant fear of the possibility of state sanction, and where by consequence individuals have developed strong internal psychological mechanisms for policing personal conduct that are regulated by judgments concerning shame, social desirability, and reputational loss. The theory of compliance proposed here therefore comes closest to that of “traditional” authority, advanced by Weber (1968). Theorists of voluntary compliance have often struggled with Weber’s third form of legitimacy: whereas “legal-rational” legitimation corresponds to the normative compliance assumed by bureaucratic states, and “charismatic” authority with that of patrimonial regimes, the concept of ‘traditional’ authority has been viewed as redundant and underspecified. By contrast, I would suggest that Weber’s inclusion of such a category was necessitated by a need to recognize the myriad forms of routine and everyday authority compliance, that appear unmotivated by either strong normative conviction, on the one hand, or personalistic ties, on the other - and the concepts of the precautionary principle, compliance norms, and compliance equilibrium
come closest to explaining this habitual form of compliance, which in many individuals and most of the time is the basis upon which states secure ongoing adhesion from their citizens and subjects.

To summarize, then, there are also a series of rational and non-rational bases by which the compliance equilibrium is maintained long after the initial period of state formation. These include collective action dynamics, pro-social enforcement, and ideological belief-systems, both positive and normative, which act to maintain the logic of compliance. The concept of the compliance equilibrium solves one of the longstanding puzzles in political science: why are institutions so persistent? Compliance equilibrium means that each citizen is aware of the compliance expectation among other citizens, and none can expect to make great advantage through a lone act of non-compliance, even if a group of citizens might have been able to do so collectively. Because the compliance equilibrium over time becomes sustained by beliefs, identities, and collective memories, the norm of compliance becomes “sticky,” and can withstand a relatively high degree of exogenous shock in the form of temporary incentives to non-compliance. Nonetheless, compliance and non-compliance are equilibria between which transition is possible, given a sufficiently large and sustained shift in incentives. The study of state formation is, in no small part, a study of how such a shift can be produced.

iii) Why Do Authorities Seek Compliance?

In the preceding sections I have argued that a key factor in explaining patterns of compliance is coercion, or rather, the latent yet credible threat of coercion, and that neither citizen awareness of collective action dynamics nor legitimating beliefs can be seen as the primary causes of how compliance is generated, though both have a secondary role in explaining how compliance is maintained. Yet if compliance depends upon elites applying sanctions to their subjects via their control of a state apparatus, then a question that needs to be addressed is why elites would invest significant effort, resources, and risk in securing societal adhesion in the first place. After all, to do so is not without cost. It involves a substantial upfront investment of resources in order to establish monitoring systems that provide information on which to act when deciding whether or not to punish potential
deviant subordinates, to change perceptions of state power, to audit available fiscal and human resources, and of course to distribute acts of punishment (Levi 1998). Assuming the sovereign uncovers instances of wrongdoing, the act of punishment entails a further cost, as beyond the cost of finding and training replacement officers of state, there is also a risk of rebellion. It would only be rational for a sovereign to invest in such resources where the marginal benefit from ensuring additional compliance is greater than their expected cost as a result of these factors, and this is likely to be the case only when there is an extreme opportunity cost to not investing in state capacity\textsuperscript{29}.

The costly nature of institution-building means that the initial strategy for most elites, and indeed, leaders in all organizations, is to adopt patrimonialism - i.e., the provision of selective benefits to known individuals, who in turn act as patrons within their own personal fiefs, and in turn apply a pattern of individualized benefits and sanctions to the known community of individuals under them. The patrimonial strategy has clear and immediate short-term advantages over the impersonal, bureaucratic method, and in particular in situations where there is greater local over global information, where existing informal networks (such as clans or tribes) can be mobilized into service, and where states have limited infrastructure for penetrating local communities and interests. In the context of political authority, by apportioning patronage to key individuals and devolving to them responsibilities for local governance, fiscal mobilization, and military recruitment, it becomes less costly for elites to claim sovereignty over a given territory, as local lords or chiefs are in a better position to know how to maintain political order within their respective domains. In addition, where authorities have large personal discretion how to reward subordinates, individuals have an additional motive for complying with authority, because individuals not only face individualized sanctions for non-compliance, but also receive selective benefits for their record of complying with the ruler\textsuperscript{30}. Where individuals are presumed by the

\textsuperscript{29}There is, in short, an optimal level of non-compliance: an optimal degree to which elites would tolerate the existence of engagement by subjects in non-sanctioned activities.

\textsuperscript{30}Though as Rudolph and Rudolph (1979) and Derluguian and Earle (2010) observe, all societies are patronage based at a sufficiently high point in the political system: even in democratic states governed by the rule of law, premiers must decide how to reward or punish their ministers, and civil servants can influence the career trajectories of their junior staff. The difference, of course, is that local authorities have limited discretion to reward individual citizens, yet even here, this is only a relative statement of fact; police may
patron to be loyal, i.e., bound to a ruler by affective ties, a single act of defiance has the potential to invalidate an entire record of loyal service, as past acts of compliance will be seen to have been motivated by instrumental motives, and not by loyalty. The limitations of the patrimonial strategy are that it entails a greater risk that local potentates may simply declare their independence from the realm, or join under competing powers, and that it places a strict limit on the extractive potential of the central state itself: only a small portion of the revenue mobilized locally within each fief is transferred to the central polity as tax or tribute, while the lion’s share will be enjoyed locally by landowners, rights-holders, and the prebendal officers appointed by the sovereign. Yet patronage-based mechanisms of governance are far more efficient in maintaining rule in polities that lack effective mechanisms of bureaucratic control, or where agents of the state cannot be relied upon to implement rules and regulations\textsuperscript{31}.

The bureaucratic strategy, by contrast, entails large upfront cost. It takes a great deal of material and human resources to recruit, train and equip agents with direct responsibility to the sovereign, and even greater potential cost to wrest power from local potentates and install public officials in their place. Moreover, generating compliance necessarily results in large upfront resistance, both from former holders of patronage, and from broader social actors who may prefer to rebel rather than accept subordination on new terms and incur the likelihood of accelerated state demands. Given that political leaders also exhibit decisional preferences based upon loss aversion and present time preference, there is an important question to be asked: why is it that rulers would go through the effort of establishing disciplinary and coercive institutions to regulate the behavior of their subjects, when the option of delegated authority exists?

The origin of states and of state formation is the subject of a longstanding literature in anthropology and political science (Oppenheimer 1929, Carneiro 1974, Sahlins and Service 1975). Theories of state formation have generally proposed three answers as to why it is be able to offer variable treatment to offenders in exchange for cooperation, and municipalities can exert influence over private actors through the allocation of grants, outsourcing contracts, or information.

\textsuperscript{31}For example, where local lords have relative security in tenure, they may act as what Olson (2000) termed ‘stationary bandits’ and invest in the supply of patronage. Yet insecurity in office - introduced through contestation for control, such as via regular and contested elections - will increase the incentive for local powerholders to behave rapaciously.
Figure 2.5: **States and Compliance.** A simplified schema for understanding the evolution of states and compliance. (1) In pre-state, patrimonial societies (‘chieftaincies’), elites possess few means of detecting the degree of compliance among subordinates, as well as limited means of sanctioning acts of deviance or defiance. (2) Under early despotism, rulers apply large sanctions, such as torture or execution, to raise the degree of compliance among subjects. (3) As states increase their degree of infrastructural power, including powers of surveillance, policing, and discipline, autocratic regimes combine large sanctions with heightened capacity to detect acts of non-compliance. (4) Modern ‘regulatory’ states engage in a selective deregulation of compliance demands, and a substantial reduction in the sanctions applied to non-compliance, while continuing to augment the state’s capacities of monitoring and surveillance.
that elites seek to gain control over their subject population, which we call the “extractive,” “rebellion” and “warfare” hypotheses:

i) First, many theories of state formation have laid emphasis on the extractive incentives that elites face (Marx 1969; Oppenheimer 1929; Olson 2000; North et al. 2006). By building a coercive apparatus, elites can enjoy rents from the taxation of the populace, which can then be used for consumption. The state is therefore a coercive apparatus whose foundational goal is to exploit a population, at the outset largely agrarian, that is capable of being taxed (Scott 1969).

ii) A second school of thought has emphasized revolt, and the threat of revolt, as providing the incentives to heighten compliance, alongside the more minor and continual threats that elites face in the form of rebellion and usurpation (Skocpol 1979). In a period of rebellion, an outlying people refuses to accept the sovereignty of the ruler, who is then required to assemble an armed coalition to re-establish his territorial control. Elites in early states invest heavily in building the army of the central court, and reinforcing alliances among military elites. The suppression of rebellion provides a fairly strong incentive for compliance, though only in respect of those areas for which the sovereign is willing to fight (principally, the payment of tribute, taxes, and soldiers for war). Second, usurpation means the displacement of the sovereign by another member of the elite, through intrigue, assassination, or a palace coup. It provides a very strong incentive for the sovereign to invest in surveillance and sanctioning among the members of his inner circle, and even very early states have had some form of secret police whose attention is directed at this group. Finally, in revolution, there is total mobilization against the regime by mass movements, a phenomenon specific to the modern era, when social mobilization is widespread due to the spread of education, knowledge, and widened political participation (Deutsch 1961, Huntington 1968). The threat of revolution requires regimes to invest in a significantly more extensive apparatus of surveillance and control, not one that is directed merely at members of the ruling circle, but one that can penetrate the entire society. It is for this reason that in the modern era states have adopted more ambitious goals regarding the monitoring of individuals and ideological warfare to maintain systemic legitimacy. Moreover, if
elites will invest in surveillance and control in order to forestall the threat of revolution, this threat is doubly apparent in the case of revolutionary regimes themselves, which from the outset are beset by the risk of “counter-revolution” by the displaced status quo ex ante. New regimes are therefore the greatest investors in increasing the state’s capacity to secure compliance, and it is not accidental that the white and red terrors of the French or Russian revolutions led to the development of conscription, propaganda, and attempts to repress or replace prevailing religious belief-systems (Skocpol 1979). The same is also true regarding ethnonationalist movements in societies where national identities are fragile, or former elites can claim an alternative basis for legitimate rule, or in which there are large numbers of potentially secessionist minority groups (Weber 1978, Akturk 2014).

iii) Third, warfare has also been seen as one of the primary stimuli to investments in state capacity (Tilly 1992, Parker 1996). In the case of interstate war, the elite incentive is clear – military failure entails loss of land, legitimacy, and quite often, one’s life. Where interstate warfare requires increasingly capital intensive weaponry, this requires greater levels of fiscal mobilization, leading to reform of tax administration, and where warfare is based on mass mobilization, such as the deployment of infantry soldiers, it also requires higher levels of manpower mobilization, such as through conscription (Roberts 1955). The intensity of interstate warfare, and the interaction of warfare with military technology, are seen as key drivers of state capacity. Due to the requirement that effective military modernization also requires economic resources, and therefore economic development, “defensive modernization” refers to the adoption by elites of policies designed to accelerate economic growth so as to more effectively counter external threats (Evans 1995).

Which of these hypotheses best helps to understand the circumstances under which political leaders make investments to build state capacity? Extractive incentives have most often been cited with respect to theories of early state formation, such as Wittfogel’s (1957) and Carneiro’s (1970) argument that ease of taxation led to early political centralization, or Oppenheimer’s argument that early states arose where pastoralist groups, with mobility and specialization in violence, encountered agarian societies that they could dominate and
However, the fiscal-extractive hypothesis is limited by the fact that extractive regimes have often remained as patronage states, with limited need for further bureaucratic centralization or universal taxation. *Ancien regime* France and the Mughal Empire in India, at its apex at least, are two frequently cited examples (North and Thomas 1973; Maddison 1975). Many early polities achieved this extractive model of political organization without ever progressing to a more centralized form of political order, characterized by features such as bureaucracy, law, and the centralized payment of salaries and offices throughout the realm, or achieve these only briefly and in part before sliding into political decay, patrimonialism, and collapse. Extractive incentives only touch upon one aspect of compliance - fiscal compliance - and are largely neutral with respect to other forms of compliance, such as regulatory or bureaucratic control.

If extractive-consumption incentives alone are not enough to explain why states enforce compliance, both the revolution and the military mobilization hypotheses make greater sense, and in particular, given the fact that elites experience the same motivations of loss aversion and the precautionary principle that were outlined earlier with regard to citizen behavior. For any leader, the two most high-impact negative outcomes are insurrection, and total military defeat, either likely to result in death and the deaths of many close individuals. We can summarize this view as the “threat hypothesis”: that elites are primarily motivated to invest in securing compliance and raising state capacity when they face threats to their rule. These threats may be external, such as via military or colonial domination, in which case elites must invest in compelling greater compliance with fiscal or manpower demands, or, they may be internal threats, such as insurrection or revolution, in which case elites must invest in surveillance, monitoring and control of their own populus. Often one leads to the other, for as the state’s fiscal and enlistment demands mount, citizens and local elites have a stronger incentive to resist compliance, necessitating even stronger counterpolicies by the central state, which aim at coercing obedience. But whenever the threat exceeds the potential cost from internal rebellion as a result of measures taken to face it (such as imposing new taxes), elites are certain to pursue such measures.

---

32 Oppenheimer’s argument has most often been applied to explain the large number of early states that arose from the domination of central Asian pastoral and nomadic groups, such as the Aryans in India or the Hun in Europe.
The state, then, is an organization with the coercive means to compel compliance from other organizations and individuals, and over course of many generations, elites in charge of this organization have invested in enhancing its capacity to enforce compliance. Because of the long, accretive nature of the “state-building” process, this enhancement of state power is not a single event, but takes place over the course of a myriad of successive, and often individually minor, reforms. However the prompt for such ratchets in state capacity is most often the existence of a credible threat to the political elite, of which the two most important are interstate warfare, and internal insurrection. If we are to explain the variation in patterns of compliance across the world today, one of the most important factors is likely to be the legacy of longstanding historical states in that region. This is particularly the case where such states have engaged in repeated warfare requiring the mobilization of fiscal and human resources, and where pressures from within the elite and society as a whole have exposed authorities to the risk of political instability.

Conclusion: Explaining Historical States and Legacies of Compliance

This chapter has sought to provide a theoretical basis for answering the question: why do individuals in some places largely choose to comply with the demands of public authority, while individuals in other places do not? I have argued that legacies of state coercion are the most important factor in explaining why compliance norms exist, and that good theoretical reason can be adduced to explain this relationship. Addressing limitations with three of the theories of voluntary compliance, collective action, and rational choice, in this chapter I have propose a modified rational choice approach, which I have augmented using insights from behavioral economics, to argue that sustained latent patterns of coercion lead to the formation of compliance norms. I have then examined the theoretical basis for persistence: why it is that compliance norms, or cultures of compliance, are sustained among communities over time despite changes in policies, political regimes, and other alterations in the general incentive structure. Alongside, I have also addressed why is it that historical political authorities have sought to generate compliance from their subjects through means of bureaucratic control and enforcement. A simple index of compliance values, shows that the best predictor of worldwide variation on such norms is a country’s legacy of historical
state formation.

In the following chapters, I put this theory to the test by examining patterns of historical state formation in the South Asian subcontinent and their institutional legacies for districts of the Indian subcontinent.
The preceding chapter has identified historical states as one of the key factors in generating norms of compliance, and shown in a simple cross-country test that compliance norms are significantly associated with a legacy of long-term historical state formation. The purpose of this chapter is to establish how patterns of historical state formation varied within the South Asian subcontinent at the time of colonization, and in particular during the century leading up to the final British consolidation of power in 1858. What were the patterns and configurations of political control that existed in the subcontinent prior to the consolidation of the colonial regime?

An appreciation of the diversity of precolonial India’s political legacies has often been missing from the early historical and sociological literature, which has tended to view the subcontinent as stateless until foreign arrival. India was characterized by nineteenth century observers as a “village” society, each like a grain of sand: empires might make of these grains a castle, but sooner or later, like the statue of Ozymandias, only sand would remain. Thus Marx wrote of how “just as Italy has, from time to time, been compressed by the conquerors’ sword into differential national masses, so do we find Hindustan, when not under the pressure of the Mohammedan, or the Mogul, or the Briton, dissolved into as many independent and connecting States as it numbered towns, or even villages” (Marx

---

1Notable exceptions have been the works of Kosambi and Romila Thapar. Marxist Historians such as Sharma (Indian Feudalism 300-1200 A.D, 1965, Urban decay in India 300-1000, 1987), have also pioneered a debate regarding the origins and existence of feudalism in South Asia.
Colonial administrators and authors like Henry Maine described India as a sea of village communities, each a separate “little republic” disjoined from the rest (Dumont 1980: 158-9; Maine 1871: lecture IV). This view of India as recurring from village to empire to village is also found in twentieth century political sciences. Barrington Moore, for example, describes Indian society as “the starfish whom fishermen used to shred angrily into bits, after which each fragment would grow into a new starfish” (Moore 1966: 458). A recent monograph on global political development covers a millennium of Indian history in just four paragraphs, arguing that Indian state formation consists in “foreigners’ efforts to transplant their own institutions on to Indian soil” and that “after the tenth century, the political history of India ceases to be one of indigenous development” (Fukuyama 2011).

However the indigenous political development of South Asian states did not end in the tenth century, and by the time of imperial consolidation, there was a high diversity in the range and development of India’s precolonial political institutions, from pre-state hunter-gatherer societies in the interior northeast, to feudal regimes in the northern Rajasthani region, to the more centralized, bureaucratic southern states of the Deccan. As a result the British East India Company did not encounter India as a sea of villages, but rather as a set of hardening states which it had to defeat sequentially over more than one century of conflict. Some of these, such as Mysore or the Marathas, could only be subdued following several wars, while others, such as the Gorkha Empire, Bhutan, and Siam, were never fully conquered, instead surviving to form the modern independent states of Nepal, Bhutan, and Thailand, respectively. Other polities, notably the Mughal successor states of Awadh and Bengal, were already so weak by the time of British expansion that they sought protection from European powers, and willingly ceded their powers of revenue collection to officials considered more capable of exerting these capacities. But what was the pattern of state formation at the point of contact? How can we measure the historical political development of South Asian regimes before the 1757-1858 territorial expansion of the East India Company? This chapter seeks to answer this question.
Patterns of Precolonial State Formation

Suppose that we were to conduct an “audit” of precolonial state capacity on the eve of the East India Company’s consolidation of power, taking each of India’s polities and assessing precolonial institutional strength on the following aspects:

i) **The fiscal system.** How were revenues collected? Was it the responsibility of agents of the state, or was it subcontracted to gentry (zamindars), or local lords (jagirdars)? How universal was the taxation system? Were all lands included, or was the assessment system balkanized between tax-free concessions (*inam*) and plots with differential rates of taxation based on caste or religion? How was taxation calculated? Were rates based on land cadasters and collected in dues, or were rates assessed as a percentage of output and paid in kind?

ii) **Military structure and recruitment.** To what extent were military and administrative recruitment organized bureaucratically and meritocratically based? Was there a recruitment procedure, or were military and administrative roles assigned on the basis of inheritance or patronage? Were salaries paid from the central exchequer or based on land grants and tax concessions?

iii) **The structure of the executive and administration.** Did the sovereign rule directly, or was there a chief minister or diwan? How were he and his ministers appointed? Were there mechanisms of executive replacement? Was there a separation between royal (privy) finances and the public accounts? To what extent was there bureaucratic autonomy?

Evidence for India’s precolonial polities can be found in a wide variety of historical sources, including studies of the Maratha peshwa and princes (Gordon 1993; Kulkarni 1974; Sen 1976; Gune 1953), the successor states of the Mughal Empire such as Awadh (Fisher 1987), Bengal (Sarkar 1973; Eaton 1993), and Hyderabad (Nayeem 1985), the Sikh Empire (Singh Bajwa 1964; Singh 1932), Mysore (Rao 1948; Habib 1999), Arcot (Ramaswami 1984), Rajput states including Bikaner (Bhadani 1979), Jaisalmer (Somani 1990), Marwar (Sharma 1977) and Mewar (Paliwal 1970; Mathur 1987; Naravane 1999), the Himalayan Kingdoms of Sikkim and Bhutan (Sinha 1968; Karna 1963; Namgayal and Dolma 1908), Tamil chieftaincies, such as Pudukottai (Dirks 1987), and Travancore (Pillai 1940). Taking a schematic approach to the study of the precolonial state, we can see that there were broad differences in the scope and extent of state power before British rule.
The result of such an exercise are shown in Figure 3.1, which fits scores for 27 precolonial polities to the distributions of their territory in the latter half of the eighteenth century.

**Centralized and Non-Centralized Polities: Four Regime Types**

Considering the range of political institutions in precolonial India, I suggest a fourfold classification of regime types that reflect the fiscal, military, and executive organization of the state, and I term these the tributary, feudatory, concessionary, and centralized regimes, respectively. Of these, the first two are “decentralized” insofar as they devolve considerable powers to local givers of patronage, whether tribal chieftains or warlords; whereas the latter two are “centralized” insofar as the sovereign acts as the ultimate rightsholder over land, and is the focal point for military and fiscal mobilization. In this chapter, I describe these differing political configurations, validate this typology by reference to the existing historical work on precolonial South Asia, and argue that these territorial patterns can help explain the patterns of variation in compliance and state capacity in India that will be analyzed in the next chapter.

*The Tributary State*

Considering the range of regimes that existed in precolonial India, perhaps the most skeletal political structure is what can be termed the “tributary state”. The essence of a tributary polity is that the sovereign maintains individual agreements with local stakeholders, who are tribal or village chiefs, regarding their amounts and type of financial and military tribute, while exercising no internal power over the communities. Since the tributary state relies upon voluntary adhesion, it is liable to dissipate as soon as local chiefs might sense an opportunity to suspend tribute, or reach new terms with a neighboring protector. By the late eighteenth century, the tributary model had become increasingly rare, due to the rise of new military states in the south and the conversion of forest lands into feudatory estates in the north. However, many South Asian polities maintained tributary relations on the fringes of their empires or with indigenous communities within them².

²That this was true of the Mysore Empire under Tipu in the 1790s, for example, is demonstrated by a
Figure 3.1: State Centralization Measure among Precolonial Polities. Degree of state centralization measured for 27 polities of the precolonial era, based on eight criteria: i) the degree to which taxation was universal or had widespread exemptions (inam); ii) whether precolonial assessment was conducted on individual cultivators (2), villages (1) or lords (0); iii) whether the tax rate was above 25 per cent on land; iv) whether tax collection was subject to audit; v) whether the army was largely recruited and salaried centrally or raised from local potentates; vi) whether there was a separation between head of state and head of government; vii) whether there were ministries with a clear bureaucracy; viii) whether we have records of written laws and regulations.
The only significant “pure” tributary regimes of eighteenth-century India were those of the Maratha princes in Central India, and, in particular, the Bhonsle polity with its capital at Nagpur, and the Scindias at Gwalior. Laying claim to the vast delta forest of the interuvial zone from Central India to the Orissa coast, the Bhonsle polity relied upon a strong equestrian military whose impressive string of victories over the middle to latter half of the century established a network of tributaries who pledged chaut (a form of protection tax) to the Bhonsles in Nagpur (Wink 1986). As Bayly (1988: 30-1) notes, “Maratha rulers of the Deccan and central India pushed further into the forest homelands of the Bhil and Gond tribes, converting their tribal leaders into petty rajas and their tribesmen into bondsmen or tribute-givers.” This tributary network, established in the course of a generation, did not survive the waning of Maratha power a generation later, and by the mid-nineteenth century evidence of the Maratha presence in the center-east was restricted to a few scattered palaces and forts, together with a large number of independently marauding groups formed out of the deserting cavalry officers of the defunct Bhonsle armies. During the formal founding of the colonial regime (1848-1856), all four of the princely states that reverted to British direct rule under the “doctrine of lapse” were Maratha principalities left behind by this tributary regime (cf. Iyer 2010, Appendix Table 3). Even during its brief zenith, however, the authority of this state remained nominal. Blunt (1801) reports that when he led a party of Company officers through Bhonsle territory from Chunagarh to Rajahmundry in 1798, they were stopped every few miles by chiefs demanding toll, despite the fact that they carried the Maratha pass. This revealed the true authority of the Maratha seal among their eastern tributary chiefs. Such chiefs, moreover, had little cause for loyalty: the Bhonsle realm operated in a largely predatory fashion, with “no competent civil service, no stable government and no wise foreign policy” that “kept lawlessness always raging in their jagirs and the desolate land could yield no produce for feeding the population or paying the administration” (Sarkar 1932: 112).

In southern India, by the end of the eighteenth century the tributary system had experienced a rapid breakdown as rising states, such as Mysore or Travancore, had largely

---

wall tableau in his summer palace at Seringapatam portraying in a large grid of painted images, not only audiences with emissaries and rulers of neighboring realms, but also with minor poligars discussing their terms of association.
warred down petty chieftains such as the Coorgs and Nairs and converted their villages into assignments to be taxed directly by the state. Indian court historian Mir Hussein Kirmani, for example, lists the chieftains dispossessed at one time by the expanding Mysore state, which included “the Poligars of the Balaghaut,” “the Poligar of Gungoodi Pala (Kangundi Koppam); the sons of Bhyreh Koor, the Poligar of Chuk Bal-apoor; Pud Nair, the Poligar of Vinkut Giri Kote, who was residing at Charkul; Shunk Rayel, or Rawul, the Chief of Punganoor, and besides these, the Poligars of Khut Koommir, Mudunpalli, Anikul, Oonkus Giri, [and] Cheel Naik” (cited in Rao, 1948: 887-888).

Nonetheless, in the Tamil interior, a number of minor chieftaincies (or palaiyakkarar, the origin of the anglicized term “poligar”) had endured, and continued to seek tributary bonds to outside protectors, such as the Mughal Nawab in Arcot or the British at Madras. One of the best documented examples is Pudukottai, a sole Tamil chiefdom that survived as a princely state during the British Empire. Nicholas Dirks’ (1987) *The Hollow Crown*, an ethnographic and oral history of the polity, provides a detailed picture of how such a chieftaincy was governed, with the king securing the loyalty of his 10-15 sub-chiefs through “continual gift-giving,” relieving 8,000 soldiers (amarakarars) of the duty to pay taxes through heritable fiefs, taxing only about “30 per cent” of land, and never dismissing a soldier against the preferences of his chieftains (Dirks 1987: 175, 117, 185). Dirks also reconstructs the political logic of this tributary world, in which jungle warrior tribes and forest bandits strove through service for the patronage and prestige of distant great powers (Figure 3.2), deriving “meaning from its relation to the whole (the sovereignty, and the full set of emblems, of a great king)” (Dirks 1987).

**The Feudatory State**

Beyond the tributary regime, the second form of political structure prevalent across pre-colonial South Asia was can be termed the “feudatory” state. A feudatory state refers to a polity in which the territory of the realm is divided into fiefs or tracts which are held as hereditary domains by members of an aristocracy, and over which the king has limited (either de facto or de jure) ability of reassignment. I use the term “feudatory” rather than
<table>
<thead>
<tr>
<th>Regime Type</th>
<th>Fiscal Structure</th>
<th>Coercive Structure</th>
<th>Executive Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tributary</td>
<td><strong>Fiscal structure</strong>: sovereign maintains individual, tributary agreements with local stakeholders, such as tribal chiefs.</td>
<td><strong>Coercive structure</strong>: local stakeholders (chiefs) are the heads of autonomous political units; all responsibilities (e.g. judicial, fiscal, enlistment) within such entities are performed by chiefs.</td>
<td><strong>Executive structure</strong>: king acts as absolute ruler, tributaries may maintain emissaries in the court to negotiate fiscal and military support.</td>
</tr>
<tr>
<td>Feudatory</td>
<td><strong>Fiscal structure</strong>: local stakeholders are nobles who share a kinship tie to the royal household (nobility), pass ownership by hereditary succession, and remit taxes to royal court according to rank.</td>
<td><strong>Coercive structure</strong>: local stakeholders (nobles) are responsible for judicial, fiscal, and enlistment within their estates.</td>
<td><strong>Executive structure</strong>: The king is head of a royal court consisting of a council of ministers that is recruited from within the nobility. These may have specialised portfolios (e.g. for finance, religious affairs, military, or justice), though may also represent the largest estates.</td>
</tr>
<tr>
<td>Centralized (Concessionary)</td>
<td><strong>Fiscal structure</strong>: A majority of lands are acknowledged as belonging to the royal domain, rather than to tributaries (e.g. tribes) or nobles. The sovereign offers conditional concessions (e.g. the right to farm taxes from a given land) is exchange for public service to soldiers and religious institutions, which can be canceled or reassigned.</td>
<td><strong>Coercive structure</strong>: local stakeholders are responsible to the court, which has the power to appoint, promote, and replace the holders of concessionary domains, and a standing army is financed centrally.</td>
<td><strong>Executive structure</strong>: King appoints a chief minister who heads a council with specialized portfolios (e.g. finance, religious affairs, military, and justice), with ministries that supervise local agents.</td>
</tr>
<tr>
<td>Centralized (Bureaucratic)</td>
<td><strong>Fiscal structure</strong>: A majority of lands are acknowledged as belonging to the core state. Concessions are limited in number and scope, with the sovereign engaging direct fiscal agreements with individual cultivators, administered by salaried officials.</td>
<td><strong>Coercive structure</strong>: The government appoints, salaries, promotes, and replaces local officials. A standing army is recruited centrally.</td>
<td><strong>Executive structure</strong>: The king appoints a chief minister who heads a cabinet of ministries in the capital, composed of salaried officials who supervise local agents.</td>
</tr>
</tbody>
</table>

Table 3.1: Classification of India’s Precolonial Regimes.
Feudal, due to the extensive debate concerning the appropriate definition of feudalism, and, in particular, controversies surrounding its applicability beyond the European context (Gat 2012). Under the feudatory state, responsibilities for military recruitment, taxation, and administration of justice are fully devolved to the holders of land, and members of the administration, such as ministers of state, are recruited predominantly from the hereditary nobility. By the end of the eighteenth century the most prominent feudatory states were the former Mughal provinces, such as Awadh and Bengal, in which tax-fiefs had degenerated into hereditary domains, though more longstanding examples of feudatory realms would be the Rajputs of the north-west, centered on the region that under British rule became Rajputana and today forms the state of Rajasthan.

The paradigmatic examples of feudatory regimes in precolonial India were the Rajputs: extending over an arid and landlocked landmass of some 330,000 square kilometers, these principalities were founded in the late medieval era by Kshatriya military clans, who created hereditary fiefdoms in which sons of the nobility assumed suzerainty over large territories. Such estates formed the basis of what was called the pattadari system, whereby the kingdom was fragmented into holders of estates (patti), each obliged to maintain a cer-

---

3 Azar Gat (2012) for example, insists that feudalism only applies in the case of an equestrian military elite, or knighthood, in which case the term is quite narrowly restricted.

4 “Rajput” derives from the Sanskrit term rajputra, meaning “son of a ruler.”
tain number of troops, deliver revenue to the court, and fight for the king (Rajput) in the event of war. The precolonial Rajput polities were thus weak states in which taxation was fragmentary rather than universal. Responsibilities for tax collection, enlistment, and jurisdiction were devolved to hereditary feudal lords, rather than centralized, and service in the executive was based on clan identity rather than merit.

The basis of the Rajput polity was the support of the hereditary nobility. In the northern Rajput state of Bikaner, for example, the *patta bahi*⁵, a seventeenth-century document recording the structure of taxes and identities of landholders, shows that two centuries after its founding the ruling Bikawits continued to hold almost of a third (31.5 per cent) of all land, leaving the other Rathors much of the rest (Bhadani 1979). A similar pattern of landholdings existed in other Rajput states such as Amber (founded by the Kacchawa clan) or Mewar, where the nobility (*solah*), counted just 16 to 22 individuals, under whom served 32 chiefs (*batti*), with the rest of the population as their subjects (cf. Mathur 1977). Dependence upon the nobility extended to fiscal administration as a whole, with the Rajput states implementing a limited and highly balkanized tax regime. In his study of the Rajput states, James Tod (1920) for example reported that land revenue collection was so weak that a majority of revenue could be derived from caravan custom duties, rather than taxes on land; this was true of Jaisalmer (a remote polity located in the Thar desert) and the Marwar kingdom (Sharma 1977). Within the Rajput states, land taxes varied from one area to the next, with caste and locality used to determine the level of final revenue settlement.

During the Mughal period, a number of Rajput states became Mughal tributaries and entered the *mansabdhari* system, with the king taking a grade within the Mughal court and attempting to reform his polity internally along similar lines (Zeiglar 1981). This can be seen in reforms to the executive branch: Marwar, with its capital in Jodhpur, had by the early modern period had adopted the differentiated structure of the Mughal state, with the king appointing a chief minister (*pradhan*), and finance ministers (*tan-diwan* and *desh-diwan*), and the structure of the Mewari state had begun to adopt a separation between the “inner” and “outer” court, as well as separation between the royal finances, which derived from *khalsa* (crown) domains, and the privy purse. However, attempts to supplant

⁵Its full title, the *Bikaner-re-patta-re-gaon-ri Vijat Raja Karan Singhji-re-samem ri*. 
the patrimonial system of the Rajput polity with more meritocratic norms were generally frustrated. In Bikaner, the finance minister (diwan) enjoyed a form of hereditary succession, with this position passed down the Mohata Tarawi family (Somani 1990). Attempts by Rajput kings to appoint, dismiss, and reassign fiefs were generally unsuccessful and if anything, the structure of Rajput states became more feudatory in the eighteenth century (Naravane 1999).

While the Rajputs may have constituted the most typical examples of a feudatory model in precolonial South Asia, by far the most demographically significant examples were those provinces of the Mughal Empire - Awadh, Hyderabad, and Bengal - which broke from the declining imperium after the death of Aurangzeb in 1707 to form separate, self-governing entities (Bandhapadhay 2004). While each of these initially preserved the Mughal Mansabdari system that gave the emperor the right to appoint and reassign land grants based on grades of public service, by the latter half of the century these breakaway realms were already developing a feudatory system that would presage the “permanent settlements” of the colonial era. Such a transformation of tax fiefs into “property” occurred as a result of the growing reluctance of zamindars to turn over revenues to the state, out of the assumption of a right of heritability and the ability to exchange ownership “rights” on the marketplace (Alavi 2002). In the words of Bayly, “[z]amindars (landholders) began to ... seize prebendal lands which the Mughal elites had once tried to keep out of their hands [and] families of servants of the Mughals and relatives of the old nobility bought up proprietorial rights over land or quietly converted non-hereditary into hereditary rights,” while “the proprietary rights of landholders and headmen were more regularly exchanged by sale and mortgage” (Bayly 1988: 10-11).

Awadh, located in the middle Gangetic plains in what is now Uttar Pradesh, is the clearest example of this, with power dispersed among a few large landowners and a weakening royal court. By the nineteenth century “about two dozen great magnates accounted for most of the state’s revenue and much of its military strength” (Bayly 1988), while “the state was losing its control and access to its own revenue and revenue-paying classes” (Roy 2013b: 68). These landowners were responsible for the raising of troops to fight in wars, and the most prominent among them, Almas Ali Khan, could mobilize more troops than his master in Lucknow. Faced with such weakness, the sovereign was unable to compel his estate-
holders to finance the central treasury. Neither could the court mobilize military resources, by consequence of which “Awadh was a militarily weak state [that] survived by avoiding battles” (Roy 2013: 18). It may not have helped that the royal court in Lucknow operated among a narrowing elite: Awadh’s chief minister was required to be not only Muslim, but, like the Nawab, a Shi’a (only one Sunni Muslim served as first minister) and was often appointed from the members of the Nawab’s own household.

Other Mughal successor states had managed to preserve greater vestiges of the *mansabdhari* bureaucracy, though with only varying degrees of success. In Hyderabad, for example, whereas the Mughals had tried to keep separate the roles of revenue manager (*diwan*) and governor (*subhadar*), “early on Asaf Jah and Saadat khan began again to ensure that their descendants inherited these newly amalgamated offices” (Bayly 1988: 19). Eventually, the fief-holders or *talukdars* of the Hyderabad state were to become a settled feudatory elite, as “in the later eighteenth century in some parts of Hyderabad such [land] grants did, however, tend to become hereditary. This created a more settled landowning class which negotiated its revenues and military commitments through agents settled in Hyderabad city” (Ibid). By the end of the colonial period, the Nizam and his nobility had accumulated forty per cent of all land in the state (Guruswamy 2010). Similarly, in the densely populated Bengal, as in other Mughal provinces, *zamindars* and *jaghirdars* became hereditary fiefdoms, *de facto* prerogatives that were made into *de jure* rights by the East India Company in 1793, thirty years after it had first assumed management of the province’s revenues. While Bengal’s revenues were vast in absolute terms, on a per capita basis the fiscal capacity of the state was very low. Even at the height of the Mughal Empire the state had only ever been weakly present in many parts of Bengal, in particular in the upper Ganges and the vast tribal frontier of the north.

Besides the Mughal successor states and the Rajputs, feudatory regimes also existed in the northeast of the Indian subcontinent. In particular, the Ahom Kingdom was a aristocracy based on hereditary assignments of *paiks*, or landed serfs, that for six centuries was the dominant power in the Assam Valley. Guha (1991) estimates that in 1714 one per cent of the population belonged to the aristocracy, while 90 per cent were assigned as serfs. Yet even by comparison with the Rajput states, the executive structure of the Ahom polity was narrower: from the fifteenth century onwards, the prime minister (*rajmantri*) headed
a “council of ministers” (Patra Mantris) but this was simply a group of five nobles, each of them owning large territories where they exercised military, judicial, and fiscal responsibilities. This council was expanded over time, from two members to three and finally five under the reforms of Pratap Singha in the early seventeenth century, though it remained a board of aristocrats (Bhuyan 1968). The patrimonial structure of the Ahom polity extended also to its military system, as commanders assigned to govern forward territories of the kingdom were recruited exclusively from within the narrow base of the nobility. As Ahom came under increasing military pressure from the seventeenth century, this prompted the regime to engage in limited reforms such as the introduction of its first land cadaster, the formation of a standing army, and relief of paiks with military specialization from landed duties. Yet these changes proved too limited to ensure the polity’s survival, which was destroyed by Burmese occupation in 1819-1826 (Baruah 1993).

*The Concessionary State*

A more centralized form of political structure beyond the feudatory regime is what I call the centralized-concessionary, or, simply, the “concessionary” state. Under a concessionary system, though the sovereign grants tax-fiefs to stakeholders in exchange for service, the king is the ultimate rights-holder over the land and the assigned fiefs can be canceled or reassigned, based on performance. In this way the sovereign generates the incentive for his nobility to obey. European studies of state formation have examined in length, for example, how the transition from feudalism to a concessionary model occurred under absolutist kings in early modern Europe, and similar studies have been conducted on early modern Japan (Lieberman 2009). In South Asia, the Mughal Empire had operated on the basis of a military-concessionary system by which lands were assigned to generals and public servants according to a bureaucratic grade, thus ensuring responsibility to the center, though, as we have seen, such a system did not survive among the states that succeeded the Mughal polity in the eighteenth century. Under the religious-concessionary model of the later Hindu kingdoms, concessions of brahmadeya and devadeya to Brahmins and temples were also not given simply as a means of winning support from local peoples, but increasingly involved administrative obligations to the state, including record-keeping and civil
service. In precolonial India, “breakthrough” transitions from feudatory to concessionary state forms occurred where rulers of feudatory regimes successfully institutionalized control over their stakeholders, while “degeneratory” transitions occurred when the latter by contrast re-established hereditary succession over their domains. Former Mughal provinces in the north were an example of this, while on the other hand the rising Hindu “warrior states” of the south, such as the Marathas or Travancore, were able to establish a service obligation for literate Brahmins to fulfill bureaucratic tasks in exchange for concessionary agreements.

One very clear such example of transition from a feudatory to a concessionary regime is that of Travancore, located at the very southwestern tip of the Indian subcontinent and nestled behind the mountainous western ghats. Before 1729, Travancore had been a feudatory realm governed by a council known as the Ettuveetil Pillamar (the “Lords of the Eight Houses”), which represented the eight largest estates of the realm. According to official accounts, in the 1720s, members of the council intended to implement a palace coup and establish their permanent rule. The new sovereign, Marthanda Varma, reacted by executing the eight members, seizing their estates, and selling their women and children into slavery (Menon 2008). Having displaced the Nair chiefs, Varma established absolutist control over his new realm by working with the priesthood to form a basic bureaucracy, appointing Tamil Brahmins to key state positions and granting concessions to temple lands. In 1750, Marthanda Varma formalized the alliance by dedicating the entire territory of his state to the deity Sri Padmanbha, and declaring himself as the deity’s first servant. In order to raise money for his accelerating military expenses, Travancore implemented a highly variegated system of taxes, with exemptions based on caste, religion, and public service.

Such a model of fiscal organization was relatively unique in the South Asian context, as it was not based on the taxation of land, but rather on the minute regulation of social behavior and conduct, and, in particular, the rigid enforcement of the rules of the caste system. While a low rate land tax was introduced in 1739, the majority of revenues in precolonial Travancore derived from an elaborate system of personal taxes and fines, covering not only life rituals such as marriages, childbirth and death, but also the consumption and export of commodities (salt and pepper), personal and capital goods, such as country boats, ploughs, carts, umbrellas, and adornments as headscarves or moustaches. Among
the more notorious of these was the mullakaram (literally, “breast-tax”), which was imposed on lower-caste women above the age of fourteen who chose to cover their breasts in public. The combined effect of such dues was that in 1853 per capita taxation was among the higher levels in India, even as revenue from taxation of land was among the lowest. The disciplinary regime set in place by this Hindu theocratic state generated a level of social and societal compliance that was unique in the subcontinent. So much wealth was accrued by the king within the vaults of the temple to Padmanbha in Travancore that when in 2011 an Indian court forced the board of the temple to accept a public audit of accumulated assets, these were valued at $19bn, equivalent to approximately $20,000 per man, woman and child in precolonial Travancore (Hayes 2011). The flow of revenues to the center allowed the state to build a highly centralized military force, as well as to invest large sums on the construction of fortifications, temples, and palaces. After defeating the Dutch East India Company in 1741, Varma appointed its captured Flemish captain, De Lannoy, as his own general, and gave him an ample budget to develop a European-style army based on cannons and infantry.

Another example of a concessionary regime in precolonial South Asia was the Sikh Empire established in 1799 under Ranjit Singh. Rapidly expanding to incorporate much of the Punjab, the administrative structure of the Sikh state succeeded in maintaining the centralized-concessionary system of the Mughals that, in other former provinces, had decayed over the course of the eighteenth century. In accordance with the concessionary model, about 40 per cent of the territory of Ranjit Singh’s Empire was taxed indirectly, through assignment of jagirs in reward for service (Bandyopadhyay 2008). Much of the rest however came under direct (kardar) taxation, under the administration of the Sikh court. Each village was under the revenue-collecting authority of a muqaddam, with the qanungo as the keeper of fiscal records (Singh 2006). The Sikh Empire also preserved a centralized system of policing and justice, while other parts of northern India descended

---

6 Under the colonial regime the Travancore state would also make significant expenditures on education and health, though these were not priorities during the precolonial period.

7 The difficulty of classifying the Sikh case is compounded by the fact that, following the expansion of the Sikh Empire, a number of outlying polities were brought into a tributary relationship, specifically Jammu, Kashmir, the Rajput hill states, Multan, Bhawalpur, Dera Ismail Khan, Dera Ghazi Khan, and Mankhera.
into anarchy and lawlessness. Ranjit Singh appointed chaudhries (police officers) to maintain law and order, with each chaudhry supervising over a group of villages (Oberoi 1994: 73).

A final example of concessionary regimes, albeit much smaller demographically than the other cases, are the Himalayan mountain kingdoms of Bhutan, Tibet, Sikkim and Ladakh. Though only Sikkim and Ladakh were formally incorporated into British India, all of these states exhibit important similarities, notably in the domination of animist hill and mountain tribes by a proselytizing Buddhist core state, which granted concessions to monasteries and lamas. Despite the religious difference (being Buddhist, rather than Hindu) these states presented a great deal of similarities with Travancore, namely in the existence of a mutual compact between the sovereign and the priesthood: the granting of exemptions and privileges to the latter in exchange for the conduct of administrative roles, and the resultant use of political authority to enforce a disciplinary regime of behavioral control. Following the Tibetan model, the Buddhist ministry was accorded a primary role in the organization of the state, under a doctrine known as the “dual system of government” (in Tibetan, Cho-sid-nyi). There was no possible separation between temple and state, and the king, as Chogyal (literally, “religious king”) was obliged to act as head of the clergy, just as Lamas were required to act as agents of state. It is on this basis that Sinha (1966: 330) describes precolonial Sikkim as a “Lamaist theocracy” with the Buddhist Lamas operating as ministers and civil service, recruitment and providing a religious esprit de corps. So powerful was the dominance of this monastic bureaucracy that when the king of Bhutan died in 1651, his death was kept a secret for 54 years, and civil servants continued to issue orders in his name (Worden 1991).

Because of this theocratic core, the Himalayan state exhibited attributes of bureaucratic penetration and control that were well beyond others in many parts of the subcontinent, despite having states that were, in other respects, still rudimentary. Sikkim for example

---

8Specifically, the Himalayan states remained poorly organized until quite a late point of time in matters of military organization. Namgayal and Dolma (1908) observe that precolonial Sikkim had no standing army, and warriors were collected from individual households, who were expected to bring their own weapons. The military weakness of the Tibetan mountain state is evidenced through the string of defeats they incurred under the pressure of their expansionary neighbors, with the Gorkha Kingdom (what is now Nepal) invading both Tibet and Sikkim in the eighteenth century, while Kashmir, a tributary to the Sikh Empire, conquered

95
introduced a first census in 1747, and implemented a direct head tax. The Sikkimese polity also exerted high demands upon individual households, which were obliged to conduct corvée labor for local monasteries. The fiscal structure of the Himalayan state was more akin to the ryotwari (individual-cultivator) model of southern India than to either landlord or village-based assessment. All land was owned by the state, and peasants rendered service. Edgar (1873) remarked of precolonial Sikkim that “if the systems were carried out to theoretical perfection, [the peasant] would be bound to give over to the Raja all the produce of the land.”

Centralized States

Beyond the concessionary regimes of Travancore, the Sikhs, and the Himalayan kingdoms, precolonial India also contained a number of more “fully” centralized states, in which a majority of lands were acknowledged as belonging to the sovereign and assessed directly by salaried administrators. Taxes were paid to the government, and local governance - including military recruitment, taxation, and the administration of justice - was in the hands of salaried officials of state, rather than chieftains or landlords. Centralized states in precolonial South Asia were characterized by a clear cabinet government, in which the head of state appointed a head of government (a regent or chief minister), who in turn chaired a council of ministers, each at the head of a department of state with its ministry in the capital and staffed by salaried officials who would supervise local agents.

Where were these more centralized regimes found? The clearest examples of political centralization in precolonial India were the cases of Mysore and the Maratha Peshwa. Both were the dominant power across southern India during the late eighteenth century, and the preeminent challengers to British hegemony over the subcontinent: Mysore fought four

9While, in practice, centralized states may continue to have some concessions for temples or lords, and continue to operate tributary arrangements with tribes in their territories, these are not the modal form of public administration. The persistence of patrimonial forms within a bureaucratic state, of course, is not unique to the Indian context: residual tributary and concessionary agreements were common in many European nations in the twentieth century, such as exemptions of crown lands from payment of tax, exemptions to church lands and corporations, and treaty agreements with micro-states such as those of Monaco and Andorra with France, or San Marino with Italy.
wars against the British East India Company, and the Marathas three, with both polities conducting many smaller wars against their neighbors and each other. This period of military mobilization led to increasing autocratic centralism, characterized by a strong central bureaucracy with control over taxation, policing and justice.

Mysore began as a modest tributary of the Vijayanagara Empire, whose power subsided following military defeat and plundering of its eponymous capital - at the time, the second-largest city in the world - in 1565. It remained in this subsidiary role until the advent of Chikka Deva as its king in 1673, who began the process of centralization by making enhanced revenue demands from his chiefs, and, inviting 400 of his notables to Nanjangud for discussions, killed them all without negotiation (Nagaraj 2003). Afterwards, the polity adopted a more centralized regime, with a new system of taxation and the subdivision of offices of state into 18 separate ministries in the capital, a practice that, according to Rao (1948: 485) “contributed to efficiency and centralisation of control under the personal eye of the king himself of the affairs of state.” However it was only under Haider Ali, who served as Sarvadhikari (chief regent) from 1761 to 1782, that Mysore turned from rump kingdom to regional power. Haider set about raising taxes, eliminating intermediary chieftains, and establishing a minute system of land assessment, surveillance, and enlistment, which allowed for unprecedentedly high levels of taxation and military mobilization within the Indian subcontinent.

Haider’s fiscal reforms began in 1776, when he returned to the capital following a victorious campaign against the Marathas. Tax-collectors and tributaries were invited to Mysore for the purpose of adjusting their past accounts and future revenues, and, in order to ensure better compliance from his bureaucracy, Haider appointed Anche Shamaiya, a Brahmin from the Kolar district, as his new head of intelligence. Shamaiya then “appointed harikars [informants] in every taluk [subdistrict], whose duty it was to hear and report upon all complaints in revenue matters” (Karnataka State Gazetteer, 1982: 295). The methods adopted to raise fiscal compliance by Haider and his minister Shamaiya were unambiguously despotic. The most basic sanction, korade or flogging, was dispensed so liberally

10 The tale of the Nanjangud massacre has the status of a foundational myth for Mysore. Reference to the event is made in colonial documents on the Mysore polity, and the event is assumed to have occurred though without further evidence.
that officials would “bear the marks of the stripes on their backs as public warnings.” Less fortunate offenders were subject to dabbala (piercing by pack needles), patna-chavadi talagana gidangi (solitary confinement underground), ane kalige katti (being tied to the foot of an elephant and crushed), pettu (death by continual blows) or baregala hakisi (branding) (cf. Haider Nam, 70). Accounts by English, Persian, Turkish, Portuguese, and French contemporaries attest to the ruthlessness of the Haider regime when it came to discipline of its tax collectors. The visiting Reverend Schwartz, for example, stated that he:

‘saw a number of men sitting round about; their faces betrayed a conscious terror. Hyder’s Persian Secretary told me they were collectors of districts.... To me they appeared as criminals expecting death. [One] poor man was tied up; two men came with their whips and cut him dreadfully; with sharp nails his flesh torn asunder; and then scourged afresh; his shrieks rent the air’.

The effect of these policies, however, was a marked increase in fiscal resource mobilization. In Shamaiya’s first year alone 10,000,000 varahas (34,000kg of gold) in excess were delivered to the central treasury.

At the center, Haider’s regime was an elaborately organized government, building on the reforms of Chikka Deva a century earlier. There was a clear separation between civil and military spheres and, within the former, the separation of government ministries and roles. Visiting the capital in the late eighteenth century, Schwartz observed how “on both sides [of Haider’s palace] are ranges of open buildings, where the military and civil servants have their offices, and constantly attend” and that “Haider can overlook [...] from his balcony... Here reigns no pomp, but the utmost regularity and dispatch” (from Wilks 1869, quoted in the Mysore Gazetteer, p. 2538-2539).

The compliance of officers of state was secured by a system of surveillance and punishment as strict as that applied to revenue officers. With respect to highway policing, for example, it was reported that “if, by accident, any highway robbery etc. was committed, the Kavalgar (local police guard) was impaled without delay and another promptly appointed to his place” (Rao 1948: 531). The Ottoman visitor Mirza Ikbal, in his Ahwali Hyder Naik, wrote how in Mysore “mankind were held in fear and trembling... Cutting off the nose and ears of any person in his [Haider’s] territories was the commonest thing imaginable,
and the killing of a man there was thought no more of than the treading on an ant.” Other reports by contemporary visitors to the realm frequently noted the high degree of public order and that, while “official men had cause to tremble,” “the mass of the population felt that the vigour of the Government... rendered their condition comparatively safe” (account of Colonel Wilks, quoted in the Mysore Gazetteer II: pp. 2542-2545). This was because the autocratic institutions of precolonial Mysore saw to it that “the internal governance was safe and secure, that the country’s peace and order were maintained, that disturbers of peace were not allowed any sway, that the roads were safe for traffic and trade,” and that “public order was maintained vigorously” (Rao 1948: 444-5).

Under Haider’s successor, Tipu (r. 1782-1799), bureaucratic centralism was pushed to its limit, and Susanne and Lloyd Rudolph (1979) may be correct in their suggestion that the relatively “patrimonial” structures of the East India Company, with a focus on individual leadership, responsibility, and initiative, may have proved a tactical advantage over Mysore’s increasingly ossified bureaucratization. Colonel Wilks reports a tale in which Tipu had issued a book of regulations to local officials, instructing them how to act in response to any exigency and that they should only act as the book dictates. However, “if any case should occur, not provided for, and requiring reference to the resplendent presence, such reference was to be made.” In the Kankanhalli district a husbandman came out of breath once to tell the Amil that a large field of sugarcane was on fire, to which the latter replied that the book of regulations had nothing to say on the matter, and therefore nothing was to be done. Over the villager’s protestation that the village be commanded to take buckets of water and put the fire out, he was reproached that “the case is unprovided for, and must be reported and referred.” The field was left to burn, and the case referred. Eventually Tipu made his response: to add another line to the book of regulations, concerning fires and sugar-fields (in Wilks 1869, p. 604-5).

If Mysore followed a fast-track route to autocratic centralism, a more gradual path was taken by the Marathas, the dominant power in central-northern India during the century before British rule. The Marathas ruled over a vast swathe of land, but with two distinct zones, each with a very different system of governance. The Maratha “core state,” centered on western Maharashtra, was founded in the late seventeenth century by Shivaji, and functioned as a centralized regime under the chief minister (the peshwa), with a cabinet
of eight salaried ministers, and a well-trained bureaucracy and army (Gordon 1993). As the Marathas expanded into the former Mughal territories to their north, this polity was then buffered by a series of “confederates”: princes who were ethnically Maratha (e.g. the Bhonsles, Gaikwars, Scindias) and declared their allegiance to the Maratha sovereign in Satara, but acted effectively as independent rulers. Governance in the confederate states operated according to local systems and principles of rule, ranging from the centralized system of the Gaikwars in Gujarat to the predatory, tributary regime of the Bhonsles. Yet, when the centralism of the precolonial Maratha state is mention by authors such as Gordon (1993) or Roy (2013b), this is in reference to the Maratha “core state” under the peshwa, and not to its peripheral allies.

Within this core Maratha state established in the seventeenth century by Shivaji, a set of unique political and administrative institutions were developed. Shivaji divided his new kingdom into divisions (prants): Northern, Middle and Southern. The Maratha government was based on a council of eight ministers called the Ashta Pradhan Mandal, composed of prime minister (Peshwa); minister of finance (Amatya); chief of correspondence (Sachiv); secretary (Mantri); commander-in-chief (Sari-i Naubat); foreign minister (Sumanta); chief of justice (Qaziu'l-Quzat); and ecclesiastical head (Sadaru-s-Sudur). These were salaried offices of state, with the Peshwa receiving 15,000 gold hons, the Amatya 12,000 and the other ministers 10,000 each. To give a basis of comparison, an accountant in Shivaji’s administration would be paid 125 hons per annum, making the finance minister’s salary one-hundred times higher than that of his lowest subordinate (Dighe and Qanungo 1985: 544)11.

How did the Maratha core state achieve fiscal centralization? At the outset, the Maratha realm was a feudatory state, as initially Shivaji rewarded his generals by offering them large hereditary jaghir (land grants) in western Maharashtra. However, in order to wrest central control over land administration, Shivaji and the Peshwas decided to flesh out the skeleton of a feudatory regime with the muscle of central state administration. Shivaji commissioned

---

11 Another basis of comparison is to convert into current nominal value, based on the gold content. There were 2.8 grams of gold to a hon, such that the Peshwa received a salary of 42kg of gold annually. Using the last price of the gold standard before its abolition in 1973, this would translate to a salary of $55,300 for the Peshwa, versus $461 for one of the accountants.
a new land cadastre in order to assess taxation on land. An important feature of his reform was “the abolition of the intermediate collecting agents as farmers of revenue” (Balkrishna 1940: 108). This is also reported in specific local histories, as the Satara District Gazetteer for example reports how:

“As soon as Shivaji got permanent possession of any territory, every species of military contribution was stopped, all farming of revenue ceased, and the collections were made by agents appointed by himself... Shivaji never permitted the deshmukhs [village headmen] and deshpandes to interfere in the management of the country; nor did he allow them to collect their dues until the amount had been ascertained, when an order was annually given for the amount. The patils, khots and kulkarnis were strictly superintended.” (Satara District Gazetteer, p. 242)

As a consequence of the mobilization of Brahmins into the service of the fiscal administration, Stein (1996: 191) concludes that, under the Peshwa, record-keeping “exceeded any previously known in India” and the bureaucratic system was “the most elaborate system of documentary control India had known up to that point.” The reach of the state into the village extended beyond revenue collection, for as Roy (2013b) remarks, though the “primary duty of the Patel, his assistant, the Chowgule, and the accountant, the Kulkarni, was to look after land and maintain accounts,” “frequently they were also engaged in civil administrative duties,” and “amils were drawn into carrying out tasks that concerned the well-being of the village, such as policing or dispute settlement” (Roy 2013: 30).

Just as the autocratic regime of Mysore had ensured an exceptional degree of social compliance, the Marathas enforced rigid control over their subjects. The state, for example, operated a system of labour corvée for the purpose of maintaining public works, with “certain classes of village mechanics and artisans... liable to forced labour (begar) on behalf of the state” (Edwardes 1929: 386). In addition, public order was vigorously maintained. “The best urban police force at the close of the eighteenth century was unquestionably that of the [Maratha] capital, Poona, composed of foot-police, mounted patrols, and Ramosis, used principally as spies and trackers, and was described as efficient.” It was perfected with a rigidly enforced 10pm night curfew (Tone 1818: 54-5). Order was maintained by an elaborate system of punishments, as “flogging was frequently inflicted with the object of extorting confessions of guilt; and in the case of crimes against the state torture was usually employed” (Edwardes 1929: 390). “Capital punishment and mutilation were inflicted upon
persons convicted of grievous hurt, dacoity and theft, as well as upon those found guilty of murder or treason” (Sen 1925: 417-24). These sentences were implemented by means of “hanging, decapitation, cutting to pieces with swords, or crushing the skull with a mallet” though special clemency was shown to Brahmins, who were executed with poisoning or starvation (Tone 1818: 15-6). Collective punishment was commonly used to discourage deviance; “it was a common practice of the Maratha government to incarcerate the innocent wives and children of convicts, as a warning to other potential malefactors” (Edwardes 1929: 391). Finally, the Maratha regime ruled not only with large thumbs, to quash potential sources of insurrection, but also little fingers, to keep track of opponents and spies.

As with Mysore, the Marathas had an elaborate system of domestic surveillance, backed by a postal network to speed the transfer of communications, and as Sen observes, “[s]ome writers on Maratha affairs have sought to discover the germ of modern postal communications in the system of intelligence maintained by the Maratha government” which carried “messages and letters with astonishing rapidity throughout India” (Sen 1925: 496-70).

**Overlap with Other Typologies**

The classification of precolonial Indian regimes into “tributary,” “feudatory,” “concessionary” and “centralized” polities can be compared with other efforts to classify Indian precolonial states, in particular in the works of Ramusack (2004) and Roy (2013). Examining the alternative taxonomies made by historians, however, provides a useful and necessary validating exercise (King, Keohane, and Verba 1994). A classification being a taxonomic device produced for a specific goal, discrepancies in categorization can occur due to differences in purpose, though agreement in classification is unlikely to be coincidental.\(^\text{12}\)

Barbara Ramusack’s (2004) contribution to the *New Cambridge History of India* presents a threefold classification of Indian princely states and precolonial regimes between “ancient,” “successor” and “conquest” states. Seeing the historical development of South

\(^{12}\)Notably, in the agreement in this chapter and in Roy (2013) on the importance of making a division within the Maratha realm between its core state and outlying tributaries, and the agreement of all three in making the key division of precolonial Indian regimes that between the Mughal successor provinces of central-northern India, on the one hand, and the rising southern states of the Marathas and Mysore, on the other.
Asian polities in terms of three phases of history (medieval, Mughal and colonial) her key criteria is not the degree of fiscal or bureaucratic centralization, but rather how polities relate to the Mughal Empire: whether as successors, challengers, or peripheral entities in its rise and fall. The increasingly fragmented successor provinces of the Mughal Empire like Awadh and Bengal, constitute the group that are termed “successor” states; the term “ancient” is reserved for those smaller, often feudatory polities such as the Rajputs whose existence pre-dated the consolidation of Mughal power and which throughout the Mughal era continued to operate on the margins of its imperium. Finally, the group termed “conquest” or “warrior” states refer to the new expansionary regimes founded in the seventeenth and eighteenth centuries which challenged Mughal power on the subcontinent, the Maratha, the Sikh, and the Mysore Empires.

In common with the classification developed in this chapter, the main difference for Ramusack is between the feudatory provinces of the former Mughal Empire, (Awadh and Bengal) and the more centralized regimes that arose in the course of its fragmentation. As political order in the Mughal provinces degenerated, while centralization occurred among...
its challengers, the result is a typological map fundamentally similar to that sketched in this chapter (see Figure 3.3). There are some minor differences that can however easily explained. Since Ramusack’s volume is primarily concerned with the origins of the princely states of colonial India, the inclusion of a special category for “ancient” states is needed to explain the myriad of minor principalities that were preserved within British India as a result of their alliance with the East India Company during its expansion phase13. The nomenclature of “ancient” is based on the fact that many Rajput princes could trace a lineage to the medieval era, though Ramusack’s third group, the “conquest” or “warrior” regimes, developed from polities with equal, if not longer, histories (Travancore was a renaming of the Venad kingdom, which traces its origins to the eleventh century, while the Mysore kingdom, an offshoot of Vijayanara). Also, there is disagreement about the coding of one micro-polity, Pudukottai in contemporary Tamil Nadu, which is classified by Ramusack as a “conquest” state - on account, presumably, of its militarism - but that it seemed here more pertinent to classify as a tributary polity because of its very low level of political centralization14.

Tirthankar Roy’s (2013a, 2013b) economic histories of precolonial South Asia emphasize the contrast between the policies and institutions of the British East India Company and those of the indigenous regimes of the subcontinent, rather than distinctions within such regimes. As an economic historian, Roy’s discussion concentrates precisely upon the features that are relevant to the promotion of production and trade in the subsequent centuries, such as the system of private property rights, the existence of independent courts, and the (eventual) reductions in the rate of taxation - features that clearly distinguish the British colonial regime from the indigenous polities. This focus on economic institutions, rather than political development, inevitably yields to a different interpretation of the political changes and upheavals of the eighteenth century. The extreme bureaucratic central-

13 A large number of the princely states were “micro-polities,” whose survival at this time depended on their independence from India’s indigenous empires, and the fact that they were willing to ally with the East India Company as it defeated the subcontinent’s major imperial powers. By consequence, many of these micro-states were rewarded with autonomy at the moment of the consolidation of the British imperial state in 1857.

14 The difference is of little practical consequence: the principality accounts for just one of India’s 649 modern districts.
ism of precolonial Mysore, for example, is (justifiably) seen by Roy as an impediment to trade and growth, due also to the diversion of human and economic resources towards the wasteful pursuit of warfare. Seen from the perspective of economic history, the eighteenth century in India is a catastrophic era of extractive regimes, armed conflict and, across the former Mughal domains, the dissipation and lack of maintenance of public goods, trade, and investments. Yet seen from the perspective of historical state formation, as is the focus of the present study, the eighteenth century is, by contrast, one of accelerated change: the indigenous polities of the era, or at least the Marathas, Mysore, and Travancore, advanced rapidly in bureaucratic, fiscal, and military modernization - even if the upshot of such improved political machinery facilitated the pursuit of ever more destructive conflict, with only negative immediate effects upon long-run capital accumulation.

In discussing differences within precolonial regimes, however, Roy (2013b) makes a critical distinction between different structures of land administration, and, in particular, between what he terms “statism, or attempts to control and improve the land revenue administration,” and “gentrification,” or the slide towards “revenue-farming and dependence on remnants of the former nobility, merchants, bankers and military families long settled on land but who still carried arms” (Roy 2013b: 54). These two models of governance, one involving increasing concentration of power in a central state, and the other involving political decay and a new kind of feudatory regime, largely - and perhaps not surprisingly - map the distinction that I make in this chapter between centralized and non-centralized regimes. Mysore and the core of the Maratha realm are categorized by Roy as “statist,” while I classify them as centralized states. Other regions of India, and above all the successor Mughal provinces, are generally seen as practicing some form of decentralized revenue administration, here instead termed as feudatory. The distinction between former Mughal realms and their peripheral challengers is so central to this divide, mirroring Ramusack’s own distinction between successor and conquest states, that the divisional maps resulting from these three independent taxonomies show a great deal of similarity. Roy’s distinction between statism and gentrification also, conveniently, parallels the division made here between the centralized state of the Maratha Peshwa and the tributary or feudatory regimes of his Maratha allies, with Roy arguing that the “Marathas practised statism in Maharashtra,” as the Peshwa “tried to erect a statist set-up in the core domain” (Roy 2013b: 38).
Conclusion

In this chapter, I have attempted to show that India, on the eve of colonization, was not a tabula rasa, but a richly variegated system of states and regimes with large internal variation in terms of bureaucratic centralization, police and judicial control, and fiscal capacity. This chapter has examined such patterns and configurations of political control as they existed in the subcontinent prior to the consolidation of the colonial regime in 1858, and validated a simple classificatory distinction between “centralized” and non-centralized states by comparison with other recent historical works. Following the advent of the colonial regime, many of these polities ceased to exist. Finally, in 1956, India amalgamated districts on the basis of language, at last erasing centuries of public administration based on colonial and precolonial demarcations. Yet as the next chapter will show, after six decades of governance within unified ethnolinguistic blocs, sharp internal differences can still be found within key linguistic states such as Maharashtra, Karnataka, Kerala, and Tamil Nadu with respect to the responsiveness of local government, the delivery of roads, schools and hospitals, and policing and provision of the rule of law. The persistence of differences within ethnolinguistically homogenous territories that have been subject to the same post-1956 state-level policies, I contend, illustrates the enduring historical legacy of the structures and institutions of the past.
Chapter 4

Testing for Border Discontinuities in Modern India

In the previous chapter we have seen that in the precolonial era India exhibited wide variation in the degree to which states had secured rights of taxation, enlistment authority, and political authority structures enabling their elites to exert infrastructural control over subject populations. How can we demonstrate the legacies of long-extant states and regimes for the effectiveness of governments across the world today? One of the implications of seeing state capacity through the lenses of historical state formation is that the organizational resources built by past states and regimes - institutional legitimacy, infrastructural presence, norms of compliance - should continue to be evident within countries, as well as between them. In a contemporary polity in which decisions over the extent of bureaucratic recruitment, policing, the levels and method of taxation, and enforcement were historically made under a wide diversity of competing regimes that possessed widely variant patterns of warfaring and territorial control, and, where there has been high organizational continuity of the state from these regimes into their successors, evidence of such variation ought to persist. This is especially so when historical states and regimes are seen as generating compliance equilibria that are persistence through formal institutional transitions.

Accordingly, this chapter tests such an intuition by examining variation among districts of contemporary India, where during the precolonial period centralizing and expansionist polities on the peripheries of the subcontinent broke from the dominant Mughal realm
and engaged in intensive military and fiscal modernization, directed first against the form-
er Mughal realms, then against each other, and finally against European colonial powers.
Where precolonial rulers were engaged in constant warfare, elites had an incentive to in-
vest in building state capacity by displacing local chieftains and landlords, conducting land
cadastres, abolishing tax concessions, taking on the administration of justice, and recruit-
ing armies directly. Such strategies of territorial consolidation led to the removal of local
intermediaries and the consolidation of local bureaucratic capacity, which persisted into the
colonial era due to the continuities in local-level administration under British rule.

In order to test for the effects of early political development, in this chapter I use a
border specification to identify the enduring effect of these earlier regimes by comparing
districts either side of their long-extant precolonial boundaries. I find that peoples who
today narrowly live in territories that were part of warfaring states such as Mysore, Travanc-
core, and the Sikh Empire have higher confidence in local government, higher confidence
in police, and lower rates of reported kidnappings and robberies than those narrowly on the
other side. I validate this interpretation by null-testing “placebo” borders outwards of the
actual precolonial boundaries, and also show that district-level effects are robust to controls
for state-level variation, suggesting that precolonial treatment has persisted over the effect
of post-independence state-level policies.

To identify the persistent effect of precolonial states on current institutions, this chapter
devises a border specification that compares districts located around the long-gone precolo-
nial boundaries, exploiting the geographical discontinuity created by precolonial states in
South Asia. In order not to capture unobserved heterogeneity, I control for patterns of direct
and indirect rule during the British colonial era, and membership of contemporary federal
states. Because the precolonial borders of the major precolonial polities cut through several
of India’s linguistic states, including Maharashtra, Karnataka, and Tamil Nadu, this allows
for identification of discontinuities along either side of long-extant boundaries, holding
constant state-level policies.
Mapping India’s Precolonial Political Boundaries

In order to test for border legacies, we must first establish where such borders were settled. Who ruled India - and where did they rule - in the India of the eighteenth century? For the scholar seeking to reconstruct the precolonial South Asian state, there are several potential sources. The East India Company (EIC) drew extensive boundary maps in the late eighteenth and the early nineteenth centuries as it sought to consolidate its territorial control, which inform publications such as Joppen’s (1907) *Historical Atlas of India*. Next, the district-by-district historical accounts within the *Imperial* and the *District Gazetteers*, a colonial-era publication maintained under the Raj, provide a valuable source for mapping the boundaries of historical political control across each of 349 imperial districts. Finally, there is the *Historical Atlas of South Asia*, edited by Joseph Schwartzberg (1978; republished 1992), a collaborative enterprise between seventeen authors, historians, and cartographers over a period of fourteen years, that comprises 650 maps across a wide range of themes. Several include full-page illustrations to represent the fluctuating boundaries of the precolonial and early colonial periods, including prominent Indian regimes such as the Mysore or the Sikh Empires, and the expansion of the Maratha Dominion (Figure 4.1).

Each of the sources has its respective merits and demerits. The Joppen (1907) maps reflect territorial control at different point in the expansion of the East India Company (in 1751, 1784, and 1800), and in particular the times of the four Anglo-Mysore wars. Control over land revenue being central to EIC policy in the subcontinent, this placed a high premium on the possession and usage of accurate boundary maps, so as to reflect its revenue claims and those of its rivals. However, for some polities with which the East India Company did not have direct interaction, these early maps may not be accurate except as a rough guide to territorial sovereignty.

Since the 1857 edition of the *Imperial Gazetteer* the District and Imperial Gazetters have offered district by district reports, and subsequent editions, notably the 1901 compendium, substantially expand upon earlier accounts. Entries of the Gazetteer follow a schematic format, including a historical section to provide a narrative account of each dis-

---

1Often via intermediary sources, such as Aitchison’s (1892) contribution to A Collection of Treaties, Engagements, and Sanads Relating to India and Neighbouring Countries.
Figure 4.1: 1907 Joppen (left) and 1993 Schwartzberg et al. (right) boundary maps for eighteenth-century India.

trict of imperial India in the period before colonial rule, including the identity of the pre-colonial power, its extent of territorial control, and the dates when such rule was established and displaced. Such local histories are a valuable source for mapping the boundaries of historical political control, as they cover each of the 349 districts of colonial India, including many regions where control is elsewhere poorly indicated or ambiguous, and thus allow for a truly comprehensive historical cartography. They are limited only by the fact that they refers to the 349 districts of imperial India, and not to the 640 administrative districts of the 2011 census, which involves a minor loss of resolution\(^2\).

Finally, the maps of the *Historical Atlas of India* have the benefit of being determined by multiple historians working together, able to correct each others’ mistakes, and to use the latest (as of the 1960s) historical evidence to refine their position of precolonial boundaries. They also make more of an effort to show the expansion of the Marathas, and some estimate of its division between its core state (the peshwa’s domain) and the various confederate princes. However for details on the historical evolution of polities over time, there are now more specific sources, such as Gordon’s (1993) contribution to the *Cambridge New History*.

\(^2\)However, most of the boundaries have been preserved: first, new districts have in the overwhelming majority of cases been created by splitting existing districts, rather than reassigning territory across former boundaries; second, in the majority of cases, when assigning districts the East India Company worked with the administrative units of the powers it displaced, thereby preserving their boundaries.
of India, and the maps in other volumes from the Cambridge series in particular Ramusack (2003), Stein (1990) and Richards (1993).

Where there are multiple valid data sources, it is good practice to cross-reference data so as to detect discrepancies, identify sources of measurement error, and ultimately correct them. In order to ensure accuracy, historical maps from both the Joppen (1907) and the Schwartzberg (1978) atlases were georeferenced to the current (2011) district boundary lines of India, grouped by the identity of the historical power, and then, on the basis of the 349 histories contained within entries of the Gazetteer, individual imperial districts were coded according to the precolonial power that was identified in the text, and converted to contemporary districts using subsequent district splits. Reassuringly, in the overwhelming majority of cases, the Joppen (1907), Schwartzberg (1978; 1992) and Gazetteer sources were found to converge.

The resultant map of India’s prominent precolonial boundaries is shown in Figure 4.2. The boundaries of Travancore on the southwestern tip of India were agreed upon by all three sources, having remained stable since the expansionary conquests of Marthanda Varma in the early eighteenth century and maintained as the princely state of Travancore until states reorganisation of the 1950s. The core state of the Mysore Empire covered an equal sample of districts in what is now southern Karnataka and interior Tamil Nadu during the late seventeenth century, but underwent rapid expansion under the leadership of Haider Ali in the 1760s and 1770s, which is reflected in the map; it should be noted that many of

3 In a very small remainder, one source might diverge from the other two, though this was typically due a difference of dating, or whether a tributary region was classified as an independent power. Cases resolved in this manner include portions of interior Tamil Nadu (classified by both the Gazetteer and Joppen as part of the Mysore Empire, though not by the Schwartzberg Atlas); the western territories of the Nizam of Hyderabad within interior Maharashatra (classified by the Historical Atlas as Maratha due to the victory of several battles there by the Marathas in the late eighteenth century, but by the Gazetteer as belonging to the Nizam in continuity, due to the brevity of the interlude); a northwestern portion of the Mysore border in southern Maharashatra (included by Joppen as a part of Haider’s domain, due to a northern push by Haider’s armies, but not in the Historical Atlas map or in the Gazetteer histories, as the area was vacated two years later); and the boundaries of the the Sikh Empire under Ranjit Singh (the southern boundary is agreed by all sources to lie along the Phulkian states and extend across Baltistan and Kishtwar to the north, but Joppen and Schwartzberg also include the Kashmir Valley, whereas the Gazetteer clarifies that the valley existed in a subsidiary relationship and was largely autonomous).
Figure 4.2: **Boundaries of Indian Precolonial Regimes.** Shown are the boundaries of challenger states (gold text), and Mughal successor provinces (purple text). Antique states are indicated in red text in the figure.
these territories were lost in 1791 and following Mysore defeat by the EIC in 1799. The territories of the Maratha peshwa cohere almost entirely with the “core state” of Maratha kingdom under its founder, Shivaji, in the 1680s, depicted by Gordon (1993, 1994) in his histories of the Marathas. This core state underwent rapid expansion over the following 50 years into the former Mughal territories of the north, yet peshwa rule over these additions did not last long into the eighteenth century, and peshwa governorship relapsed to the core state while territorial additions fell under the rule of Maratha princes. Finally the Sikh Empire under Ranjit Singh was based primarily in the regions of the Punjab that are now part of Pakistan, and the portions of its territory that lie within contemporary India form only its northeastern bridgehead.

Data

I test the hypothesis of an enduring precolonial legacy by using data at the district level across India, using 640 districts of the most recent census (2011). India’s administrative districts constitute a logical unit for testing the persistence of precolonial institutions into the colonial era, for, as Brass (1994) notes, “the entire system of British district administration... concentrated authority at the district level” and following decolonization “was retained virtually intact, particularly the central importance of the District Magistrate and the district courts and the police.” The districts are also a pragmatic choice of unit as they allow for both the use of district level aggregates (as reported for certain categories of public order or public goods provision), but also for the delineation of precolonial boundaries within the larger administrative units of both the colonial and post-colonial period, that is, the presidencies and the federal states, respectively. Because a large and growing literature already discusses the effect of institutional legacies at these more macro levels, either as a result of the legacies of direct and indirect rule during the colonial period (e.g. Iyer 2010, Verghese 2016), or as a result of the policies implemented at the federal state level (e.g Kohli 1988, Singh 2010), identification of differences within such larger aggregative structures would constitute strong evidence of precolonial persistence. Finally, the boundaries of precolonial polities can be traced in colonial and post-colonial district boundaries, as while the number of districts has expanded steadily over time - from 356 in 1971 to 640
by 2011 - this has occurred almost entirely though the splitting of existing districts into two or more sub-entities, rather than re-apportionment between multiple districts, which would limit the transitivity of data across units. Certainly, district boundaries have been tampered to a far lesser degree than those of the federal states, which were entirely reorganized in 1956: largely via the reassignment of existing districts to new entities rather than the creation of new districts.

The tendency of postcolonial authorities to periodically split districts into new units generates a challenge that is familiar to scholars of South Asia, namely the need to keep track of district name identities over time and reapportion data between parent and derivative units in the case of splits. To this end the district divisions of the 2011 census were matched to their larger colonial parent entities, and in the few cases where districts were created across the lines of former units, these cases were dropped if they could not clearly be territorially and demographically assigned to a historical entity. In addition, India’s “union territories,” a type of administrative division ruled directly by the union (central) government, were also included alongside districts in the dataset. The union territories are demographically comparable to the districts of federal states, and in one case (the Delhi national capital territory), are also administratively subdivided into similarly sized administrative districts. By ensuring that every unit of the country is included in the analysis, rather than a limited number of case studies, or as in some analyses (e.g. Banerjee and Iyer 2005) certain subcategories for which data is easier to obtain (e.g. districts under direct rule), I seek to use a full set of cases that is not open to accusations of sample bias, either by accident or design, as well as descriptively provide a complete as possible picture of historical legacies and contemporary government performance across the Republic of India.

Analysis at the district level also has another advantage for the purpose of econometric analysis, which is that it offers greater unit homogeneity than analysis at the level of the federal state. While subnational comparative research is often advanced as superior to cross-country sampling due to the assertion of greater unit homogeneity, in that country subunits share the same country-specific factors such as nation-level policies or recent historical events (King, Keohane and Verba 1994; Snyder 2001), an important violation of this assumption occurs subnational units differ vastly in demographic weighting, ter-
Mysore kingdom expanded by Chikkadevaraja in the late seventeenth century, who restructures government into seven departments with clear ministries. Under Haider Ali, chief regent in the late eighteenth century, Mysore becomes an ‘autocracy of the Cromwellian type’ (Rao 1946), with centralized taxation and a system of domestic surveillance. Bayly estimates taxation as high 40 per cent of output under Haider; almost six per cent of adult males see active military deployment.

Maratha Peshwa

Founded by Shivaji in late seventeenth century, who rebels against Mughal rule and proclaims himself Hindu king. Founds system of cabinet government of eight ministers under a “peshwa” (chief minister); each is salaried centrally rather than rewarded through land grants. In eighteenth century, the peshwa adopts centralized authority over the “core state” of the larger Maratha Confederacy. Within the peshwa’s domain, an advanced bureaucracy is developed through the mobilization of brahmins into civilian roles: record-keeping under the peshwas ‘exceeded any previously known in India; “the bureaucratic system of the peshwa was ’the most elaborate system of documentary control India had known up to that point” (Stein 1996: 191).

Travancore

Founded by Marthanda Varma in the early eighteenth century: executes the existing Nair elites and appoints Tamil brahmins to serve as administrators. Establishes fiscal regime on the European “company” model, i.e. state monopoly over export agriculture (spices, sugar) and use of profit to finance the state. Extensive military modernization in the mid to late eighteenth century following successful war against the Dutch East India Company.

Sikh Empire

The Sikhs rise in revolt against the Mughal Empire several times in the late seventeenth and eighteenth centuries; finally establish independent state in 1799 under Ranjit Singh. The Sikh Empire rapidly expands to incorporate much of the Punjab. The political structure of the Sikh state largely followed Mughal precedent, though involved a period of intensive military mobilization against both Mughals and Afghans.

Table 4.1: India’s Challenger Regimes.
It is never justifiable to:

claim government benefits
to which not entitled

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Other states and territories</td>
<td>0.70</td>
<td>0.74</td>
<td>0.54</td>
<td>0.61</td>
</tr>
<tr>
<td>Precolonial legacy states</td>
<td>0.82</td>
<td>0.78</td>
<td>0.72</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Table 4.2: State Capacity and Compliance Among Precolonial Legacy States. Federal states with precolonial heritage are those deriving from a single dominant precolonial polity during 1757-1856: Karnataka (Mysore), Kerala (Travancore), Sikkim, Punjab (Sikh Empire), Maharashtra (Maratha Empire), Tripura, Andhra Pradesh (Hyderabad). Contemporary federal states without a single dominant precolonial polity are Bihar, Haryana, Tamil Nadu, West Bengal, Delhi, Uttar Pradesh, Himachal Pradesh, Madhya Pradesh, Chhattisgarh, Mizoram, Nagaland, and Arunachal Pradesh.

...
tricts of the bottom decile, to an average of 4,593,092 among the top decile by population of districts. The data are therefore more normally distributed and non-skewed than India’s federal states and union territories, as per the 2011 census, 85 per cent of India’s administrative districts contained between 500,000 and 5,000,000 residents, and only 18 districts (3 per cent of the sample) were more than two standard deviations from the mean. Second, as the districts are demographically significant in their own right, multi-district regression within India also has some of the generalizable attributes of multi-country analysis, perhaps among the most significant benefits of this form of analysis (Inglehart and Welzel 2008). One-third of the 247 countries and dependencies included in the comparative data of the United Nations Statistical Office have a population fewer than 1m inhabitants, leaving 159 countries above this cutoff. As of 2011, 405 of India’s 640 districts contained a population of more than 1 million: each district a Botswana, Latvia, or Lebanon. Traveling within a single federal state such as Kerala, Karnataka, or Maharashtra, differences between districts are often visible and vast, with differing ethnic and religious groups, as well as vast differences in economic and human development, and subnational analysis offers an entirely new universe of cases for testing the hypotheses that have been developed based upon cross-country data. Third, the greatly increased number of units means that any source of unobserved variation will be dissipated across multiple discrete instances, rather than concentrated in a single outlier. Finally, there are 640 districts within contemporary India (as of 2011), whereas the number of states and union territories is relatively low, with 36 discrete entities today (and only 20 in 1956), such that the estimated standard errors would be likely large, resultant coefficients would be subject to a high degree of sampling sensitivity, and only weakly robust.

**State Effectiveness Across Districts of Contemporary India**

For data on variation in state effectiveness across districts of India, I take data from several sources. First, public opinion data on citizen perceptions of institutional performance are taken from the India Human Development Survey, which covers 41,554 households in 1,503 villages and 971 urban neighborhoods across India (Desai et al., 2006). From this survey, three indicators are included: the level of confidence in local government (**pan-**
chayat) “to implement public projects”; the level of confidence in police; and reported crime victimization, in the form of the percentage of households that say they have been the victims of a theft. Second, I take a series of indicators of legal compliance and rule of law, including the number of cases of ‘cheating’ that were reported to the police, and reported numbers of kidnappings, robberies and rapes, at the district level across India. Finally, I validate these observations by including several broader indicators of public goods delivery at the village level, including the literacy rate, the male-female literacy gap, and the density of medical dispensaries, as reported in the 1981 and 1991 censuses (Iyer 2010).

**Full Sample District Regressions**

*Set 1: Basic OLS*

The first set of tests estimate an effect of precolonial heritage by means of a simple dummy variable for whether a district was a part of one of India’s eighteenth-century challenger regimes. This basic model expresses the state effectiveness variables $y$ of district $i$ in state $s$ as a function of whether the district was historically administered by one of the challenger states, $C$, and a set of district level control variables, $X$:

$$Y_{is} = \alpha + \beta_1 C_{is} + X'_{is} \gamma + \epsilon_{is}$$  \hspace{1cm} (4.1)

Included in $X$ are the following controls. First, the level of legal compliance may be positively affected by a higher level of regional income. Accordingly, the models comprise a variable for district level average household income. Second, a significant body of literature has asserted that group heterogeneity is negative for rule of law, and in particular, in the South Asian context, the percentage of members of scheduled tribes and castes (Jauregui 2013; Alesina et al. 2003; Alesina, Baqir, and Easterly, 1999; Miguel and Gugerty, 2005). Separate variables are included for the percentage of members of each district belonging to either of these two groups. Third, a longstanding literature has argued that social institu-
tions matter for local governance, in particular participation of citizens in local governance through civic activities such as protest, petition, attending neighborhood associations, writing to local newspapers and officials (Putnam 1993; Norris 1999; Tsai 2007). A variable is therefore included which registers the average membership, by district, of each of the following civic organizations: a mahila mandal (women’s organization); youth or sports clubs; labor unions or business associations; self help groups; rotating credit and savings associations; religious organizations; caste-based support groups; development NGOs; and agricultural cooperatives. Finally, the policies of India’s state-level administrations may be an important determinant of variation in institutional quality. Dummy variables are included for Kerala and Karnataka, the two federal states that are most frequently cited as examples of high quality governance in the South Asian context (Kohli 2009). This is done in order to control for the effects of policies introduced at the state level since the establishment of India’s current system of federal states in 1956.

Set 2: Border Discontinuity Design

The second set of regressions restricts the sample space to only districts adjacent to the eighteenth century borders of India’s precolonial states. In the spirit of a regression discontinuity design, the objective is to identify the effect of defensive modernization among India’s eighteenth-century challenger regimes by comparing households and villages that lie on either side of the former borders. Proximity within a common region is likely to ensure similarity in unobserved variables such as language, religion, and geography. The minimum possible bandwidth is chosen within the framework of a district-level dataset, which is to opt for districts adjacent to the boundary lines.

The models to be estimated now take the form:

\[ Y_{is} = \alpha + \beta_1 C_{is} + X'_{is}\gamma + \epsilon_{is}, \in \text{borderregion} \]  

(4.2)
Where $Y_{is}$ refers to one of $n$ metrics of state capacity in district $i$, $\alpha$ is the intercept, $C_{is}$ is a dummy variable for whether a district was governed in the past by one of India’s eighteenth-century challenger regimes, and $X'_{is}$ refers to a matrix of control vectors, whose independent association with $Y$ is estimated in the vector of betas $\gamma$. Districts selected are those within the subset of observations satisfying the border qualification, i.e. are adjacent the precolonial boundary lines.

This empirical identification overcomes a wide variety of potential observed and unobserved confounding variables. Because these regions have generally been governed under the same state government since current state boundaries were settled in 1956, post-independence policies and institutions at the state level cannot explain variation across most boundaries; nonetheless, control variables are included for the states of Kerala and Karnataka, the most frequently cited as examples of high quality state-level governance (Kohli 2009, Heller 2000). In addition, such an identification excludes entire regions from the analysis which might be seen as unfairly leveraging the results: in particular, almost all but a handful of districts from the BIMARU states (Bihar, Madhya Pradesh, Rajasthan
and Uttar Pradesh), known for their relatively higher levels of deprivation and problems of governance, are now excluded from the sample. Such an identification also controls for more subtle forms of variation that might confound empirical identification. In particular, comparison districts either side of former precolonial borders are in most cases of the same ethnic and linguistic community: the former southern Mysore border runs through the middle of the Tamil districts of contemporary Tamil Nadu, for example, while the northern boundary of the Maratha peshwa’s domain cuts across the north of the Marathi areas of current-day Maharashtra.

 Nonetheless, we may be concerned about alternative channels by which precolonial boundaries correlate with contemporary district-level outcomes. In particular, it might be that the channel by which precolonial states affected subsequent outcomes was not due to their own investments in state capacity, but rather, because precolonial boundaries patterned the boundaries of direct and indirect rule during the British Empire, and it was this which affected subsequent paths of development (Mukherjee 2013, Iyer 2010). Though, if true, this would still show an enduring effect of precolonial state-building, it would also place greater theoretical and substantive interest upon colonial policies, rather than the earlier historical processes identified in this article. In order to exclude this alternative causal pathway, the regressions in Set 2 therefore include a control measure for patterns of direct and indirect rule during the British Empire (from Iyer, 2010). Furthermore, by including control variables for the federal states of Kerala and Karnataka, this also controls for differences in cultural/language group that may predate the founding on India’s contemporary linguistic states (specifically, Malayali and Kannadiga identity), and which might serve as an alternative causal mechanism. Results are shown in Table 4.3.

**Results**

The coefficients in Table 4.9 suggest that presence within an eighteenth-century challenger regime is significantly associated with improved state effectiveness in the present day. Individuals living within such districts have higher levels of confidence in local government “to implement public projects” and in local police, are less likely to report having been a victim of theft in response to a crime victimization survey item, have lower numbers
**Independent variable:**

<table>
<thead>
<tr>
<th></th>
<th>conf: local govt</th>
<th>conf: police</th>
<th>cheating</th>
<th>victim: theft</th>
<th>robbery</th>
<th>kipnap</th>
<th>rape</th>
<th>literacy</th>
<th>gengap</th>
<th>medical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenger State (=1)</td>
<td>0.25**</td>
<td>-0.12</td>
<td>-2.06</td>
<td>-0.07**</td>
<td>-12.2*</td>
<td>-3.83</td>
<td>-4.79</td>
<td>10.71***</td>
<td>-0.02</td>
<td>0.05*</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(27.74)</td>
<td>(0.02)</td>
<td>(5.33)</td>
<td>(9.36)</td>
<td>(2.44)</td>
<td>(2.58)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Direct British Rule (0/1)</td>
<td>-0.12</td>
<td>-0.1</td>
<td>1.59</td>
<td>-0.02</td>
<td>6.87</td>
<td>-0.67</td>
<td>3.41</td>
<td>3.77</td>
<td>0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.12)</td>
<td>(31.62)</td>
<td>(0.03)</td>
<td>(6.08)</td>
<td>(10.67)</td>
<td>(2.79)</td>
<td>(2.95)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>State: Kerala</td>
<td>-0.44**</td>
<td>0.29</td>
<td>24.18</td>
<td>0.06</td>
<td>-1.84</td>
<td>-13.25</td>
<td>6.35</td>
<td>32.2***</td>
<td>-0.13***</td>
<td>0.29***</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.21)</td>
<td>(58.87)</td>
<td>(0.05)</td>
<td>(11.32)</td>
<td>(19.86)</td>
<td>(5.19)</td>
<td>(5.51)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>State: Karnataka</td>
<td>-0.5***</td>
<td>0.26</td>
<td>81.91*</td>
<td>0.14***</td>
<td>12.69</td>
<td>-1.5</td>
<td>-0.36</td>
<td>-11.33**</td>
<td>0.04*</td>
<td>-0.12***</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(38.13)</td>
<td>(0.03)</td>
<td>(7.33)</td>
<td>(12.86)</td>
<td>(3.36)</td>
<td>(3.57)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.7***</td>
<td>-1.93***</td>
<td>27.53***</td>
<td>36.58***</td>
<td>11.94***</td>
<td>47.09***</td>
<td>0.19***</td>
<td>0.11***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(29.53)</td>
<td>(0.02)</td>
<td>(5.68)</td>
<td>(9.96)</td>
<td>(2.6)</td>
<td>(2.76)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>Adj R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>0.31</td>
</tr>
<tr>
<td>48</td>
<td>0.07</td>
</tr>
<tr>
<td>55</td>
<td>0.03</td>
</tr>
<tr>
<td>48</td>
<td>0.33</td>
</tr>
<tr>
<td>55</td>
<td>0.09</td>
</tr>
<tr>
<td>55</td>
<td>-0.06</td>
</tr>
<tr>
<td>56</td>
<td>0.04</td>
</tr>
<tr>
<td>56</td>
<td>0.63</td>
</tr>
<tr>
<td>56</td>
<td>0.43</td>
</tr>
<tr>
<td>56</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Table 4.3: Discontinuity Test, Precolonial Boundary Districts Only. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level. Includes only districts bordering the precolonial challenger state boundary lines.
of officially reported crimes across a wide variety of areas, from allegations of cheating and 
 fraud to kidnappings and rape, and better human development outcomes, including a higher 
density of medical dispensaries, higher literacy rates, and a lower gap between male and 
female literacy outcomes. These effects are robust to the inclusion of controls for contem-
porary federal state, the pattern of rule during the British Empire, income, social capital, 
the proportion of lower castes, the proportion of tribal peoples, and a measure of land in-
equality. They are, further, substantively large: membership of a centralizing precolonial 
regime affords a reduction in the rate of kidnappings by 40.6 per 100,000, almost the same 
as the all-India district mean (49.5 per 100,000) and the rate of robberies by 25.9 (against a 
district mean of 59.7); it also is associated with a 11.4 per cent increase in the literacy rate.

Do these effects remain when we consider only the reduced-sample discontinuity ef-
fects reported in Table 4.3? With the sample space reduced to only those districts on the 
precolonial border zone, the estimated effect sizes are substantially reduced, yet, remain 
significant across most areas. Even controlling for legacies of British direct rule and con-
temporary federal states, individuals living within districts situated narrowly outside of the 
borders of India’s centralizing precolonial states, namely Mysore, Travancore, the domain 
of the Maratha peshwa, and the Sikh Empire, have higher levels of confidence in local 
government “to implement public projects,” are less likely to report having been a victim 
of theft in response to a crime victimization survey item, have lower numbers of reported 
robberies, and villages within such districts have a higher density of medical dispensaries 
as well as higher literacy rates. Within this sample the estimated reduction in the rate of 
robberies is 12.2 per 100,000 (to an average of 30.0), there is a 10.7 percentage point in-
crease in the literacy rate, and a 7 per cent estimated reduction in the rate of households 
reporting having experienced a theft (to a subsample average of 1.56 per cent).

Sensitivity and Robustness Tests

A concern we may have with the results reported in Table 4.9, is that one individual 
polity may leverage the reported results. This section tests whether the results are robust 
to the casewise deletion of each of the four challenger polities (Mysore, Travancore, the
Maratha peshwa and the Sikh Empire), so as to ensure that the results are not simply due to leveraging effect of districts within one individual polity, but rather reflective of a pattern of greater institutional effectiveness across all of the former challenger regimes. Accordingly the “Challenger State” coefficient from the boundary regressions in Table 4.9 are re-estimated, using four additional samples, each excluding one of the precolonial challenger regimes. These coefficients are reported in Table 4.4.

Almost all the coefficients remain in the expected direction, such that districts within precolonial challenger regimes exhibit greater state effectiveness than those outside. In some instances, the effect size is considerably reduced to the point of losing statistical significance, however, there is no consistent result such that overall institutional effectiveness can be attributed to any one polity. For example, the coefficient for the reduction in the rate of kidnapping is only non-significantly negative after excluding the territories of the Sikh Empire, yet many other estimated effects (e.g. on confidence in local government and police, surveyed crime victimization, or reported robberies and rapes) remain robustly significant in this circumstance. Similarly, on the exclusion of districts within the Mysore Empire, the coefficient for the reduction in crime victimization loses significance; however, in this case, the coefficients for differences in cheating, robberies, kidnappings, rapes, and density of medical dispensaries are all significant and in the expected direction. While there is evidence that particular precolonial polities may have left a more enduring impact across performance in particular policy domains (e.g. control of crime, literacy, or access to healthcare) there is no evidence that the results of Table 4.9 are due to one individual polity rather than reflecting a general legacy of eighteenth-century challenger regimes. Furthermore, in all cases the coefficients for the literacy rate, density of medical dispensaries, and the rate of robberies, kidnappings, and rapes are robustly significant after casewise deletion.

**Placebo Test Using Pseudo-Borders**

A possible limitation of the reported results, even in the border discontinuity sample, is that these only capture differences between northern and southern India, and not a specific legacy of precolonial regimes. Were this the case, however, we should find an effect also when spuriously moving the border outwards. I implement this falsification test by altering
### Table 4.4: Sensitivity Tests for Challenger State Coefficient, with Casewise Deletion. Sample: All Districts.

<table>
<thead>
<tr>
<th>Independent variable:</th>
<th>conf: govt</th>
<th>conf: police</th>
<th>cheating</th>
<th>victim:theft</th>
<th>kidnap</th>
<th>robbery</th>
<th>rape</th>
<th>literacy</th>
<th>gengap</th>
<th>medical</th>
</tr>
</thead>
<tbody>
<tr>
<td>All districts</td>
<td>0.19**</td>
<td>0.12*</td>
<td>-27.46*</td>
<td>-0.04***</td>
<td>-25.97***</td>
<td>-40.61***</td>
<td>-16.52***</td>
<td>11.4***</td>
<td>-0.03***</td>
<td>0.05***</td>
</tr>
<tr>
<td>Excluding Maratha Peshwa</td>
<td>0.21**</td>
<td>0.14*</td>
<td>-34.56**</td>
<td>-0.05***</td>
<td>-26.2***</td>
<td>-45.3***</td>
<td>-16.34***</td>
<td>10.32***</td>
<td>-0.04***</td>
<td>0.04***</td>
</tr>
<tr>
<td>Excluding Mysore</td>
<td>0.11</td>
<td>-0.04</td>
<td>-33.76*</td>
<td>0.000</td>
<td>-25.23**</td>
<td>-30.89*</td>
<td>-13.23**</td>
<td>12.7***</td>
<td>-0.03**</td>
<td>0.05***</td>
</tr>
<tr>
<td>Excluding Sikh Empire</td>
<td>0.13</td>
<td>0.23***</td>
<td>-2.5-0.06***</td>
<td>-14.26-26.86*</td>
<td>-12.33**</td>
<td>8.99***</td>
<td>-0.02</td>
<td>0.04***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excluding Travancore</td>
<td>0.2***</td>
<td>0.12*</td>
<td>-24.36</td>
<td>-0.05***</td>
<td>-25.76***</td>
<td>-40.24***</td>
<td>-16.63***</td>
<td>9.02***</td>
<td>-0.02*</td>
<td>0.04***</td>
</tr>
</tbody>
</table>
the treatment to a pseudo border that is located an additional district outwards from the actual outer limit of the precolonial states used in Table 4.3. These results are reported in Table 4.5. All of the effects either side of the pseudo border are statistically insignificant, while the share of the variance that can be explained, as indicated by the adjusted r-square, is zero or close to zero (or even negative) in all cases. This supports the hypothesis that there is a genuine discontinuity along the borders of precolonial challenger regimes, and not simply the continuation of a gradual incline from the “Hindi heartland” of northern India outwards to the peripheral regions of the subcontinent.

**Data on Historical State Capacity during the Colonial Era**

In general, areas of India that were governed by more centralized states have better public service delivery, rule of law, and local governance, and these differences persist even when making matched comparisons within contemporary Indian federal states. Yet we cannot know whether the differences observed today between regions of federal states are a consequence of historical divergence in the past, or a more recent public policy outcome. For example, politicians from regions with stronger precolonial states may have been more successful in capturing public resources and distributing these to their supporters in the form of jobs, investment, and subsidies (Weiner 1967, Wilkinson 2007). Alternatively, the takeover of less developed regions by historically more centralized provinces may have “underdeveloped” the former, by removing political responsibility and authority from their areas, as has been alleged in comparable cases such as the Italian mezzogiorno (Aprile 2012).

Accordingly, this section assesses within-state differences at a point before the creation of India’s federal states, in order to rule out post-treatment bias in the assessment of regional disparities within these units. If differences in public goods and rule of law outcomes are a persistent legacy of precolonial states, then we should expect to observe these differences at any point between the abolition of India’s precolonial regimes in the first half of the nineteenth century and the present day, and not simply at one time interval. To do this I reconstruct statistics of public investment for the whole of India in 1900-1, taking data from the *District* and the *Imperial Gazetteers of India*, accessed in hard copy.
### Independent variable:

<table>
<thead>
<tr>
<th></th>
<th>conf: local govt</th>
<th>conf: police</th>
<th>cheating</th>
<th>victim: theft</th>
<th>robbery</th>
<th>kipnap</th>
<th>rape</th>
<th>literacy</th>
<th>gengap</th>
<th>medical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo Boundary (0/1)</td>
<td>0.07</td>
<td>-0.12</td>
<td>-12.03</td>
<td>0.01</td>
<td>-5.42</td>
<td>-0.84</td>
<td>0.35</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.1)</td>
<td>(18.41)</td>
<td>(0)</td>
<td>(5.98)</td>
<td>(9.76)</td>
<td>(4.88)</td>
<td>(2.45)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Direct British Rule (0/1)</td>
<td>-0.04</td>
<td>0.08</td>
<td>3.62</td>
<td>0</td>
<td>8.05</td>
<td>-15.25</td>
<td>0.42</td>
<td>14.18***</td>
<td>0</td>
<td>-0.05*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.1)</td>
<td>(18.5)</td>
<td>(0)</td>
<td>(6.01)</td>
<td>(9.8)</td>
<td>(4.91)</td>
<td>(2.5)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>State: Kerala</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>State: Karnataka</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.82***</td>
<td>-1.93***</td>
<td>60.75***</td>
<td>35.26***</td>
<td>49.09***</td>
<td>23.29***</td>
<td>39.12***</td>
<td>0.21***</td>
<td>0.1***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(14.13)</td>
<td>(0)</td>
<td>(4.59)</td>
<td>(7.49)</td>
<td>(3.75)</td>
<td>(1.98)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

| N                     | 43               | 43           | 54        | 43            | 54      | 54     | 54   | 54       | 54     | 54      |
| Adj. $R^2$            | -0.03            | 0            | -0.03     | 0.04          | 0       | 0.01   | 0.02 | 0.4      | -0.04  | 0.05    |

Table 4.5: Discontinuity Test - Placebo Test using Artificial Boundary, One District Removed from Actual Boundary. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level. Includes only districts bordering the precolonial challenger state boundary lines. Dummy variables for Kerala and Karnataka due to absence of cases in the sample space.
at the India Archives in New Delhi and London, and hand-copied in late 2014. Collected at the height of the British Empire, data for each of 349 imperial districts were published under individual entries for population, spending on education, spending on healthcare, the length of paved and unpaved roads, numbers of schools, medical dispensaries, literacy, and the numbers of police. If the differences between Indian districts today are a legacy of divergences in the precolonial period, then the within-state disparities shown in Table 4.12 should also be visible in governance and service provision indicators from the turn of the twentieth century.

A summary of such indicators is shown in Table 4.6. Not only are the precolonial divides visible in 1900, but if anything, they are even more apparent at this point than they are today. Each of the areas that had been under the rule of the expansionary, centralizing states of the eighteenth century had higher numbers of schools per capita, literacy, and medical dispensaries per capita, and most have greater education and health spending, road provision, and policing in 1900. The differences are striking, as in a number of cases the districts on either side of the extant precolonial border were at this point ruled directly by the British Empire, as in Punjab or northern Karnataka. These differences were substantively very large; former districts of the Mysore Empire had in 1900 three times the per capita spending on healthcare and over double the per capita spending on health than those beyond its northern limit, Tranvancore’s per capita spending on health was more than double that of Cochin and Malabar, and the former domains of the Maratha peshwa spent almost twice as much on education and health, in 1900, than the Marathi-speaking areas that were then under indirect rule.

**Bureaucratic Responsiveness Tests**

The evidence presented in Tables 4.9 to 4.4 suggests strong support for the hypothesis that precolonial states have left an enduring legacy upon patterns of compliance and rule of law in contemporary India. Nonetheless, we may have concerns regarding the validity of public goods indicators as indicators of state capacity, as well as the reliability of survey items of institutional confidence as a reflection of the objective performance of these institutions. This section provides one additional source of evidence by presenting the results of
<table>
<thead>
<tr>
<th></th>
<th>Karnataka</th>
<th>Kerala</th>
<th>Maharashtra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mysore Other Travancore Other Peshwa Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Educ. spend. / cap.</strong></td>
<td>0.19 0.08 0.22 0.17 0.19 0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health spend. / cap.</strong></td>
<td>0.06 0.02 0.09 0.04 0.10 0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Roads / resident</strong></td>
<td>0.33 0.19 0.34 0.27 0.3 0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>School / capita</strong></td>
<td>0.64 0.22 1.26 0.89 0.42 0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Medical disp / cap.</strong></td>
<td>0.02 0.01 0.04 0.01 0.02 0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Police / cap.</strong></td>
<td>1.05 1.02 0.59 0.51 1.17 1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lit. rate</strong></td>
<td>0.06 0.03 0.12 0.10 0.05 0.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Tamil Nadu Travancore Other Sikh Empire Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Educ. spend. / cap.</strong></td>
<td>0.22 0.28 0.11 0.08</td>
</tr>
<tr>
<td><strong>Health spend. / cap.</strong></td>
<td>0.09 0.08 0.03 0.04</td>
</tr>
<tr>
<td><strong>Roads / resident</strong></td>
<td>0.34 0.39 0.1 0.12</td>
</tr>
<tr>
<td><strong>School / capita</strong></td>
<td>1.26 0.71 0.21 0.11</td>
</tr>
<tr>
<td><strong>Medical disp / cap.</strong></td>
<td>0.04 0.01 0.01 0.01</td>
</tr>
<tr>
<td><strong>Police / cap.</strong></td>
<td>0.59 0.65 0.65 0.98</td>
</tr>
<tr>
<td><strong>Lit. rate</strong></td>
<td>0.12 0.08 0.04 0.03</td>
</tr>
</tbody>
</table>

Table 4.6: Within-State Comparisons: State Capacity Data from 1900-1. All data from the *District Gazetteer.*
a bureaucratic responsiveness test conducted within the state of Karnataka in 2013-2014. Bureaucratic responsiveness is a direct indicator of state effectiveness, as it measures the willingness of local bureaucrats to comply with their legal obligations; and has been used in many previous studies of local state capacity (e.g. Putnam 1993). India is an especially appropriate context for such a design, as the Right to Information (RTI) Act 2005, directs all public officials to respond in writing any appropriately framed request for information. The failure to do so within 30 days incurs a possible legal penalty. A letter was sent to the Karnataka State Government in October 2013 requesting information about the number of officials charged under a recently introduced law against caste discrimination. This order was passed to police stations in size I-III cities in Karnataka, which subsequently responded directly. While the state of Karnataka is known within India for its generally high degree of state capacity, examining internal variation within the state allows us to examine a natural experiment in legacies of precolonial state formation. This is because following the defeat of the Mysore Empire by the East India Company in 1799, a rump Mysore Kingdom continued under the British Empire as an indirectly governed princely state, in 1947, it was made into a state of the new Federal Indian Republic. In 1956, this state acquired new territory under the process of States Reorganisation, which organized states along linguistic lines, and was rebaptized from Mysore to “New” Mysore. During this reorganisation, Mysore regained all of the Kannada-speaking territories that it had lost in 1792 and 1799 to the East India Company. It also regained several Kannada speaking districts in the north that, had never been a part of the erstwhile Mysore Empire, were predominantly Kannada-speaking. This variation between districts that were formerly part of the Mysore Empire and districts that were never part of Mysore provides a good context for testing the natural experiment provided by precolonial treatment. These districts in both northern and southern parts of Karnataka share the same language, the same ethnic and religious composition, and have been governed in common as members of a single federal state since 1956. We therefore immediately control for many sources of observable and unobservable heterogeneity between regions that might provide an alternative explanation for regionwide differences in state capacity. Across size I-III cities in Karnataka, 25 per cent responded with the requested information within 2 weeks to 3 months of receiving the request. This distribution of bureaucratic responsiveness is shown in Figure 5.
Figure 4.4: Cities surveyed by the RTI test within Karnataka. Positive responses in red; non-responses in white. Northernmost border of the Mysore Empire indicated by green line.

One can note that only two districts north of the 1782 Mysore boundary line responded to the requested information. This discontinuity is especially stark given that many cities immediately south of the border provided their responses unlike those immediately to the north. We can test for the evidence of such a discontinuity formally by using a regression discontinuity design, with the treatment effect as experience of Mysore rule in the period of the late eighteenth century. As incorporation into the Mysore Empire was largely a function of the steady outward expansion of Mysore acquisition in the period from the late 1600s to their defeat in the third Anglo-Mysore war of 1791, it is effectively randomly assigned, and almost all variation in treatment is explained by distance from the capital, Mysore. I first test using a full bandwidth that incorporates all cities of Karnataka north and south of the border cutoff. After that I perform a second regression with a narrower bandwidth, taking only those cities within 70 miles of the precolonial boundary line. There are 158 observations in the unrestricted sample and 56 observations in the bandwidth sample. I also include tests using a placebo boundary located 100 miles to the south of the actual
northern Mysore Empire border as a further check upon the robustness of the findings.

Local average treatment effects are estimated around the threshold:

$$\alpha = E[Y_i(1) - Y_i0 \mid X_i = c]$$

Where $c$ is the boundary cutoff between cities that were under or not under Mysore rule, illustrated in the figure showing the map of Karnataka.

For the dependent variable, I take the number of days taken for municipalities to respond to the RTI request from the date of initial deposition. Because three-quarters of municipalities did not provide a response, these are right-censored at 62 days (two months). This is also the formal deadline after which a person can initiate legal proceedings against an agency for failing to provide a reply. Therefore, this is the point at which a request for information can be assumed to have failed.

Results are shown in Table 4.7. Using the full sample of cities, separate regressions estimated either side of the border show a discontinuity in the intercepts at the border, at 64.5 days on the non-Mysore side, versus 60.4 days on the Mysore side. The slopes are also different. The undivided sample (Model 1) shows a downward slope overall, with southern
cities (within the former Mysore Empire) having better responsiveness, separate regressions show an inversion of the slope coefficient at the border. These differences are illustrated in Figure 4.5 below. The results are easier to interpret when one takes the reduced sample within a 100 mile bandwidth of the former Mysore northern border. The intercept is 55.4 days until response south of the border. There is also a slight estimated downward slope such that more distant cities are more bureaucratically responsive, though this coefficient is not significant. As there were no responses at all within a 100 mile radius north of the border, the estimated slope is flat and lies along the right-censored value of 62. There is therefore a large discontinuity in the estimated value at the border (55.5 versus 62) and a change in the slopes. Finally, the placebo test, placed on a 120 mile bandwidth 70 miles south of the Mysore border, shows no substantial discontinuity in estimated values. The discontinuity at the cutoff is only 59.4 against 57.2, and the changes in slope coefficients are also not significant.

Examining Extant Borders Individually

This section explores these comparisons in greater detail, by examining individual pre-colonial boundaries in the data from within Karnataka, Kerala, and Maharashtra. By looking within each federal state to show how regions under centralized precolonial rule differ from those that did not undergo such rule, it is possible to show the persistent effect of historical states can be found in each instance. Within-state comparisons may be more revealing, considering that all parts of India were grouped into ethnolinguistic units following 1956, regardless of whether they demanded it. The districts that were so amalgamated comprised the same dominant ethnolinguistic group, lived in similar geographic conditions and were subject to following reorganization to the same statewide policies. They differed systematically in only one fundamental regard, which is their legacy of public administration from the past. Looking internally to these new entities, and the persistent discrepancies in government effectiveness among their component districts, it is possible to make inferences about institutional trajectories and the long shadow of the past upon the present.

Karnataka: The Mysore Border Separating North and South
Figure 4.5: Illustration of the Regression Discontinuity. Response times of cities in Karnataka to RTI requests, with 1782 Mysore border as the discontinuity cutoff. Slopes on either side slide down towards the border, with a distinct discontinuity.
Born as “New Mysore” on November 1, 1956, and renamed “Karnataka” in 1973, the state of Karnataka promised a revival of the “greater” Mysore that existed in southern India of the late eighteenth century. After a hiatus of two hundred years, the territories of the rump Mysore princely state were reunited with the Kannada-speaking districts that were ceded to the East India Company following the third and fourth Anglo-Mysore wars. However, as the basis of unification was linguistic rather than historical, this “new” Mysore also found itself in charge of a set of Kannada-speaking northern districts that had historically never been a part of the Mysore polity. These areas instead had been governed by declining Deccan sultanates and local potentates in the eighteenth century.

A priori, there was no reason to expect that these new districts would have any difficulty integrating into the rest of the state, as they were of the same Kannadiga ethnolinguistic origin. Moreover, there were also similarities in terms of religious composition (87.5 versus 91 per cent Hindu), caste structure (both sides with 3-4 per cent upper-caste brahmins, and a little over 20 per cent scheduled castes), and land inequality (with gini indices of 0.483 and 0.496)\(^4\). The experiences of the colonial period also did not offer any reason to expect a difference. During this period, half of the districts of the north were, like princely Mysore in the south, subjected to indirect rule. The other half were part of the Bombay Presidency, just like the central districts that lie within the former Mysore Empire boundary\(^5\).

However, a deep internal divide cuts across modern Karnataka. Since independence, “the region belonging to the northern part of the state has been lagging behind in almost all facets of development relative to the southern region of Karnataka” (Prabhakar 2010). As a whole, Karnataka is known as one of the better governed states in the country, with compact, efficient and responsive civil service compared to most other states (Rao, 2005), one of India’s best fiscal management records (Raju 2012)\(^6\). Independent surveys of ser-

\(^4\)District-level data on household income from the India Human Development Survey (Desai et al. 2006), show similar averages (48,000 Indian rupees, in 2005) though official statistics show a large discrepancy; data on land inequality are from Iyer (2010).

\(^5\)Bellary district was assigned to Madras, again under direct British rule, while Raichur was part of Hyderabad. The southern portion of Karnataka was part of the rump princely state of Mysore, the remaining territories left by the defeat of the Mysore Empire in 1799.

\(^6\)Combining both high and regular tax revenues (at 9.8 per cent of gross state product, the highest in India) with minimal state debts, as annual interest payments on debt are just 1.8 per cent of state GDP, the lowest
service delivery across “health, education, bus transport, and subsidised distribution” place “Karnataka in the top decile of India’s states” (Paul et al. 2004). However, in northern districts, areas where the Mysore Empire never ruled, assessments of local government are among the lowest decile in all of India (Desai et al. 2006). This perception gap is shown in Figure 4.6, which shows the difference in the levels of confidence in local government between northern and southern districts.

Perceptions of poor service reflect the mediocre reality on the ground. Karnataka’s state government may be rated as one of the best investors in education, health and infrastructure in India. At the same time, the data on public goods provision in 1980 and 1990 show that in the 1980s, there was not one single tuberculosis clinic, nursing home, specialized health center, or primary health subcenter in any of the northern district villages. On the contrary, all of these facilities existed in the south, including in districts relatively close to the northern boundary line. As far as the more basic healthcare services, such as medical dispensaries or hospitals, the differences in village-level access were stark. Over three times among Indian states.

7When asked to rate their level of confidence in a selection of public officials, ranging from the police, to politicians, to courts and judges, the only group rated highly in the north is the army.
as many possessed hospitals, and five times as many had maternity homes\textsuperscript{8}. The “north” of Karnataka trails the “south” across every category of service provision, from roads, to hospitals, to recruitment of teachers, public employee absenteeism, student dropout rates, urban sanitation, or water supply (Das 2005). A qualitative assessment of schools across Karnataka in 2005 found northern institutions consisting of “dilapidated school buildings, often consisting of a single room with bare floor, with no seating facilities, no storage facilities to maintain teaching aids and equipments, if any, no playground, toilets.” However, “the situation dramatically changes as one moves to developed localities in southern [regions of the state]” (Honnalli 2006: 104).

Why is there such a deep divide between the north and the south of one of India’s most economically developed states? One theory might be that southern politicians have more successfully monopolized the distribution of public resources since the state’s founding in 1956, depriving the north of much needed social investment. Yet while clientelist explanations are useful for explaining public goods outcomes in the south Asian context, where political coalitions are built to reward supporters, they cannot explain the position of northern Karnataka. This is also, considering that the, political elites of northern Karnataka have, been overrepresented rather than underrepresented in the state’s modern political history (Keefer and Khemani 2004; Kitschelt and Wilkinson 2007; Kohli 1987). In the 50 years since Karnataka’s reorganization, there have been many chief ministers from North Karnataka, and the former editor of the \textit{Economic Times} of Bangalore, E. Raghavan, has argued that northerners often determine the outcome of state electios (Manor 2006). The North was also more engaged in the linguistic states movement. During the states reorganisation commission of 1954, almost all organisations and political leaders in the north submitted memoranda in favor of a unified Kannada-speaking region, unlike the southern Mysore government and Mysore Congress which refrained from giving evidences (Honnalli 2006). Alternatively, the agricultural conditions of northern Karnataka are sometimes cited as an explanation for the region’s poverty. But, even on the count if one goes by the Karnataka Human Development Report (2005), “the south is more prone to severe droughts

\textsuperscript{8}In the north there was one hospital per 1321 villages, against one per 482 villages in the south, and one maternity center per 3522 villages, against one per 642 in the south.
than the north, contrary to popular impression” (Das 2005). It can also be argued that state level politicians have ignored the plight of the north. In recent decades public policy in Karnataka has been highly attentive to the problem of its northern “backward districts,” and extensive sums have been outlaid in connecting the region with road links and rail, upgrading its educational and health infrastructure, and promoting its economic development. A plethora of initiatives since the mid-1990s have sought to bridge the growing north-south development gap, including a “Committee for Redressal of Regional Imbalances in the State,” the creation of a North-East Karnataka Road Transport Corporation (NEKRTC) to invest in regional transport links, initiatives to promote tourism in the region, the relocation of government institutions to the north, or more tailored initiatives, such as a “Mid-Day Meal Programme” across government schools of the north, to redress the region’s high dropout rate.

Despite efforts at reform from above, the results on the ground have failed to close the gap. A recent study concludes that the ultimate cause for the failure of the northern districts to deliver, is not a lack of commitment from the center, but “above all, lack of political will” from “the political elite of the region.” Interviewing the members of this elite, the author finds that “the political leadership of the region... blame bureaucracy for inefficiency, corruption, and poor implementation,” while they themselves have “withdrawn themselves from the government operated education and health institutions and have raised their own [private] institutions with high quality infrastructure” (Honnalli 2006: 166). The absence of accountability between northern elites and constituents is part of a wider anomic syndrome, characterized by apathy, disengagement, and lack of solidarity. Even the former Prime Minister Manmohan Singh, having no particular ties to the region, has remarked that “in minority pockets like Raichur, Bellary, Gulbarga and Bidar” there is “lack of communal harmony and [a] sense of insecurity” (cited in The Hindu, April 30th 2013).

The “lack of communal harmony” in northern Karnataka becomes very evident from the survey data. The modal response among village respondents, when asked whether they “generally get along with each other” or whether “there is conflict,” was the latter. This

---

9The study substantiates this charge with a very long list of private schools and hospitals in the north, listing their owners, and alleged connection to political elites.
Figure 4.7: Precolonial Mysore Boundary (Black) and Households Reporting Experience of Theft, 2005. Karnataka Districts Only. Data from the India Human Development Survey (Desai et al. 2006). Red/green cutoff at 1.2 per cent of households reporting theft.

contrasts strikingly with responses given in the south of Karnataka (Desai et al. 2006). The “sense of insecurity” is also tangible from both the crime surveys and the official data. In districts north of the 1782 boundary, 22.2 per cent of households, or over one-fifth, report that they have experienced theft in the past year. Across the south of Karnataka, this figure is only 4.1 per cent, or 1 in 25 (figure 4.7). The homicide rate in northern districts is fifty per cent higher than in the south, and there is also a higher number of riots and twice as many highway robberies.\(^\text{10}\)

The combination of unaccountable elites, apathetic citizens, and social anomie recalls the “social capital” arguments of Putnam (1993) or Knack and Zak (2002) regarding southern Italy or the American South. They show that entrenched inter-regional differences in administrative patterns are a consequence of differences in norms and networks of collective action, that persist due to the transmission of social institutions and values. However, to paraphrase Joel Migdal (1988), the problem in northern Karnataka is not so much that society is insufficiently strong, but rather, that the state has been historically too weak. Superficially, the civil society in the region appears vibrant and citizens in the north have 2.1

\(^{10}\)The murder rate per 100,000 was 3.3 in the north as against 2.2 in the south, and there were 8.9 cases of dacoity per million persons as against 4.4 in the south.
memberships per person (when asked about their memberships across a range of category of voluntary associations) compared to 0.76 in the south. Yet this apparent civic vitality is accounted for by an astonishing 77 per cent of northerners being members of “caste associations,” against only 13 per cent in the south, 75 per cent being members of “religious” societies compared to just 9 per cent among southerners (Desai et al. 2006). Together, these two forms of association together account for the entire north-south gap in voluntary activity. Social bonds despite being dense are framed around what Putnam (2000) has termed “bonding” as opposed to “bridging” ties - networks of survival, that are more concerned with the challenge of “getting by” than the goal of “getting ahead” (Woolcock and Naryan 2000).

What then are the origins of this persistent difference in societal structure? The defensive civil society of northern Karnataka, rather than being a cause of its underdeveloped public institutions is its consequence. The patrimonial ties upon which everyday life is structured - as reflected in the strong majorities participating in sectarian and caste associations - are a substitute for the existence of an impartial service provision, security, and the rule of law. As Prabhakar (2010) writes in describing the situation of northern Karnataka, ascriptive identities such as caste thrive in an environment where impartial institutions are unavailable. In the north of Karnataka, caste has “found a fresh lease of life in modern institutions such as panchayats, schools, factories and government offices” with individuals taking advantage of identity networks to access excludable goods and benefits. The response of citizens to a historical deficiency in public order, public goods and accountability has not been to “hunker down” but rather to mobilize segmentally in order to capture private gains. This has been done with the cooptation of political elites, as in the 1970s, in order to build a political clientele. The Chief Minister “used illicit funds to establish and then build up caste associations among disadvantaged groups that had previously been poorly organised or entirely unorganised” (Mason 2005). As will be explored in the next chapter, the origins of this pattern lie in the divergent paths taken by the two regions in the eighteenth century. In this period southern Karnataka experienced centralization under the Mysore Empire and the extension of bureaucratic administration, while the north underwent a steady path of state failure and the balkanization of political authority. The origin of the north-south divide in Karnataka has therefore deep historical origins. A recent report
The State of Travancore within Kerala and Tamil Nadu

The Kingdom of Travancore ruled the southwestern tip of India from its founding in 1729 up to its abolition following States Reorganisation in 1956. This also included a brief period from 1948-1956 spent in union with Cochin. Among the states classified by Ramusack (2003) as “warrior,” or challenger regimes, Travancore uniquely survived the ascendency of the British Empire, a result of a strategic alliance that was made following Mysorean invasion in the late eighteenth century. The Kingdom of Travancore cuts across two states of contemporary India, comprising the southern half of Kerala along with the district of Kanyakumari in Tamil Nadu. Following States Reorganisation in 1956, the district of Kanyakumari was prized from Travancore and made a part of the new Tamil state. The rest of Travancore was dissolved into the new state of Kerala and thereafter joined with the Malayali-speaking districts of Malabar that had mobilized for the creation of a united Malayali homeland.

Figure 4.8: Precolonial Travancore Boundary (Black) and Hospitals per Village, 1980-90. Source: District Gazetteers of Kerala (1995). Green/brown cutoff set at 25 per cent of villages with hospital access.
The district of Kanniyakumari: An Enclave of Travancore in Tamil Nadu

The district of Kanniyakumari, at the southern tip of Tamil Nadu, is a Tamil-speaking district that was separated from Travancore and joined to the new state of Tamil Nadu in 1956. Once known as Cape Comorin, it is located at the southernmost point of the sub-continent where the Indian Ocean meets the Arabian Sea. Its long cliffs and beaches lined by coconut groves, similar to those of neighboring Kerala, earned it a description among European explorers as “a fairy land” (Aiyar 1920: 1)\(^1\). In common with southern Kerala, historically Kanniyakumari was part of the domain of the Travancore kings. Before the Travancore kings, it was under their Venad predecessors, who first established rule over the area in the first quarter of the 12th century, and established a capital there until 1747 (Gopalakrishnan 1995: 71)\(^2\). The separation of Kanniyakumari from Kerala arose following the movement for “States Reorganisation” in the 1940s and 1950s. Several parts of Travancore were mobilized into the united Kerala (\textit{Aikya Kerala}) movement that called for a united Malayali homeland. However, Tamils in Travancore were faced with the possibility of becoming a marginalized minority, unlike in Travancore where Tamil had enjoyed an official status as a language of state. As the prospect dawned of a Malayali federation, an All Travancore Tamil Congress was founded that won all seats in the Tamil-speaking areas during the 1954 elections. Under States Reorganisation two years later, these wards were subsequently adjoined to Tamil Nadu.

Much has been written regarding the superior social and governance record of Kerala with respect to other states of India (e.g. Heller 1996, 2000; Singh 2012), as well as the extent to which this derives from the institutional legacy of Travancore, or from the policies of the state’s post-1956 left governments (Sen 2004; Franke and Chasin 1989). The existence of a portion of Travancore which was never incorporated into the post-1956 Kerala state allows us to test such arguments under very strict control conditions. Today there are no particular reasons to expect Kanniyakumari to be especially better governed than its neighboring Tamil districts to the east: Kanniyakumari is poorer than the rest of

\(^{11}\)The description of Comorin as a “fairy land” is attributed to Lord Commemera, though repeated by other explorers.

\(^{12}\)The Venads did not rule continuously from this point, however, until the foundation of Travancore in 1729, and independent Tamil nayaks also ruled Kanniyakumari during the late medieval era.
Tamil Nadu, with an average household income of Rs. 36,042, against 39,141 in the rest of the state. It is also more socially fractionalized, with a large Christian as well as a smaller Muslim minority. Nonetheless, on every single metric of public service delivery or rule of law, Kanniyakumari is far in advance of other Tamil districts, so much so that it is not within a several standard deviation range. Kanniyakumari’s outperformance covers almost every category of public goods, from high schools (82.6 versus 17.6 per cent of villages with access), to hospitals (41 per cent versus 3.6 per cent), medical dispensaries (37.6 per cent versus 2.8 per cent) or nursing homes (5.8 per cent against 0.3 per cent). The gap between Kanniyakumari and the rest of the state on these measures is extremely wide: for example, the provision of hospitals or nursing homes is 12 standard deviations above the district average for the rest of Tamil Nadu, the provision of medical dispensaries, 18 standard deviations from the Tamil average, and, since joining the state of Tamil Nadu, the literacy rate in Kanniyakumari has consistently been 25 percentage points higher than in the rest of the state. The record of public service provision in this former Travancore enclave of Tamil Nadu is also reflected in citizen assessments of public institutions. While Kanniyakumari residents express substantially lower-than-average confidence in “the Tamil Nadu state government,” they are well above the Tamil average when asked about how they assess their own schools, hospitals and courts.

How can we account for the outlier position of this small slice of Travancore, alone within the Tamil state? Besides the obvious inference that this was a direct consequence of Travancore policies and institutions, legacies of direct versus indirect rule have also been suggested as one reason for the outperformance of certain regions of India. It is true that much of Tamil Nadu came under direct rule as part of the Madras Presidency, while Kanniyakumari, as part of Travancore, remained self-governing (Iyer 2010; Mukherjee 2013). However the other part of Tamil Nadu that was governed indirectly during the colonial era

---

13 These estimates are from the Human Development Survey of 2006.

14 In 1980, the literacy rates of Kanniyakumari and the other districts of Tamil Nadu were 48 and 71 per cent, and in 2011, 55 and 80 per cent respectively.

15 While the Kanniyakumari average confidence assessment regarding the Tamil state government is 1.2 standard deviations below the average in other Tamil districts, the assessment for hospitals and courts are one standard deviation above the Tamil average. This would suggest that Tamils in Kanniyakumari have a qualitatively different judgement of their local institutions than the institutions of the federal state as a whole.
- Pudukottai - performs worse on most metrics of governance than either Kanniyakumari or indeed the rest of Tamil Nadu. It has even fewer schools, social services, bus services or paved roads than other districts. It may be a better representative of the trajectory that the rest of Tamil Nadu might have taken, had it remained independent of British influence.\(^{16}\)

Kanniyakumari’s public goods provision may be an outlier in Tamil Nadu, but not in comparison to southern Kerala, with whom Kanniyakumari shared a government for more than two centuries prior to 1956.\(^{17}\) More than a generation after Tamil unification, Kanniyakumari remains overwhelmingly more similar to its Malayali-speaking former compatriots in Kerala, than to its Tamil-speaking counterparts in its own state. As the 1995 Kanniyakumari District Gazetteer states in its introduction, before going on to list the district’s achievements, “the areas that came under Kanniyakumari now... were for a very long time, with the Travancore State [and everywhere] the traces of influence of the culture of the neighbouring Kerala State are still to be seen” (Gopalakrishnan 1995: 8).

In a pattern that is very similar to the comparison between northern and southern Karnataka, the social fabric of Kanniyakumari is qualitatively different from the other regions. It is more law-abiding, stable, and not characterized by the predominance of civic associations based upon caste. “Kanniyakumari is generally a tranquil district where the law and order problems were few and far between,” and despite being religiously diverse, there is “social amity and harmony among the peoples of this area” (Gopalakrishnan 1995: 631-3). Almost half as many households in Kanniyakumari as in other Tamil regions report have experienced an act of assault in the past 12 months, while official statistics show lower levels of reported crime across every single category.\(^{18}\) In a country riven by police cor-

\(^{16}\) As a minor poligar which survived into the colonial era completely “by chance” - arising from a circumstance alliance with the East India Company during the Anglo-Mysore wars - Pudukottai is more similar to the poligars of interior Tamil Nadu than Travancore, a relatively centralized and modernizing regime already in the eighteenth century.

\(^{17}\) Taking 2 standard deviations as the significance threshold for a two-tailed significance test at \(\alpha = 0.05\), the Kanniyakumari level of service provision is not significantly differently from other former Travancore districts with respect to primary and secondary education, hospitals, medical dispensaries, maternity homes, or child welfare centers.

\(^{18}\) The categories reported by the Tamil statistical office include not only major crimes such as homicide, rape, and kidnapping, but also minor misdemeanors such as accusations of cheating, counterfeiting, and breaches of the peace. Even given low crime rates within Tamil Nadu with respect to India, Kanniyakumari remains below the Tamil average on every measure.
ruption, official statistics even go so far as to report that of all the items handed over to the police as missing or lost, more than 50 per cent are successfully returned to their rightful owners (Gopalakrishnan 1995: 640-1). Over this, almost no respondents (less than one per cent) report being members of caste associations, compared to 1 in every 15 respondents in Tamil Nadu.

**Northern and Southern Kerala**

Just as the Travancore enclave of Kaniyakumari is an outlier within the state of Tamil Nadu, a similar if less pronounced “Travancore effect” cuts through the state of Kerala itself. Former Travancore districts in the south have substantially better public goods provisions than in the districts of Malabar to the north. While the southern half of Kerala was ruled consistently by the state of Travancore from 1729 to 1956, the northern half - Malabar - was ruled by independent Nair chiefs until the eighteenth century, when it came under the subjection of Haider Ali from 1763 to 1782. Thereafter, it was made a district of the Madras Presidency, to be governed directly by the British Empire.

Despite being on average relatively poor, the annual household income of Rs. 62,649 versus Rs. 68,515 across the north, the former Travancore districts in the southern half of the state have significantly lower poverty rates (8.5 per cent against 28 per cent of households), and much better access to school, hospitals, and other public services. 91 per cent of villages in former Travancore districts had a high school (against 73 per cent in other regions), 30 per cent had access to a hospital (against 14.5 per cent elsewhere), and 70 per cent access to a nursing home (against 37.5 per cent outside of former Travancore). The data also show better access to a range of communications services in the south, including post offices, telegraph, and public phones. Respondents in Travancore had similar confidence in public institutions as those outside. Travancore residents also had a higher confidence in courts, police, schools and medical services.

There is a very extensive literature on the outlier position of Kerala with respect to the rest of India, and a very wide number of theories to explain this gap ranging from the poli-

---

19 These data are for the period from 1980 to 1990, during which over two-thirds of villages had access against less than half in the north. Telegraph services were finally ceased in India in 2013. The data show three-quarters of villages in former Travancore with public phones as against 50 per cent in other districts.
cies of post-independence left governments (Franke and Chasin 1989; Heller 1996, 2000), the influence of gender and matrilineal institutions (Chacko 2003; Robin 1987), subnational mobilization in the early twentieth century (Singh 2012), the effect of Christian missionaries in providing schools and challenging caste structures (Lankina and Getachew 2013), the region’s exposure to trade (Gaikwad 2015), and of course, the legacy of Travancore policies and institutions (Sen 1991).

However, while these theories may be empirically valid as explanations for Kerala’s outlier status relative to the rest of India, none of these theories save the last can explain why regions of Travancore within Kerala perform better than Malayali-speaking regions in the same state - or why Travancore’s Tamil offshoot in Kanniyakumari, lacking Malayali family institutions, subnationalist mobilization, or the same state government since independence, should so significantly outperform the rest of Tamil Nadu.

*The Maratha Peshwa in Western Maharashtra*

In the second half of the eighteenth and the first few decades of the nineteenth century, the Maratha peshwa in Pune directly ruled the section of land that now covers the western portion of the state of Maharashtra. The eastern districts were under the formerly Mughal region of Hyderabad, and the competing Bhonsle and Scindia princes. As the heart of the Maratha alliance dominating eighteenth-century India, the system of governance in the
peshwa’s lands was “the most elaborate system of documentary control India had known up to that point,” with an extensive administrative apparatus radiating from the royal court in Pune (Stein 1996: 191). Following the Anglo-Maratha wars, the peshwa was defeated in 1818 and his territories came largely under direct British rule as part of the Bombay Presidency. With the rise of the linguistic states movement, an agreement was reached in September 1953 between the leaders of Marathi-speaking Marathwada (under the rule of the Nizam of Hyderabad), Vidarbha (directly ruled by the British following the annexation of Nagpur in 1853), and Bombay State. These formed a single Marathi entity with separate allocations of funds for each of the three regions and separate protections with respect to matters concerning education and employment in government services. With States Reorganisation, the Bombay State thus lost its Gujurati-speaking areas to the north, and instead, acquired new Marathi-speaking districts to the east, equal in size to the rest of Maharashtra.

Today, there remains a clear east-west division within the state of Maharashtra that is broadly coterminous with the distinction between areas that were or were not governed by the Peshwa in the early nineteenth century. In regions formerly under the peshwa’s direct rule, there is substantially greater confidence in local government, a lower rate of crime, and higher provision of schools, hospitals, healthcare clinics, public bus services, and paved roads. Such regions are among the most developed in India today, while by contrast, as Kanbur (2010) observes, the districts of “inland Eastern Maharashtra are examples of lagging regions within prosperous states [that] form a contiguous corridor with deprived areas of Andhra Pradesh, Orissa, Chhattisgarh, Jharkhand and Bihar.” “Although Maharashtra’s Infant Mortality Rate (IMR) is near the best in the country, its worst districts have IMRs that are higher than those of much lower ranked states” (World Bank 2007: 21-24). While the new districts were from the start less economically developed than western Maharashtra, this disparity has even increased over time. The share of eastern Maharashtra in gross state product fell from 30 per cent in the early 1990s, to 25 per cent in the present day, despite massive government intervention through regional development boards (Government of Maharashtra 2013).
Figure 4.10: Precolonial Peshwa Boundary (Black) and Medical Dispensaries, 1980-90. Districts of contemporary Maharashtra, peshwa’s territory prior to 1818, and percentage of villages with medical dispensaries, 1980-90. Green/brown threshold set at 2.5 per cent. A clear divide is visible between western Maharashtra, then under the Bombay Presidency, and eastern Maharashtra, split between the Bombay presidency, the Central Provinces, and Hyderabad.

The East-West gap within Maharashtra covers the full range of indicators, from crime rates (with over twice as many households in eastern districts areas reporting that they have been victims of theft or assault), to the numbers of schools and health facilities, to bus services, paved roads, or post offices. The gap in village-level access to medical dispensaries, for example, is shown in Figure 4.10. While there are enclaves of service delivery in the eastern portion of the state, the overall pattern changes markedly across the eastern boundary. Similar disparities are seen in a range of indicators, such as paved road provision, they are shown in Figure 4.11.

Why is there such a deep divide between the east and west of Maharashtra? As in northern Karnataka, politicians in less developed regions allege “that this underdevelopment is a result of deliberate neglect on the part of the state government” (Phansalkar 2003). However, the evidence points towards an opposite conclusion. The requirement for equal representation of regions, specified in subclause 6 of the Nagpur Agreement, has largely been honoured. State governments since 1995 have had between 36 and 43 per cent of ministers from eastern Maharashtra since 1995, consistent with the balance of demography (Palshikar
Figure 4.11: Precolonial Peshwa Boundary (Black) and Percentage Paved Roads, 1980-90. Districts of contemporary Maharashtra, peshwa’s territory prior to 1818, and proportion of paved roads in 1980-90. Green/red threshold set at 44 per cent of paved roads. A clear divide is visible between western Maharashtra, then under the Bombay Presidency, and eastern Maharashtra, split between the Bombay presidency, the Central Provinces, and Hyderabad.

2012). At the same time, “many Chief Ministers have been from Marathwada,” a region in the east of the state (Government of Maharashtra 2013). Consistent with political representation, substantial allocation of resources has been made to overcome the disadvantage of the east. Though there have been ongoing controversies regarding backlogs in disbursement of funds from the centre to the regions, the underdeveloped parts of the state have been allocated preferential funds overseen, since 1994, by regional development boards.

And yet, despite massive development schemes and “a scheme of lagging area industrialization primarily designed around sales tax exemption for five or more years,” “the measures applied to redeem this condition seem to have worked much feebly than expected.” Considering the reasons for the failure of microeconomic incentives to encourage industrial relocation, a recent official report cites “above all [the] high costs of doing business in these regions” – a byword for political interference, corruption, and regulatory barriers (Government of Maharashtra 2013). Evidence of the scale of corruption in the governance of eastern Maharashtra is being exposed by the ongoing “Maharashtra Irrigation Scam”: which,

---

20Maharashtra is divided with about 40 per cent of the population in the east and 60 per cent in the west, the same as the regional representation in the state government.
with a total cost of misallocated funds at Rs. 350bn (US$5.6bn), is rapidly becoming one of the largest corruption cases in modern Indian history. Following the establishment of regional development boards in the mid-1990s, vast sums were allocated to irrigate areas of eastern Maharashtra, with the purpose of promoting economic development. After two decades, while “320 projects have been undertaken by the Vidarbh Irrigation Development Corporation” official figures indicate that only 26,000 hectares have been irrigated out of a target of 1,580,000 since 1997 i.e. – just 1.6 per cent of the amount promised (Pallavi 2012). Another estimate by the state’s 2012 Economic Survey estimated an increase in irrigation potential of just 0.1 per cent during the past decade (Menon 2012). Deliberate graft is a principle cause of budgetary mismanagement. As the president of a local citizens’ group alleges, repeated project alterations are “made with the express purpose of inflating costs.” This alone has allowed politicians to give up to Rs. 4bn ($64.5m) in “mobilisation and machinery advances years before work can be started.”

What are the roots of these differences in governance and accountability between regions? The Maharashtra state government’s own report into the east-west divide notes that in western regions there is “greater political cohesion and consensus among different communities.” It also states that there is “a longer legacy of progressive social movements” in the west, which “may have also exercised some impact on the political process and administrative system” (Government of Maharashtra 2013). This however, still does not explain the roots of these differences in social institutions. Academic and policy discussions of the east-west division in Maharashtra typically note that the “three regions that were merged to form the Marathi speaking state of Maharashtra had very different social-political histories.” The origins of this historical divergence go beyond the colonial period. The regions of Maharashtra which were under direct British rule are the ones that are today better governed than the ones that were ruled indirectly, whether by princely rulers in Hyderabad or those of Berar. The latter suffer from crime, corruption, and poor service delivery. Iyer (2010) shows that this cannot work as a generalized hypothesis across the subcontinent.

---

21 Other estimates have been higher, as high as Rs. 1,700bn ($27.5bn).

22 This reflects the history of the region’s colonization, as the Maratha peshwa was one of the principal opponents of the East India Company in the early nineteenth century. Upon his defeat at the hands of the British in 1818, his territories were assumed by the Company, while weaker princes that had allied with the
To understand the persistent divide between the eastern and western parts of Maharashtra, it is necessary to go back to the eighteenth century, and the original divide between two systems of government: a system of bureaucratic centralism in the west, based around the Maratha peshwa in Pune, and the feudatory governance of the east, under the Bhonsle princes.

**Conclusion**

This chapter has tested for the legacy of precolonial states upon district and local government effectiveness in contemporary India. By tracing the precolonial boundaries of eighteenth-century South Asian polities, this chapter has shown that districts of India which lie narrowly within the boundary lines of historically centralized states perform significantly better on a wide variety of district-level indicators of state effectiveness than those narrowly outside these boundaries – despite the fact that these borders largely ceased to exist in the early nineteenth century. These estimated effects are robust to a wide variety of controls, placebo tests for border displacement, the exclusion of individual polities, and controls for the boundaries of India’s contemporary federal states. In the next chapter, I will argue that following the decline of the Mughal Empire in the eighteenth century, warfare between “challenger states” prompted an accumulation of bureaucratic and fiscal capacity at the local level, and that this capacity has persisted through the colonial era to the present day. In contrast to “bottom-up” theories of state capacity which root institutional strength in societal characteristics such as ethnic homogeneity, social capital, or land equality, it is argued that government effectiveness is cumulatively built through long-term historical investments in state capacity, and that, in India, an important phase of investment occurred during the warring states period of the eighteenth century.

---

British were allowed to continue as princely domains. In this way, the pattern of state capacity from the precolonial past transferred to the colonial era in a manner that saw the most developed regions integrated into the colonial state, while peripheral provinces retained relative autonomy.
It is ever justifiable to:

<table>
<thead>
<tr>
<th></th>
<th>government benefits</th>
<th>take a bribe</th>
<th>cheat on taxes</th>
<th>tax revenues as % state GDP</th>
<th>court cases lasting 5 years or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precolonial legacy states</td>
<td>16.5%</td>
<td>12.0%</td>
<td>14.0%</td>
<td>7.13%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Other states and territories</td>
<td>40.8%</td>
<td>21.6%</td>
<td>24.5%</td>
<td>5.41%</td>
<td>17.3%</td>
</tr>
</tbody>
</table>

Table 4.8: State Capacity and Compliance Among Precolonial Legacy States. Federal states with precolonial heritage are those deriving from a single dominant precolonial polity during 1757-1856: Karnataka (Mysore), Kerala (Travancore), Sikkim, Punjab (Sikh Empire), Maharashtra (Maratha Empire), Tripura, Andhra Pradesh (Hyderabad). Contempora- rary federal states without a single dominant precolonial polity are Bihar, Haryana, Tamil Nadu, West Bengal, Delhi, Uttar Pradesh, Himachal Pradesh, Madhya Pradesh, Chhattisgarh, Mizoram, Nagaland, and Arunachal Pradesh.
<table>
<thead>
<tr>
<th>Independent variable:</th>
<th>conf: local govt</th>
<th>conf: police</th>
<th>cheating</th>
<th>victim: theft</th>
<th>robbery</th>
<th>kipnap</th>
<th>rape</th>
<th>literacy</th>
<th>gengap</th>
<th>medical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenger State</td>
<td>0.19**</td>
<td>0.12*</td>
<td>-27.46*</td>
<td>-0.04***</td>
<td>-25.97***</td>
<td>-40.61***</td>
<td>-16.52***</td>
<td>11.4***</td>
<td>-0.03***</td>
<td>0.05***</td>
</tr>
<tr>
<td>% Scheduled Castes</td>
<td>-0.07</td>
<td>-0.3**</td>
<td>86.26***</td>
<td>(0.01)</td>
<td>(7.29)</td>
<td>(11.17)</td>
<td>(3.83)</td>
<td>(1.89)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>% Scheduled Tribes</td>
<td>-0.01</td>
<td>0.14</td>
<td>-35.7</td>
<td>-0.01</td>
<td>-23.16</td>
<td>-3.17</td>
<td>13.84*</td>
<td>-1.47</td>
<td>-0.04*</td>
<td>-0.03*</td>
</tr>
<tr>
<td>British direct rule (0/1)</td>
<td>0.01</td>
<td>-0.07</td>
<td>-13.88</td>
<td>0.01</td>
<td>2.97</td>
<td>20.19*</td>
<td>5.89*</td>
<td>7.3***</td>
<td>-0.01</td>
<td>-0.02***</td>
</tr>
<tr>
<td>Income (Rs, 000)</td>
<td>0.01</td>
<td>0.58***</td>
<td>0*</td>
<td>0.29**</td>
<td>0.36*</td>
<td>-0.04</td>
<td>0.07*</td>
<td>0</td>
<td>0.0***</td>
<td></td>
</tr>
<tr>
<td>Social Capital Index</td>
<td>-0.2</td>
<td>0.24</td>
<td>18</td>
<td>0.18***</td>
<td>-32.88</td>
<td>-115.54*</td>
<td>-14.78</td>
<td>31.2***</td>
<td>-0.12**</td>
<td>0.02</td>
</tr>
<tr>
<td>Land Inequality (Gini)</td>
<td>0.73**</td>
<td>0.45</td>
<td>48.27</td>
<td>-0.04</td>
<td>-36.77</td>
<td>-61.93</td>
<td>71.07***</td>
<td>-9.01</td>
<td>0.04</td>
<td>-0.03</td>
</tr>
<tr>
<td>State: Kerala</td>
<td>0.13</td>
<td>0.33*</td>
<td>42.21</td>
<td>-0.01</td>
<td>-29.11</td>
<td>-17.41</td>
<td>25.7***</td>
<td>33.13***</td>
<td>-0.14***</td>
<td>0.37***</td>
</tr>
<tr>
<td>State: Karnataka</td>
<td>-0.24**</td>
<td>0.28***</td>
<td>28.5</td>
<td>0.08***</td>
<td>-3.07</td>
<td>-0.03</td>
<td>-6.57</td>
<td>-0.67</td>
<td>-0.01</td>
<td>-0.08***</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.06***</td>
<td>-2.14***</td>
<td>0.05*</td>
<td>53.87**</td>
<td>72.77**</td>
<td>33.7***</td>
<td>0.25**</td>
<td>0.05*</td>
<td>0.05*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>325</td>
<td>325</td>
<td>323</td>
<td>325</td>
<td>323</td>
<td>323</td>
<td>324</td>
<td>325</td>
<td>325</td>
<td>325</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>0.05</td>
<td>0.21</td>
<td>0.11</td>
<td>0.13</td>
<td>0.11</td>
<td>0.12</td>
<td>0.17</td>
<td>0.52</td>
<td>0.36</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 4.9: Basic OLS, All Districts. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level. All districts.
Table 4.10: Descriptive Statistics Along Border Regions of Precolonial States. Sample: All Districts.
<table>
<thead>
<tr>
<th></th>
<th>Confidence: local govt.</th>
<th>Confidence: police</th>
<th>% Survey reported thefts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>inside</td>
<td>outside</td>
<td>diff</td>
</tr>
<tr>
<td><strong>Challenger States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sikh Empire</td>
<td>8.101</td>
<td>7.099</td>
<td>1.002</td>
</tr>
<tr>
<td>Maratha (Peshwa)</td>
<td>6.931</td>
<td>7.152</td>
<td>-0.221</td>
</tr>
<tr>
<td>Mysore</td>
<td>7.228</td>
<td>6.976</td>
<td>0.252</td>
</tr>
<tr>
<td>Travancore</td>
<td>7.031</td>
<td>7.503</td>
<td>-0.472</td>
</tr>
<tr>
<td><strong>Mughal Provinces</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyderbad</td>
<td>7.67</td>
<td>8.094</td>
<td>-0.424</td>
</tr>
<tr>
<td>Oudh</td>
<td>7.202</td>
<td>6.927</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>Robberies</td>
<td>inside</td>
<td>outside</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Challenger States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sikh Empire</td>
<td>14.682</td>
<td>33.6</td>
<td>-18.918</td>
</tr>
<tr>
<td>Maratha (Peshwa)</td>
<td>31.6</td>
<td>22.4</td>
<td>9.2</td>
</tr>
<tr>
<td>Mysore</td>
<td>33.591</td>
<td>39.231</td>
<td>-5.64</td>
</tr>
<tr>
<td><strong>Mughal Provinces</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyderbad</td>
<td>29.733</td>
<td>30</td>
<td>-0.267</td>
</tr>
<tr>
<td>Oudh</td>
<td>44.967</td>
<td>68.381</td>
<td>-23.414</td>
</tr>
<tr>
<td></td>
<td>Literacy Rate</td>
<td>inside</td>
<td>outside</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Challenger States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sikh Empire</td>
<td>59.609</td>
<td>52.68</td>
<td>6.929</td>
</tr>
<tr>
<td>Maratha (Peshwa)</td>
<td>54.38</td>
<td>44.564</td>
<td>9.816</td>
</tr>
<tr>
<td>Mysore</td>
<td>52.173</td>
<td>45.954</td>
<td>6.219</td>
</tr>
<tr>
<td>Travancore</td>
<td>88.16</td>
<td>57.533</td>
<td>30.627</td>
</tr>
<tr>
<td><strong>Mughal Provinces</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyderbad</td>
<td>35.82</td>
<td>45.881</td>
<td>-10.061</td>
</tr>
<tr>
<td>Oudh</td>
<td>36.637</td>
<td>36.976</td>
<td>-0.339</td>
</tr>
</tbody>
</table>

Table 4.11: Descriptive Statistics Along Border Regions of Precolonial States. Sample: Bordering Districts Only.
Table 4.12: Within-State Comparisons: Contemporary Data. Data on confidence in local and state government, and survey reported crime victimization, are taken from the 2006 Human Development Survey (Desai et al. 2006). Data on kidnappings and hospitals are from Iyer (2010), and literacy rates from the 2011 India Census. “Relative confidence in local government” is the gap on a confidence scale between scores assigned to the state government, and a separate item regarding local government in the village or neighborhood.
Chapter 5
Trajectories of State Formation

Contrary to the population flows of the twentieth century, many northerners migrated southwards during the sixteenth century, drawn by the relative prosperity of the South, coupled with the dismal downfall of the North... The collapse of communal republicanism led to a kind of “re-feudalization” of the Italian peninsula”

- Putnam, Making Democracy Work (p. 135), 1993

The preceding chapters have shown that in the late eighteenth century large differences existed between parts of India in the degree of state formation and political centralization, and that districts of India narrowly inside the boundaries of centralized precolonial regimes exhibit higher levels of state capacity today than those narrowly outside such boundaries. An important question remains, however, in understanding how such patterns came into existence. Why and how it is that some regions adopted more centralized regimes during the late precolonial era? And why did other parts of India slide into political decay? For India on the eve of colonization was neither stagnant nor universally in decline. While in the center-north, Mughal Hindustan had began a descent into political fragmentation that was in many ways accentuated during the early colonial era, in the south of India eighteenth-century states were undergoing extensive defensive modernization, and their administrative capacity, the legitimacy of their public institutions, and reformed fiscal and land tenure systems persisted through the imperial era.
The rest of this chapter traces these diverging trajectories. First, I show that as the Mughal imperium began to fragment, “challenger regimes” on the periphery of Mughal power began to mobilize against the status quo ante, and that relative patterns of military mobilization across the subcontinent during the eighteenth century led to the accumulation of fiscal and bureaucratic capacity. Second, I then show how political order fragmented within the successor polities of the Mughal empire, leading to the low state capacity, corruption, and deficiency in rule of law that has persisted within the region from the late precolonial era to the present.

In the context of the general dissertation, this chapter also makes two analytical contributions. Firstly, I demonstrate that one of the motivations for elites to build state capacity is defensive modernization, and in particular the pressure of warfare. Secondly, I argue that this is a conditional hypothesis: elites invested in military capacity where the coalitional structure of the regime facilitated the exclusion of the warlord and landowning classes, and by consequence the centralization of fiscal and military organization. By contrast, where the structure of rightsholding left warlords as powerful actors, political centralization did not occur, and in many cases a process of “repatrimonialism” could be observed as Mughal institutions were steadily eroded.

India on the Eve of Colonization

Over the course of the sixteenth and seventeenth centuries, the Mughal Empire had consolidated its hold over India’s north and center. This imperium reached its zenith under Aurangzeb in the late seventeenth century, before entering a stark decline after his death in 1707. Though historians in the last two generations “have begun to ask questions about the nature of the Mughal empire, particularly about the degree to which it was ever a centralized state” nonetheless there is little doubt about the period that followed (Bayly 1988: 8). As Stein concludes, “after 1707, many imperialist and nationalist historians agree, an era of devastating wars commenced, causing economic decline and anarchy” (Stein 1996: 196). In a short period of time, the fragile order generated by the Mughal polity had evaporated. At the start of the century, Aurangzeb had been confident enough to wage war in the south and reimpose the jizya (poll tax) on Hindus; one century later, “the Mughal [emperor] was
reduced in European eyes to the status of a pathetic ‘tinsel sovereign’, surrounded by the
emaciated ladies of his harem and chamberlains who maintained the shadow of his author-
ity through the recitation of court rituals”. A popular quotation of the day exclaimed that
the “empire of Shah Alam extends from Delhi to Palam” – the current site of Delhi’s Inter-
national Airport. This was no mere figure of speech: at its height, the “Maratha tributary
regime extended itself to within fifty miles” of the Mughal throne (Stein 1996: 190). The
great Mughal could claim titular authority, but no power; he had a comparable control over
regions of the Indian subcontinent as a medieval Pope over western Christendom.

The demise of the Mughal imperium sets the context for the diverging trajectories of
state formation in eighteenth-century India. As the power of the Mughal court in Delhi
waned, the territories of the old Mughal realm could be divided into two camps. On the one
hand were those areas, such as the Punjab or western Maharashtra, which rose in defiance of Mughal order, often based on peasant mobilization and based on sectarian rejection of Islamic rule, to establish a new regime cleansed of the status quo ante – with both the Sikhs in Punjab and the Marathas explicitly claiming a theological support for their struggle to found a new Sikh or Hindu polity, respectively. Second, there were those areas, such as Awadh or Bengal, in which governors of the Mughal empire, nominally subordinate to the great Mughal in Delhi, gradually founded their own independent provincial realms. These areas were subject to a steady process of “repatrimonialism” that evident across a number of areas: the steady breakdown in security and the rule of law due to warfare between tax administrators; the spread of prebendalism among former offices of the Mughal state; and the steady undermining of fiscal capacity.

**Rising States of the South**

In no part of India did political development proceed as rapidly in the eighteenth century as it did across the south. As Stein (1984) remarks, “centralization was the signature of regimes of eighteenth century India.” Nowhere did state formation proceed as fast as among the states that dominated the southwest: Mysore, Travancore, and the Maratha Confederacy. To some extent these changes were prompted by conflict with European powers, as was the case for the Mysore Kingdom during the period of the four Anglo-Mysore wars - in 1767-1769, 1780-1784, 1789-1792 and 1799 - as well as the Marathas during their East India Company wars in 1775-1782, 1803-5 and 1817-8. It is also true of Travancore, which marked the first colonial victory of an Asiatic state by defeating the Dutch East India Company in 1741. However, the new states of southern India just as often fought wars against other South Asian polities, for example, Mysore and Travancore in 1782-4 and 1789-90, the Marathas and the Afghans in 1761, or the Marathas and Mysore in 1785-87. So extensive was warfare in the second half of the eighteenth century that there were very few years in which one or other major southern Indian powers was not in a state of active military deployment. Just as social scientists and historians examining western Europe in the sixteenth century have found the intensification of warfare (Tilly 1992), and the “mil-
itary revolution” brought about by the use of the musket and rifle to prompt new forms of state organization (Roberts 1956, Parker 1996), the same observation holds in India about one century later. The growing wave of warfare prompted challenger states, such as the Marathas for example, to develop new methods of troop and tax mobilization and the deployment of the first armies based on gunpowder weaponry, thanks to the spread of light arms (Gordon 1993). Equipped with new weapons and techniques, Indian polities began to adopt similar requirements regarding infantry conscription, cannon fortification, and the taxes required to pay for the new technology. As Bayly (1988: 97) has remarked of the Mysore regime in its last days: the “realm was not a decaying eastern despotism, but an attempt to face European mercantilist power with its own weapons: state monopoly and an aggressive ideology of expansion.”

The New Contenders: Marathas, Mysore, Travancore, the Sikh Empire

Who were the new contenders in eighteenth century India? The first newcomer to the political order of post-Mughal India was the Maratha kingdom, founded by Shivaji in the late seventeenth century in the arid plains of the western Deccan. Founded in rebellion against Mughal rule, the Marathas built an efficient core state in the Konkan interior, before expanding their rule across India through outlying principalities. While Shivaji was laying the foundations for a new political order in the western Deccan, further south, a political refoundation was also taking place in a forgotten offshoot of Vijayanagara, the last Hindu empire of the Indian subcontinent, that had collapsed in 1565 (Stein 1990). In 1673 - just a year before Shivaji’s coronation ceremony in Maharashtra - the new ruler of Mysore, Chikka Deva, refounded the realm as a centralized kingdom, killing his own chiefs at a meeting in Nanjangud and assuming then powers of an absolute monarch (Nagaraj 2003). A century later, under a new chief regent (Sarvadhikari) named Haider Ali, Mysore became the dominant power in the south of India, and the main rival to the East India Company.

Though never subject to Mughal rule, in common with Mysore and the Maratha kingdom, Travancore in the southwest of India was refounded in the 1729 as a centralized Hindu monarchy, and after defeating a Dutch invasion in 1741 also engaged in defensive modernization by building its armed forces and naval fleet along European lines. At the other end of the continent, in the northwestern plains of the Punjab rivers, the Sikhs presented a more
direct challenge to Mughal hegemony. Founded in the fifteenth century, initially the Sikh religion had no ambitions of territorial control, yet when the ninth Sikh leader (guru) was executed by Mughal emperor Aurangzeb in 1674 - the same year as Shivaji’s coronation in Maharashtra – the Sikhs, likewise, began to mobilize militarily. An army of baptized Sikhs, the khalsa, was founded in 1699, and launched a major revolt following Aurangzeb’s death in 1707, which though quashed in 1716 formed the basis of continued mobilization against Mughal rule throughout the century. In 1799, under the command of Ranjit Singh, the khalsa established a single Sikh Empire expanding across the entire Punjab region.

While the years of 1707, when the Mughal empire began to fragment after the death of emperor Aurangzeb, and 1757, when the East India Company began its territorial incursion into India proper, are often taken as the decisive “turning points” for Indian political history, in many ways 1673-1674 was as significant. The coronation of Shivaji as king of the breakaway Maratha kingdom, the refoundation of the Mysore kingdom in the south, and the execution of the ninth Sikh guru by Aurangzeb, were foundational events for India’s insurgent polities. In the space of four decades separating 1674 and 1729, the rise of new regimes completely transformed the political landscape of precolonial India. As discussed in Chapter IV, Ramusack (2004) refers to these new states as “warrior” or “conquest” states, while others have called them, more neutrally, regional states. I refer to them as “challenger” states, in recognition of the challenge they posed, first to Mughal hegemony over the subcontinent, and latterly to European attempts at subcontinental control. How did these new regimes differ from the earlier status quo? First, unlike the successor states of Awadh, Bengal, and Hyderabad, the challenger states did not participate in the governance of the Mughal empire on either provincial or imperial level. Second, these polities invoked an explicitly religious basis for governance: the Marathas and Travancore as explicitly theocratic Hindu states, and the Sikhs as an ethno-religious liberation struggle. Third, in the desire to form a direct bond with the peasantry and to differentiate themselves from Islamic suzeraignty, all of these new regimes engaged heavily in religious patronage: by the end of its rule, the Maratha peshwa dedicated 10 per cent of public spending to the temples, Travancore, which declared the Hindu god Padmanbha as its sovereign and the king as his first agent, accrued assets now worth $19bn within the vaults of the Sri Padmanbha temple in Trivandrum (Hayes 2011), and the Sikh Empire of the early nineteenth
century “alienated more revenues now in favour of the religious grantees than in Mughal
times” and sixty per cent of these to the Sikh temples (Grewal 1990: 99, 115). Fourth,
they were engaged in extensive warfare during the late precolonial era. Initially, in the
case of the Sikhs or Marathas, these battles were against the Mughal empire or their allies,
though increasingly, they became engaged in warfare against one another - and finally of
course against the European powers seeking to establish control over the subcontinent. The
pressure of constant warfare led these regimes to engage in wide-ranging administrative
and fiscal reform, which included the removal of intermediary landlords and chieftains,
improved mechanisms of bureaucratic accountability, the use of literate brahmins in ad-
ministrative service, and the introduction of European military techniques. It is this fourth
aspect which concerns us here.

War and State Formation: 1757-1818

A substantial literature rooted in the works of Charles Tilly and Otto Hintze, represent-
ing the “bellicist” approach to state formation, has argued that warfare in early modern
Europe was central to the increase in state infrastructural power (Hintze 1906, 1975; Tilly
1975, 1990; Downing 1992). According to this view, “states made war, and war made
states,” in that warfare required growing numbers of soldiers, and new technologies of war
entailed growing fiscal expenses (Tilly 1992; Roberts 1956; Parker 1976, 1996). However,
bellicist theories have rarely been successfully applied outside of the European context
(Centeno 1997). Given the large number of covariant processes that were occuring in Eu-
rope at the same time, this presents a problem for causal inference. A proper test of the
bellicist argument would require a context which was not subject to the same conditions of
socio-economic change as early modern Europe, and in which warfare and the introduction
of gunpowder weaponry had occurred relatively exogenously (Eaton and Wagoner 2014).

Such conditions do pertain, however, to eighteenth century India. Following the col-
lapse of the Mughal empire, the late precolonial era in India war a time of extensive war-
fare, to the extent that, in view of the scale of the armies deployed, the subcontinent may
have been the most important site of military conflict in the world. For the better part of
a century, war raged in the south of India, where “throughout the eighteenth century up to
the last decade no power in South India felt itself secure” (Gwynn 1932) and indigenous powers, such as Mysore and the Marathas, were mobilizing in force. Second, the qualitative nature of warfare in South Asia was changing. While the Mughals had introduced heavy artillery to the subcontinent, starting in the early 1700s the use of more rapid-fire, light armaments was becoming widespread, and European musketry was being introduced to the region. This arrival was largely exogenous, with European traders at Goa, Bombay, Cochin, Pondicherry, Madras and Calcutta being the key source of arms.

Examining the development in indigenous states during this period, I find that, in accordance with the arguments of the bellicist school, rising levels of warfare did indeed have a number of profound consequences for the internal structure of eighteenth-century Indian states. First, it increased their enlistment demand, as well as their need to invest in better technologies of warfare. Second, there was an increase in fiscal demand, in order to pay for both quantitative and qualitative improvements in warmaking capacity. Third, states had to engage in administrative reforms in order to sustain higher revenue demands on a permanent basis, and engage in bureaucratic development, which became more based on written records by trained civil servants. Finally, there was a change in the nature of political elites, which shifted from patrimonial lords to political entrepreneurs of varying kinds. However, in common with the arguments made by Ertman (1997) in the European context, I also find that warfare and the spread of gunpowder weaponry had a conditional effect: it was only in states in which executive-level elites had relative autonomy from local clientelistic interests, and leaders were able to reallocate resources away from local elites (such as landlords, warlords or chiefs) towards a central army, in which defensive modernization effectively took place. This was the condition that pertained in much of southern India, and in particular applied to the Marathas and to Mysore. On the other hand, where polities were patrimonial and dominated by powerful vested local interests able to resist moves toward centralization, rulers were unable to raise fiscal resources centrally. Such states often collapsed, or were able to survive by paying tribute to outside powers to prevent invasion. This was the condition which prevailed in Mughal successor states such as Bengal and Awadh. The second group of states will be considered later. In this section, we consider how state-making occurred sequentially in the first set.
The Growth of Military Mobilization in Precolonial India

First, the eighteenth century in India was an age in which war occurred very frequently, and chief among the belligerents were the challenger states which arose on the margins of Mughal authority, and contested for supremacy over the precolonial subcontinent. This can be seen from Figure 5.2, which shows a loess chart of the years in war for successor states of the Mughal empire, and challenger regimes in comparison.

Second, as well as the event-occurrence of war, another indicator of the rise of military mobilization is the quantitative growth of armies over the course of the eighteenth century. Though the Mughal army at its height during the preceding century may have been able to mobilize 150,000-200,000 in wartime, by the eighteenth century the key successors and challengers of the Mughal empire across the subcontinent were each raising armies about half this size, with the implication of a very significant overall increase in the scale of military mobilization.

The overall rise in the level of military mobilization in precolonial India can be seen in Figure 5.3, which shows levels of field strength by power, for the period from the late sixteenth century, when the Maratha and Mysore kingdoms formed, to the end of the eigh-
eenth century, when the Mysore empire was defeated. The overall level of military demand expanded moderately during the period of Maratha and Sikh insurgency from 1674 to 1716, notwithstanding a large drop in the Mughal empire’s own military demands; leading to a significant upswing in the late eighteenth century, as the Marathas, Mysore, Nizam of Hyderabad, and the East India Company fought for supremacy over the subcontinent.

The expansion in the scale of arms can also be seen from each of the rising powers of the subcontinent. At the time of the foundation of the Maratha kingdom under Shivaji in the 1660s, for example, its army numbered just 20,000 in total. By the time of his death in 1680, its field strength may have reached 40,000, already “a big army, by any standards of the day” (Gordon 1993). By the 1740s, the Marathas were sending a much larger an army

---

1 Shivaji’s army had initially consisted of a cavalry of 7,000 regular horse, infantry of approximately 10,000, and 3,000 independent troopers.

2 We know from his will that he owned 30,000 horses, suggesting perhaps 15,000-20,000 cavalry, and a total (from Sarkar, House of Shivaji, 188-89, 194).
of 80,000 to Bengal under Bhaskar Pandit (Bolts, 1772), and a similar force in the battle of Panipat of 1761, in which the Marathas fielded a core force of around 70,000 supported by hundreds of thousands of irregulars (Khan 1902). As the state began to decline after Panipat, this was the peak of Maratha mobilization, yet even then, the Marathas were still able to mobilize a still substantial core army of 50,000 during the Anglo-Maratha wars of the 1770s and 1800s.

A similar trend of rising military mobilization occurred in precolonial Mysore. In his early battles Haider had perhaps 23,000 troops under his command (Bowring 1893). However by 1780 Rao (1948: 326) reports that Haider’s army was 100,000 in size, “probably not equalled, and certainly not surpassed, in strength and efficiency, by any Indian army that had ever been assembled in the South of India.” In the battle against the British at Pollicore in 1781, “Coote’s information was that Haider had on the day of action in the field 150,000 men, with 80 pieces of cannon” (Rao 1948, citing Wilson, II. 42). Similarly sized armies are reported under Haider’s successor, Tipu, in the 1780s. For example, Tipu is reported as having an army of 150,000 during the siege of Mangalore in 1783 (Rao 1948: 580).

In other parts of the subcontinent, armies also grew in size and strength. According to records reported by Sita Ram Kohli (1967), the Sikh Empire under Ranjit Singh grew from a force of 12,075 in 1823, to a more substantial force of 41,502 by the end of his rule in 1838. Finally, in its first battle in 1845 against the East India Company, the Sikhs fielded a force of 80,000 soldiers (Hernon 2003). Last but not least, the East India Company itself became the pre-eminent recruiter and mobilizer of arms on the subcontinent, with its army growing from 35,000 in 1770, to 70,000 in 1790, and 153,000 in the early 1800s (Roy 2013a).

The growing scale of warfare is best demonstrated not by considering the growth of numbers, however, but the growth in recruitment intensity. Ever larger armies were being raised on narrow bases. Consider, for example, that the population of Mysore in 1857 was 3.46m; we might assume the core state to have been similar in 1781, while the territories of the kingdom at this time covered perhaps double the territory. If the population of the entire kingdom at this point was 7 million, then an army of 150,000 would have implied 4.5 per cent of males were under arms. For comparison, Napoleon took a peak of 483,000
soldiers on his Russia campaign, at a time when French population was around 30 million, which would have implied only 1.6 per cent of a population mobilized under arms.

A third indicator of rising warfare was not simply the quantitative transformation of Indian armies, but also, their qualitative transformation, as gunpowder weapons became more widespread. The Mughals had long had experience with cannon, an advantage that had allowed them, in common with the other so-called “gunpowder empires” of Safavid Persia and the Ottomans to subdue minor princes and establish a large-scale polity (Streussand 2010). But with the arrival of Europeans bearing light matchlock armaments, the eighteenth century brought a huge transformation in military economy. When Nadir Shah invaded North India in 1739, his forces used relatively rapid-firing, small-bore guns against cavalry for the first time in the subcontinent. At the Battles of Plassey in Bengal and Buxar in Bihar, and the Anglo-French war in the south (1757-63), European military techniques were on full display. Indigenous rulers began to adopt the new technology, setting an arms race for possession of the best weapons.

Indian rulers had invested heavily in gunpowder weaponry in a bid to catch up with European rivals. Initially, these were fleeting; during the first Anglo-Mysore wars, Rao (1948) reports a siege incident in which Haider, during the first siege of Mangalore in 1763, had equipped 20,000 peasants with “wooden muskets of ebony and standards of coloured cloth” to make it appear as if he had a vast infantry support (Rao 1948: 79). By the time of the second Anglo-Mysore war, a decade later, appearances were no longer necessary. Haider had supplemented his cavalry with 70,000 infantry, including 10,000 artillery, 5,000 rocket men and 40,000 irregulars. Of particular note to colonial soldiers was that Mysore was not only an adopter of gunpowder weaponry, but also an adapter. Mysore was unique in using rockets, a gunpowder-propelled “missile” that could be launched towards oncoming armies. Such projectiles, or “Mysorean rockets,” were the first such weapon to be deployed anywhere. English commanders took great interest in the new armament; on the defeat of Mysore in 1799, 600 were taken from the Mysorean capital at Seringapatna and sent to the royal arsenal at Woolwich (Winter 1990). There they were developed by an engineer named Congreve, who in the 1800s rebaptized it as the “Congreve rocket”, which saw its

---

3 Angus Maddison (2008) offers a figure of 31m for 1820.
first appearance in Europe by British armies against Napoleon in 1805, and again against Americans in the War of 1812. Yet - like gunpowder itself - the rocket was invented by an Asiatic power, and brought to Europe for development only later.

The Marathas and Sikhs too, were also investing in light artillery. According to Gordon “there is evidence in the Peshwa’s correspondence that he was aware of his lack of superior artillery; he was certainly aware that Europeans had it” (Gordon 1993: 129-30), and after 1740, “the Marathas hired Europeans - first Portuguese, later mainly French and English - who had expertise in casting and deploying cannon, and built up speciality artillery units” (Gordon 1993: 149). Likewise, feeling the pressure of the Afghans at one end of their domain and the British at the other, the Sikhs were also determined to build a modern gunpowder army. Their artillery numbers increased from 1,046 in 1819 to 5,003 by 1838 (Kohli 1967).

In order to achieve this reorganization of domestic armies to meet the challenges of modern warfare, Europeans were frequently hired to assist in the task of reorganizing armed forces, and adopting both light infantry and heavy artillery weapons and tactics. The most famous such case was in Travancore, where after capturing de Lannoy, the Flemish general of the Dutch East India Company forces, in 1741, agreed a deal: de Lannoy was allowed to serve the rest of his life as a general in the Travancore army, if he would train Indians to use European firearms weapons and tactics. In this capacity de Lannoy enjoyed a remarkable second career, acting for 37 years as a commander for the kingdom that he was once sent to conquer; to this day, his fortifications dot the Kerala coast (Koshy 1989). Similarly the Sikh army under Ranjit Singh brought in outside help as it expanded its artillery in the 1820s and 1830s, transferring the artillery command to a French officer, General Claude Auguste Court, in 1827, and the American Colonel Alexander Gardner in 1832 (Singh 1963). Mysore’s patronage of European officers was also well known: Innes Munro reports that Haider “endeavoured, by every means possible, to allure to his standard military adventurers of all nations and tribes, but particularly the European artificers and sepoys that had been trained up in the Company’s service, to whom he held out the most tempting...

\[4\] Meriting a mention in the “Star Spangled Banner” - “and the rockets’ red glare, the bombs bursting in air”.

169

Were such efforts in vain? It is easy to take such a view in light of Indian regimes’ eventual defeat. Yet it did not seem so at the time, when the military mobilization of India’s indigenous regimes did not appear a forelorn attempt to forestall the inevitable domination by Europeans. Indian armies frequently won their battles against outside powers, and there are quite a number of instances in this regard. The battle of Colachel, between Travancore and the Dutch East India company in 1741, has already been mentioned (Menon 1996). There were many other examples. The French were defeated at the hands of the Marathas at the Battle of Bhalke in 1751, and more significantly, the Portuguese lost all of their coastal territories along the Konkan coast bar settlements at Goa, Mangalore, Daman and Diu following their defeat at Bassein in the Portuguese-Maratha war of 1737-40 (Majumdar 1977). The Portuguese lost again to Mysore, when their settlement at Manglore was sieged and seized by Haider in 1763. In 1779, during the first Anglo-Maratha wars, the British were defeated by the Marathas and forced to cede “every bit of income and territory they had been promised since the death of Madhav Rao in 1773” (Gordon 1993: 164). Mysore also engaged in major battles against the British East India Company, winning its early battles during the second Anglo-Mysore War (1780-84) and securing a treaty recognizing its prewar territorial claims. The fact that the Marathas and Mysore could still notch up battle victories against European powers even at this late stage in imperial consolidation is testimony to the extent of their military modernization in the preceding decades. That Europeans would eventually conquer the subcontinent, then, was not predetermined. The situation of the East India Company certainly seemed precarious to its officers; figures like Clive were considered heroes because of the risks they took, not in spite of them. There is a plausible alternative history in which Mysore and the Marathas, in common with Bhutan, Nepal, Afghanistan and Siam (now Thailand), would have entered the modern era as independent states.

The Consequences of War: Rising Fiscal and Bureaucratic Capacity

A key element of the “fiscal-military” hypothesis is that rising military costs lead to pressure for new revenue sources (Tilly 1992). Similarly, in precolonial India an important
consequence of military intensification, whether as a result of increasing scale of warfare or the cost of gunpowder weaponry, was a significant new mobilization of fiscal resources. Commenting on the shift to gunpowder weapons in the Maratha armies, for example, Gordon (1993) remarks upon how gunpowder weaponry:

“...tremendously increased the cash needs of a commander. Artillery was expensive to produce. Both guns and shot required a full-time and professional foundry. The metals, especially copper, were imported and expensive. The best gunpowder was foreign-made and only acquired from the Portuguese or the English for cash. Soldiers who operated artillery were specialized, relatively well-paid professionals. Even the support personnel, drovers and such, had to be paid in cash”.

Such cash had to be found. One solution was to reallocate among existing expenditures, and Roy (2013b) estimates that in late eighteenth century India, seventy to eighty per cent of all money raised in taxation was eventually dedicated to military outlays. Yet while possible as a temporary expediency, there were limits to this course of action, and sooner or later more revenues were required. Most states were “fully occupied with the task of suppressing rivals and rebels and raising the armies and revenues necessary for this end” (Gwynn, 1932). A solution that is often first deployed by patrimonial states is meet short-term fiscal needs through measures such as an increase in rates, the sale of prebends, assumption of debt, and impositions of ad hoc indirect taxes, and there is much evidence of this in eighteenth century India. However, these can only be short-term expediencies, and when the fiscal pressure is large and continual enough in duration, only better auditing of economic resources and control over agents through bureaucratic reforms can ensure a steady supply of material resources (Ertman 1997; Wolfe 1972). States undergoing increased military pressure tend to respond in two phases: an initial period characterized by rising rates; and a second period marked by attempts at better auditing, surveillance, and sovereign control over the agents responsible for collecting and remitting revenues.

We see this most clearly among the southern challenger states, such as the Wodeyars of Mysore or Marathas in the Deccan, who began their new dynasty by increasing the intensity of fiscal extraction so as to raise new funds for war. In Mysore king Wodeyar (r.
1672-1704) increased the general revenue demand upon state officials, to such an extent that it led to “an uprising of the Lingayat peasantry against increased taxes collected by the state and centralization”\(^5\). Temporary at first, the pressures for revenue mobilization did not abate: “Between 1600 and 1800... the old Hindu state of Mysore progressively upped its nominal tax revenue from under 10 per cent of the gross produce to about 40 per cent under Tipu Sultan in the 1790s,” as “the growing costs of warfare and a desire for a new grandiose form of kingship spread across the subcontinent” (Bayly 1988: 25). Likewise the founder of the Maratha Confederacy, Shivaji, financed his incursions against the Mughals by installing a tough land revenue system, which demanded that peasants give over half their produce to government storehouses; and in order to enhance this amount further, his government provided rural credit to enhance production (Kulke and Rothermund 1986: 194). Under growing pressures of warfare against Mughals, Afghans, and the British, this fiscal extraction continued apace in the eighteenth century. “In the 1750s there was a sharp increase in the state’s land-revenue claim and a rationalisation of management. Again in the 1790s and 1810s the desire of the peshwas for increased revenue to pay for their armies... caused an expansion of revenue farming and an increase in the state’s demands” (Bayly 1988: 100).

This can be seen by examining the progression of the revenues of the Mysore kingdom and empire from its founding in the seventeenth century to its final defeat in the fourth Anglo-Mysore war of 1799, whereafter the Mysore kingdom persisted as a rump princely state under the British Empire (Figure 5.4). While revenues grew steadily from the founding until Haider’s rule in the 1760s, it is clear that during the period of warfare in the late eighteenth century – a time when Mysore was not only at war with the East India Company, but also the Nizam of Hyderabad, the Marathas, and Travancore - fiscal mobilization increased substantially. Even taking account of the increase and loss of territory from 1761-1799, taxes rose considerably over the short period of time that Haider and Tipu’s regime existed - roughly doubling within the core state of Mysore control before and after this period. And despite evidence of a decline in the early nineteenth century, there appears to have been a strong “ratchet effect” whereby this increase was not reversed.

---

It is also clear from the level of taxation raised by the Maratha peshwa (chief minister), whose territories were even more highly assessed, on a per capita basis, than those of the Mysore empire. In the late eighteenth century, the annual level of taxation per capita stood at 0.8 sterling in the core Maratha state, against 0.5 in the Mysore empire (Table 5.1), and much lower levels in the other South Asian polities that existed at the time, such as the northern Mughal successor states of Bengal, Awadh, and Hyderabad.

Raising revenue, of course, is not a switch that elites can simply flip at will. It requires a first bureaucratic infrastructure to assess and collect, and a second bureaucratic structure, to enforce compliance among the state’s own revenue agents in the fulfillment of these duties, and ensure that funds are not defalcated or diverted. How did states go about securing the additional revenue that they required? In order to understand how new regimes reorganized the fiscal apparatus, it is necessary to begin from the structure of taxation in medieval India. As explored in chapter 3, until the eighteenth century, the most common form of fiscal organization in South Asia was the “concessionary” state, in which grants of
<table>
<thead>
<tr>
<th>Polity</th>
<th>Revenue (£m)</th>
<th>Population (m)</th>
<th>Tax per Capita (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Centralized States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maratha Peshwa</td>
<td>4.0</td>
<td>5.0</td>
<td>0.80</td>
</tr>
<tr>
<td>Mysore Empire</td>
<td>4.0</td>
<td>8.0</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Concessionary States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travancore</td>
<td>0.25</td>
<td>0.9</td>
<td>0.27</td>
</tr>
<tr>
<td><strong>Feudatory States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bengal</td>
<td>5.8</td>
<td>40</td>
<td>0.15</td>
</tr>
<tr>
<td>Hyderabad</td>
<td>1.5</td>
<td>10.7</td>
<td>0.14</td>
</tr>
<tr>
<td>Mewar</td>
<td>0.14</td>
<td>1</td>
<td>0.14</td>
</tr>
<tr>
<td>Awadh</td>
<td>0.79</td>
<td>6.5</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Tributary States</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berar</td>
<td>0.49</td>
<td>4.6</td>
<td>0.11</td>
</tr>
<tr>
<td>Gwalior</td>
<td>0.32</td>
<td>3.2</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 5.1: Revenue and Taxation per capita in Precolonial India. Taxation figures refer to dates from the period 1770-1820. Data on levels of taxation from the *Moral and Material Progress of India*; Irfan Habib, *The Agrarian System of Mughal India 1556-1707*; James Mill, *A History of British India 1805-1835*; Tirthankar Roy, “Rethinking the Origins of British India”; the *Karnataka State Gazetteer*. 
land (termed *inam* or *patti*) were made to soldiers and generals who served in the army, or otherwise perform functions useful for the sovereign, and tributes were paid by chieftains and warlords in order to avoid punishment. In this model, the king was acknowledged as the ultimate rights-giver over the land, and allowed certain groups to avoid tax payment in exchange for service. In a centralized concessionary arrangement, the sovereign could effectively assign and reassign concessions of land, whereas in more feudatory regimes, the sovereign’s power to do so was merely nominal and rights-holders held assignments that were effectively hereditary. Meanwhile, under a tributary arrangement, chiefs offer fealty to the sovereign, whose sovereignty is not recognized as absolute. By the early modern era, most Indian polities engaged in a mixture of tributary and concessionary forms; the Rajputs of Rajasthan or the poligars of the South are examples of mixed systems in this regard (Dirks 1991). While tributary and concessional forms are efficient from the point of view of requiring a low degree of enforcement and information from the sovereign, they are inefficient as a means of revenue mobilization, as they leave most of the surplus from taxation of the land in the hands of intermediaries, rather than the state. Modernizing regimes, therefore, had to implement an alternative. The Mughal empire and its successors (such as Awadh, Hyderabad, Agra and Bengal) had initially replicated the ‘concessional’ model, allowing very large land grants (*jaghir*) to generals who were treated as military nobles, who then farmed out tax collection to a middle strata of zamindars. Under the Mughals, such a system was centralized by ensuring that the allocation of *jaghirs* corresponded to military rank and grade, though as explored in greater detail later in this chapter, the system steadily broke down as land grants increasingly became hereditary entitlements.

In order to improve their extractive potential, southern rulers, by contrast, developed new forms of fiscal organization. Presaging the reforms of Read and Munro in the nineteenth century, this required cutting out the middle-strata of tax collector and implementing a new method of direct taxation, known as *ryotwari*, or individual-cultivator based assessment. The ryotwari system was predicated on direct collection from millions of individual small landholders, called in Persian *ryot*, meaning “subject” or “cultivator.” The “ryot” under this system paid to the state a share of his crop calculated from the productive capacity of the fields that he engaged to cultivate each year (Stein 1996: 209). Often in practice, the landholder might be a village rather than an individual tenant, but nonetheless Maratha of-
ficers were able to assess directly, rather than operate via a long chain of quasi-autonomous intermediaries:

“In their search for revenue the regional and petty states almost always tried to gain a closer control over the headmen or their village proprietors. Sometimes direct representatives of the state acquired revenue-collecting rights at the village level and eroded the power and perquisites of the headmen families. This was particularly true where the state was strong as in the Maratha territories around Poona between 1751 and 1818, or in Mysore where Haider Ali and Tipu Sultan tried to eliminate all intermediaries between themselves and the peasant farmers, subjecting them to new demands for the punctual payment of revenue. Elsewhere the absence of powerful authorities outside the village enhanced the authority of the elites as protectors and petty rulers of the villages” (Bayly 1988: 28).

Such a move was facilitated by the fact that these “challenger” states no longer depended upon the old political elites, and were antithetical to those who worked with the Mughals and their predecessors. “The new rulers of Hindu states were less often scions the medieval ruling Rajput clans of kshatriya lineage, and more often of dominant regional peasant stock, such as the Maratha Kunbis or the Sikh Jats” (Stein 1996: 197). The Marathas were part of a more general peasant revolt against Mughal rule in the late seventeenth and eighteenth centuries; yet the new model implemented by such dynasties was also aided by the fact that, in the dry uplands of the Deccan for example, settler agriculture was still common: enterprising settlers were encouraged to cultivate new lands and establish themselves as the deshmukh or village chief, and these new settlements could then be organized in a direct relation with the state, rather than via intermediaries. In this way, the “socio-economic underpinnings of ‘the state’ in South Asia... were changing as state-building processes moved towards the creation of more unitary and centralized orders which displaced the segmented political formations of the past” (Stein 1996: 198).

6 The degree to which the early Marathas were motivated by religious, and specifically, anti-Mughal or anti-Islamic imperatives remains a source of intense controversy. The founder of the Maratha polity, Shivaji (r. 1674-1680), had deliberately cast himself against the Mughal ancien regime, by crowning himself in a Hindu coronation, claiming to fight for Hindus against Muslim rule, and mobilizing support directly from Marathi chieftains. However, it is also true that the Marathas frequently allied with Muslim states, and included significant numbers of Muslims in their own armies.
In the pursuit of this new model, the Marathas for example appointed a figure called a *kamasvidar*, who had the right to collect revenue and had to maintain a military garrison, alongside whom a small staff of brahmin clerks were employed to maintain accurate records directly for the government. Tax-farming contracts were auctioned annually after the revenue was first estimated by the peshwa’s servants; in this way, the margin for extractive gains were narrow. Such a localized system of assessment was essential in legitimizing the peshwa’s move to marginalize local zamindars. “‘Regimes like the peshwas’ look distinctively modern in comparison with the Mughals’”, comments Burton Stein; as “bureaucratic management began to be as important to states as military and charismatic lordship. Accurate record-keeping had been introduced in numerous local settings and institutions, but not until the eighteenth century did the principle find expression at the apex of a political order, thanks to brahman managers who constructed a state form that matched the challenges of the age” (Stein 1996: 191).

A yet higher level of fiscal centralization was achieved by the Mysore Kingdom, towards the last of its many wars with the East India Company. Wide-ranging reforms were implemented with the goal of enhancing fiscal extraction; as the British administrator Munro noted, “the whole course of [Haider’s] administration was a series of experiments to discover the utmost to which land revenue could be raised without diminishing cultivation”\(^7\). Rather than collect tribute and offer concessions for military support, under the Mysore Kingdom the entire realm was divided into districts called Subahs; and the local chief executive officer was referred to as the Amildar or Amil. This system of governance had dated from the founding the Mysore regime the previous century, when the local chiefs had been executed by Chikka Deva, who had instead founded a centralized government with 18 ministries. In the late precolonial period, however, this system was supplemented and superintended to an unprecedented degree so as to ensure maximal extraction from the peasantry and minimal subversion of revenues.

The key elements of Mysore fiscal policy under Haider were as follows. A first aspect was the cancellation of concessions, which previous regimes had offered to members of the nobility, temples, and families of those in the royal service. In the northernmost province

\(^7\)Quoted in the Uttara Kannada District Gazetteer, 1985: 160.
of Bellary, for example, Haider Ali and Tipu Sultan “increased the revenue by a large resumption of inams” (tax-free concessions), forcing all landholders to pay tax\(^8\). Under Haider, the province was due to pay 1,977,776 gold pagodas, as “[Haider] Ali endeavoured to augment the revenue by the resumption [cancellation] of inams and rusums and in some instances by the conversion of peshkash [tribute, paid by poligar chiefs] into rent. The resumption of inams and rusums added to the revenue from 5 to 10 per cent... the revenue continued to increase from 1779 to 1789,” when the demand for Bellary was estimated at 2,277,998 gold pagodas, a 15.2 per cent increase (Karnataka Gazetteer 1982 vol. II: 207). In doing so, Haider’s state frequently encountered opposition from the older Deccani nobility, members of which attempted, unsuccessfully, to assassinate him in 1779.

Second, the Mysorean state ruthlessly destroyed the tribal chieftains, or poligars, wherever they resisted state authority. The number of chiefs subject to this treatment was extensive, such that Mysore “rid Karnataka of many warring minor principalities,” “putting an end to many small principalities like Doddaballapur, Chikballapur, Sira, Harapanaballi, Ballery, Ramdurga and others”\(^9\). The degree to which Mysore had eliminated the chiefly strata is relayed in an account by the Persian chronicler Kirmani, who described how Mysore had pushed into exile “the Poligars of the Balaghaut... the Poligar of Gungoodi Pala (Kangundi Koppam); the sons of Bhyreh Koor, the poligar of Chuk Balapoor; Pud Nair, the Poligar of Vinkut Giri Kote, who was residing at Charkul; Shunk Rayel, or Rawul, the Chief of Punganoor, and besides these, the Poligars of Khut Koomnir, Mudumpalli, Anikul, Oonkus Giri, Cheel Naik, etc. all being dispossessed of lands” (Kirmani, quoted in Wilks 1817).

Local tales of political displacement can be found across the territories conquered by the Mysore regime. In Dakshina Kannada, after 1763 when the area came under the control of Haider Ali, “a general investigation into revenue policy was ordered. An order was issued repudiating all claims of waste lands and imposing a full fifty per cent addition to the 1618 fixation”\(^10\). The dispossession of landlords was so complete that during the 1947

\(^8\)Imperial Gazetteer 1908, Madras, p. 443.

\(^9\)Karnataka Gazetteer: 323.

\(^10\)Karnataka Gazetteer 1982-3 vol. II: 306.
Karnataka land reforms, it was noted that ‘there was no zamindar estate or inams in Dakshina Kannada district’ because ‘all small inams had been resumed [abolished] by Haider Ali and Tipu Sultan’\(^{11}\). Finally, similar evidence is reported in the gazetteers of Ramdurg, a minor princely state of northern Karnataka: “About 1753 the estates yielded 2 1/2 lakhs [Rs. 250,000],” but in 1778, when the country was brought under the subjection by Haider Ali, this level was increased; and in 1784, Tipu Sultan made further demands\(^{12}\). On assuming control of South Canara, Tipu is reported to have “completely suppressed all of the old local chiefs and dispossessed them of all but their private lands.” One such dispossessed poligar, the Kumula raja, had initially fled and then unwisely returned to try and reclaim his rights. He was “captured and hanged,” along with a brother and nephew for good measure. Two further princes of the region (of Nileshwara) were similarly hanged by Tipu in 1787\(^{13}\).

This is also true for the districts situated outside of Karnataka, and that now lie in other states. In Anantpur, whereas under the Mughals “cultivation had undoubtedly decreased owing to the depredations of the poligars... Haider Ali fixed the assessments, and endeavoured to increase them by the resumption of inams and other grants” while “Tipu Sultan made further efforts to increase them” (Madras Imperial Gazetteer, Anantpur section, p. 481). Meanwhile in Malabar, in northern Kerala, “unlike other Districts of Southern India, the Hindu rulers appear to have levied no regular land revenue, but to have contented themselves with customs and tolls and with the occasional levy of special contributions... Haider Ali, on his conquest of the District at the end of the eighteenth century, proceeded to introduce a regular system of land revenue... the share appears to have varied from 10 per cent on some ‘wet’ lands in North Malabar to 100 per cent on gardens in South Malabar” (Madras Imperial Gazetteer, Malabar section). The imposition of taxation in Malabar could not be completed without encountering local resistance, however; a rebellion of Nairs was put down in 1768 and settled by treaty; in 1773, a more determined effort was made when 40,000 troops were dispatched to Malabar following the killing of Haider’s tax-collectors (harikars). Following this, the Nairs were “reduced to a state of dependency” (Rao 1948:

\(^{11}\)Karnataka Gazetteer 1982-3 vol II: 315.
\(^{12}\)Imperial Gazetteer 1908, Bombay: 534.
\(^{13}\)South Kanara Gazetteer, 1973: 64.
Bureaucratic rationalization

The contenders for regional control in post-Mughal India were therefore in the business, not only of waging wars, but of raising taxes to pay for them. However, what made the new states successful in fiscal-military expansionism were not their policies, but rather, the new administrative institutions which allowed for an increasing flow of resources to the center. In the Maratha core domains governed by the Maratha peshwa, rising fiscal mobilization required improvements to the ability of the state to measure and assess taxable resources, using a new category of civil servants and a documentary system to keep track of land cadastres and rights. As Stein (1991) remarks,

Record-keeping under the peshwas exceeded any previously known in India, judging from their daftars, which were ledgers of correspondence and account books. These have provided a rich resource for modern historians as well as a model for local administration, imitated by British India in the next century. The Maratha regime at its zenith was headed by literate brahmans whose caste practices were not priestly, but secular, working in the world of politics and commerce (Stein 1996: 191).

As an insurrectionary Hindu state, high-caste and literate brahmins were mobilized from sacral to secular service in a manner unprecedented in the history of the Deccan. “Their scribal abilities were at a premium as conquests were followed by the establishment of civil rule” (Stein 1996: 189). The use of literate brahmins for political ends enabled a revolution in fiscal management, as “financial knowledge and institutions were mobilized to realize the prompt transmission of tribute from an increasingly extensive empire, and [the peshwa] Bajirao adopted the policy of centralizing all fiscal functions in Pune by 1740” (Stein 1996: 190). In his study of state formation Fukuyama (2011) observes that the first genuinely Weberian bureaucracies in human societies are often not in the form of states, but rather in the form of organized religion; and that early states often undergo bureaucratic development by mobilizing religious clerics into their service. This was also the case in South India in the eighteenth century, when literate brahmins were mobilized
out of their traditional priestly roles into a then less conventional form of political service: this occurred in the Maratha domains, in Travancore, and in Mysore. Though of lesser significance for the political development of South Asia as a whole, it was also true for the Himalayan states, such as Sikkim and Bhutan, where “lamaist theocracies” adopted the Tibetan model of yoking Buddhist lamas to the service of a royal monarchy, by collecting taxes, organizing the labour corvee, and functioning as a bureaucracy united by a strong esprit de corps (Sinha 1966). By consequence, these mountain states, sitting on the very fringe of the Indian subcontinent, were also able to achieve a level of political order by the early modern that was enviable by comparison: Sikkim and Bhutan for example conducted a first census in 1747, and as Ardussi and Ura (2000) note the “list of 1,149 functionaries seems surprisingly large for the presumed population size” of just 27,223 households.

A key criterion of administrative rationalization, of increasing importance from the sixteenth century on, was the extent to which writing and print were used as tools of bureaucratic management. This is, of course, why we know so much more about the polities of eighteenth century India: written copies of promulgations by Tipu Sultan found in Mysore provinces, for example, or auction records of tax contracts in the the Maratha domain (Stein 1984, 1996; Gordon 1966). These tax documents, or dehazada, listed every village, recorded the name of every farmer, every fraction of land, and both the summer and the winter crops grown (Gordon 1966; Sen 1976). By contrast, in polities that lacked administrative rationalization, the only methods available are the accounts of European voyagers, and diplomats or reconstruction of ancient political institutions via oral history (Vansina 1985; Dirks 1987). The details of administrative protocol uncovered by such enquiries unsurprisingly reveal systems of rule that were largely patrimonial and without clear legal or accounting institutions (Dirks 1987).

For Gordon (1993), the “long-term trends in information gathering and revenue collection represent one of the major accomplishments of the Maratha polity”.

“From the 1720s, there was a steady expansion of area under direct administration, largely at the expense of non-assessed lands under armed, local lineages, but also seized from opposing powers, such as the Nizam. The trend was toward greater quantity, quality, and frequency of information reaching the centre. In these records, we can track the trend toward predictability, as
areas acquired a reputation for payment or recalcitrance... We can track, for example, cropping patterns, cities and trade routes, and the minute details of caste and family conflict. In addition to revenue records, the peshwa regularly received news from all parts of India, produced by professional newswriters at all the principal courts. In a wider sense, the revenue collectors and newswriters have given us an economic and social record for the eighteenth century unmatched in detail and scope outside of Europe.” (Gordon 1993: 188).

In areas that were directly under the peshwa’s rule, the administrative calendar followed a strict cycle. First, the peshwa’s clerks prepared an estimate of revenue. Once the collector accepted the estimate, he raised an advance payment from the financial community, and before the monsoon season, set off for the field. They then toured across the land, stopping at villages for a day or two to “examine the local land records and note any changes in the actual fields under cultivation”, before preparing “documents confirming the taxes due” and signing these with village headmen (Gordon 1993). A balance sheet was prepared at the end of the year, and if a shortfall should accrue, then this would be assigned for collection in the following year. As Gordon (1993) notes, “lest this be construed as some sort of idealized system, recall that there are tens of thousands of working documents of this system in the Pune Daftar, and every researcher who has worked on them has found extraordinary levels of thoroughness, predictability, and detail.”

Turning to Mysore, the regime of Haider and Tipu in the second half of the eighteenth century also underwent extensive efforts to centralize authority, with the aim of establishing “unbridled autocracy of the Cromwellian type, dominated by the civil and the military elements and entirely subservient to [Haider’s] personal will” (Rao 1948: 285). Many aspects of the Mysore polity, including the separation between civil and military spheres, and within the former, the separation of government ministries and roles, had in fact dated back to the seventeenth century, when king Chikka Deva organized the state into 18 offices or departments. These had provided for an efficient operating bureaucracy, and visiting the capital in 1779, the German Lutheran missionary Christian Schwartz observed how “on both sides [of Haider’s palace] are ranges of open buildings, where the military and civil servants have their offices, and constantly attend” and that “Haider can overlook them from his balcony... Here reigns no pomp, but the utmost regularity and despatch” (from Wilks...
Together with bureaucratization, another key element of the fiscal state is enforcement. The regime of Mysore in its final decades was the clearest example in this regard, constructing an elaborate police state that was targeted primarily at public servants so as to ensure the maximal flow of revenues to the treasury. In addition to constructing a state apparatus of tax administration, therefore, Haider superimposed a system of surveillance. This occurred as a direct outgrowth of the fiscal needs of the state. In the first decade and a half of Haider’s rule, which began in 1761, there were a large number of military engagements, including the seizure of Mangalore from Portugal, and a lengthy war against the Marathas, that came to an end in 1772. Despite increasing the revenue demand, this was, however, insufficient to cover the mounting cost of military expenditures, forcing Haider to devise a new solution. This he did on return to his capital in 1776, when he summoned the whole of his tax collectors to the capital in person, “for the purpose of adjusting their past accounts and future revenues, and levied upon the whole country a forced contribution under the name of free gift (nazarana) for the support of the war” (Rao 1948: 290).

Knowing that the demand was unlikely to be fulfilled without very strong incentives for compliance, Haider simultaneously sought to construct a parallel system of surveillance and punishment to supplement the fiscal administration, by the appointment of harikars (police officers) to each district. The History of Hyder Naik, by Kirmani, tells of how “In all the cities and towns of his territory, besides news-writers, he appointed separately secret writers and spies to patrol the streets at night, and from them he received his intelligence.” (Mysore Gazetteer: 2541-2542). The development of this system was entrusted to Haider’s postmaster, Shamaiya, a brahim from the Kolar district. Under Shamaiya’s charge, Mysore developed a system of control that was unambiguously despotic, yet highly effective in reducing defalcations and embezzlement, beginning what Rao (1948) terms “a reign of terror”.

Colonel Fullarton, who led British troops during the second Anglo-Mysore war, also describes the effectiveness of the Mysorean secret police in his account of late eighteenth century politics. According to Fullarton,

“The minutest circumstances of detail, the produce of a crop, the cultivation
of a district, the portion paid to the circar, and that reserved to the inhabitants, were accurately known to him. Not a movement in the remotest corner could escape him, nor a murmur or intention of his neighbours but flew to him. It will hardly appear exaggeration to say that he was acquainted with every person of his empire, when we consider that he was in a continued round of inspections. In his Durbar, during the hours of business, reports from all corners were received, his secretaries successively read to him the whole correspondence of the day” (Fullarton 1787).

The primary purpose of the secret police was to ensure fiscal compliance. Haider made sure that harikars were assigned to investigate suspicions of abuse of office. Two were assigned to each taluk (local district), and their duty was to hear all complaints, and to report these to the office of the Revenue Department at the head-quarters of Government. They also reported all waste lands within their areas. Meanwhile, an elaborate typology of punishments was applied to officials suspected of expropriation, embezzlement, or disobedience. The effect of these policies was a marked increase in fiscal resource mobilization, for “new and horrible contrivances of torture, spread a terror, which sometimes rendered their application unnecessary” (Rao 1948: 292).

The same mechanisms were used to secure civil justice; with respect to highway policing, for example, it was reported that “if, by accident, any highway robbery etc. was committed, the Kavalgar (local police guard) was impaled without delay and another promptly appointed to his place” (Rao 1948: 531). Mirza Ikbal, in his Ahwali Hyder Naik, wrote how in Mysore “mankind were held in fear and trembling... Cutting off the nose and ears of any person in his [Haider’s] territories was the commonest thing imaginable, and the killing of a man there was thought no more of than the treading on an ant”. Other reports by contemporary visitors to the realm frequently noted the high degree of public order and that, while “official men had cause to tremble,” “the mass of the population felt that the vigour of the Government... rendered their condition comparatively safe” (Wilks 1817). In precolonial Mysore “the internal governance was safe and secure, that the country’s peace and order were maintained, that disturbers of peace were not allowed any sway, that the roads were safe for traffic and trade,” and “public order was maintained vigorously” (Rao 1948: 444-5).
Over the course of the eighteenth century, therefore, challenger regimes to the Mughal empire, led by Mysore and the Marathas, were engaged in defensive modernization in the Indian subcontinent. Such policies were brought on by the intensive fiscal demand that arose from continual interstate warfare, as well as by the rising costs of war due to changes in battlefield arms and technology. Polities such as precolonial Mysore and the Marathas were not bureaucratic states in the same sense as would be understood today, and perhaps, not even by comparison with contemporaneous European examples. Yet they clearly constituted a degree of administrative centralization that was unprecedented in the Indian subcontinent. One telling indication of this fact is what occurred upon the death of the Mysorean executive Haider, from cancer in 1782.

“The whole of the arrangement of the army, the weekly relief of the 2,000 horse which constantly hung around Madras, the issue of pay, the adjustment of military accounts, the answers to letters received from the envoys of different courts, and all the business of the state, went on as usual. The principal officers of the army and foreign envoys were separately and quietly taken into confidence, and all inquiries were answered to the effect that Haider was better but weak” (Wilks 1817: II, 168-69).

Such a response - the continuation of political institutions to function as before, in the absence of the leader - is characteristic of bureaucratic states such as imperial China or the Ottoman court, where a permanent bureaucracy handled matters of state with or without the presence of the sovereign. Yet until the late eighteenth century, it would have been unthinkable in India for a regime to continue functioning without its head of state.

The Origins of State Weakness in the North

If the eighteenth century was an era of state formation in India’s south, it was an age of collapse across the north. If we consider India’s precolonial institutions on a continuum from bureaucracy to patrimonialism, it is clear that post-Mughal regimes saw a substantial shift towards “repatrimonialization” following the demise of the Mughal polity (Bandhapadhay 2004). Under the Mughal empire, the mansabdhari system did at least give the emperor power to resize and reassign jagirs; religious and judicial roles, moreover, did
follow certain basic meritocratic norms. The occupancy of public offices such as the *kazi* (Islamic judge and notary), *kotwal* (police officer), and *muhtasib* (supervisor of bazaars) was based on merit and religious observance. “The author of the Riyaz recalled that in the reigns of Aurangzeb and in the Nizamat of Jafar Khan, only the nobility, the scholars, the learned and the excellent, who passed the examinations were appointed to the post of kazi. The office was never bestowed on ‘the low and the illiterate’” (Riyazu-i-Salatin, p. 284, quoted in Singha 1998: 18).

By the mid-eighteenth century, this was no longer the case across much of the former Mughal empire. Declining tax revenues and the rising cost of warfare meant that steadily smaller endowments were provided by rulers to their kazis and muftis; by consequence “uncertainty of resources from the centre” had “impelled Mughal-appointed kazis and muftis to enhance their fees” (Singha 1998: 17). These offices were steadily abused, and the compiler of the Mirat complained of “illegal exactions since the early eighteenth century,” that is, since Mughal power began to wane following the death of Aurangzeb (Mirat p. 343). Gholam Hussain, for example, highlighted “the pressure put on the poor for the payment of fees at a death, circumcision or marriage” (Seir, 1783 edition: 160, 166). As fees increased, these offices became substantial sources of revenue in their own right, and these positions attained an economic value that could be inherited or traded; outright prebendalism (the sale of office) became the norm as “statuses and office were leased and sub-leased” (Bayly 1988: 11). This in turn brought ‘bureaucratic entrepreneurs’ into the post, of limited training and vocation, who saw their station as a means to extract gains, if only to recoup their initial investment. The kazi could withhold the performance of religious offices for non-payment of fees; the sale of this office had brought in officials of limited religious inclination. A similar situation prevailed among justices: when the Sikh chief Nihal Singh told one Sukh Dial to administer justice with mercy and religious honesty, the latter, who had bid thirteen lakhs for this office, “folder his hands and said justice by contract made this difficult”14. Increasingly, administration of justice was not conducted by Muslim scholars at all, but instead “local zamindars and Mughal satraps,” who “positioned themselves as the

14From Newsletter of 18th September, 1813, in Garret and Chopra (eds.), *Events at the court of Ranjit Singh*, 1810-17, 1935, reprint 1970, p. 95.
dispensers of ‘justice and protection’, levying fees in this capacity” (Singha 1998: 17). The East India Company similarly reported that “the farmers of revenue and the zamindars had usurped the administration of justice from Mughal offices” (Singha 1998: 18). The reason for this usurpation was partly due to the failure of Islamic justices to do so once their grants were discontinued, but also for the substantial economic gains to be had from the abuse of such positions.

By the time the British extended their rule across northern India, fees attended virtually every judicial outcome. Fines were scaled to the income of the offender, so as to better generate revenue. “Heinous crimes such as murder, Company officials complained, were compromised by the ‘purchase’ of a pardon” (Singha 1998: 22). And in order to have the government reclaim a debt, fully one-quarter of the amount was taken as “commission” by the ruler: Robert Orme remarking in 1752, before the extension of Company rule across India, that “this is so avowed a practice, that if a stranger should enquire, how much it would cost him to recover a just debt from a creditor who evaded a punishment, he would everywhere receive the same answer – the government will keep one-fourth, and give you the rest” (Orme, Of the government and people of Indostan: 31).

In the shadow of Mughal decline, the few bureaucratic and meritocratic norms that were respected in the Mughal order faded and were succeeded by patrimonialism and prebendalism (Bandhapadhay 2004). By contrast, in the south of India, legal and administrative reform were tightening control over the populus. “Stewart Gordon suggests that by 1740-50 a marked differentiation had taken place in the [Maratha] Peshwa’s bureaucracy and the Peshwa’s faujdars and kotwals were producing cases on every aspect”15; Gune suggests that “from the 1730s or so the Peshwa’s government was using its administrative personnel to bring hereditary village officials under tighter control” (Gune, The judicial system of the Marathas: 126). The state failure observed by Orme was not generalized across India, but specific to those areas of the former Mughal Empire where political authority had declined, and offices of state had become subject to prebendalism.

**Fiscal Erosion in Precolonial North India**

Just as the institutions of justice, regulation and policing saw as gradual “repatrimonialization” in northern India of the eighteenth century, tax officials - who, in any event, had often usurped the former responsibilities - also assumed greater control over the collection and enjoyment of revenues. As a result, there is evidence that revenue collection declined across former Mughal realms of the eighteenth century. Whereas in the sixteenth century overall state income had more than doubled (Robb 2002: 93), ‘from the mid-eighteenth century Mughal coffers were virtually empty, indicating the seepage of power to the localities’ (Stein 1996: 190). Yet notwithstanding the obvious collapse in revenues paid to the Great Mughal in Delhi, even in the Mughal successor states revenue collection appears to have declined. While Bayly reports that the “volumes of collection in the Mughal empire’s northern heartlands do not appear to have fallen below 65 per cent of the nominal demand of the later Mughals” (Bayly 1988: 36); nonetheless a one-third shortfall in revenue collection is striking considering that southern challenger states such as Mysore of the Marathas were doubling or tripling revenue mobilization in these years.

Underlying this fiscal decline was a steady power grab by local officials, including those responsible for collection of taxation. “In Awadh and Bengal clients and family members of the new ruling houses were also able to amass large bundles of proprietary rights and rights to farm revenue from the state which in course of time became hereditary estates (zamindaris in Bengal; talukdaris in Awadh and Hyderabad)” (Bayly 1988: 21). By consequence, in Awadh, “about two dozen great magnates accounted for most of the state’s revenue and much of its military strength” (Bayly 1988: 91). Similarly, “In the later eighteenth century in some parts of Hyderabad such grants did, however, tend to become hereditary. This created a more settled landowning class which negotiated its revenues and military commitments through agents settled in Hyderabad city” (Bayly 1988: 20).

There is a large literature on precolonial processes of gentrification, and in the century before the permanent settlements of the East India Company, a proto-‘gentry’ had already formed across northern India whose land rights approximated a form of property, in nature if not in name. Strictly speaking, the conversion of tax fiefs into ‘property’ required three institutional changes; first the ability of zamindars to refuse to turn over some portion of their revenues to the state; second, the assumption of a right of heritability; third, the ability to exchange ownership ‘rights’ privately. Evidence of all three is common across northern
India in the precolonial period: Thus “zamindars (landholders) began to ... seize prebendal lands which the Mughal elites had once tried to keep out of their hands. Families of servants of the Mughals and relatives of the old nobility bought up proprietorial rights over land or quietly converted non-hereditary into hereditary rights” (Bayly 1988: 10). Furthermore, “the proprietary rights of landholders and headmen were more regularly exchanged by sale and mortgage” (Bayly 1988: 11).

A consequence of gentrification was that some of the responsibilities which formerly had been the domain of the central state, such as policing and public order, fell upon the shoulders of these landlords. “Fiscal weakness at the top led to extensive revenue farming and dependence on remnants of the former nobility, merchants, bankers and military families long settled on land but who still carried arms. Individual members of this mixed club became known either as zamindars or by other names” (Roy 2013b: 54). In some ways, the stability of the landlord class following Mughal decline may have increased the incentive to invest in public provision, and formed more durable bonds of patronage between the landowner and the communities over which he exercised revenue rights.

_Neglect of Public Goods in Precolonial North India_

One indicator of the degree to which political order collapsed across northern India in the eighteenth century is the evidence that earlier investments in public goods went into severe neglect. Contemporary travel accounts and colonial reports indicate that, together with the collapse in central state revenues, public investments in roads, canals, irrigation, and public safety withered. Roy (2013b) for example notes how “[a]round 1800, archival documents on public works and travelogues report discoveries of virtually abandoned large-scale river embankments near Patna on the Ganges, Cuttack on the Mahanadi, Sylhet on the Surma and Murshidabad on the Ganges, and disused fourteenth-century canals in the western Gangetic plains”, while “[n]orth Indian canals had been unused for decades, even a century; roads in the central uplands were unsafe not least because of the constant demands for tolls made by the local chiefs; no-one knew who repaired the river and seafront embankments in Bengal, and the urban infrastructure of the major cities were in a state of collapse” (Roy 2013b: 44).

With the withering of the Mughal mechanism of extracting and distributing patronage,
there were simply a great deal less resources to expend. No measure is more powerfully indicative of the decline of the Mughal client-state than the depopulation of its major urban centers, with Delhi, Agra, and Lahore each falling in population from approximately 400,000 to just 100,000 over the course of the eighteenth century, and Benares, on the eastern Ganges, falling from 582,000 to just 175,118 from this point until the time of abolition of the last Mughal emperor in the mid-nineteenth (Roy 2013b: 108). Yet it also reflected a breakdown in political order, with the responsibilities assumed by the central state no longer finding any responsible party further down the hierarchy with the capability to reprise its functions. In this way, the “decay was the price paid for the state losing control over money and the military machine at the same time” and “the public goods were left, by default, in charge of the communities. Law and security were provided by groups of peasants for themselves or groups of merchants for the members of their own clubs” (Roy 2013b: 44-5).

While there is an important debate to be had concerning the degree to which, in a patrimonial and predatory regime such as that of the Mughals, any layer of authority had ever been capable of investing in public goods, the fact that canals had existed is evidence that certain aspects of infrastructure provision were within its capacity, particularly those relating to transportation and trade. The observation of neglected roads and canals is evidence that during the high phase of the *pax Mughala* highways and rivers had once been navigable, kotwals policed the highways, and economic output may have expanded. Yet even during this halcyon period, there were contrary signs. Francois Bernier, a French visitor to the empire at the peak of its power in the 1660s, remarked upon how even at this time, the insecurity of the Mughal empire’s convoluted and extractive land tenure system acted as a strong disincentive from investing in public goods. In his words, “As the ground is seldom tilled otherwise than by compulsion, and as no person is found willing and able to repair the ditches and canals for the conveyance of water, it happens that the whole country is badly cultivated and a great part rendered unproductive for want of irrigation” (Bernier 1916). Nor did the patrimonial structure of the Mughal realm offer much in the way of an incentive to invest; with the peasant constantly at fear of being squeezed by a rotating cycle of jagirdars, it would have been reasonable to ask: “Why should I toil for a tyrant who may come tomorrow and lay his rapacious hands on all I possess without leaving me, if such be
his humour, the means to drag on my miserable existence?” (Bernier 1916). Roy (2013b) may even be correct to suggest one of the initial effects of the decline of the Mughal imperium could have been to improve the lot of the peasant cultivator, no longer squeezed so effectively by a predatory and venal state. These benefits were to fade in time, of course, outweighed by the effects of instability, lawlessness, and across much of the former Mughal realm, predation by outside powers. Yet it underlines the extent to which non-cooperation was prevalent even during the high phase of Mughal power. We can summarize by observing that, under the pressure of a prebendal and predatory regime, local communities had limited incentive to invest in public goods needed for agrarian production, such as irrigation or storage: given their right to the surplus, the only group who might benefit from such an investment were the landlords, but these agents were rotated too frequently to benefit from undertaking any investment on their own account. The state may have solved this problem by undertaking direct intervention and financing of agricultural investment itself, and in the challenger regimes of the south, and in particular in Mysore, such interventions did in fact occur during the eighteenth century. However such actions were neglected in the Mughal case, most likely due to the comfortable surplus already enjoyed by the urban elite, which in any case found itself distant from the daily realities of agricultural production.

The core Mughal state had concentrated public investment in the one form of public good most central to its military prowess, namely the existence of functioning roads, and canals to connect its gangetic heartland; with a secondary allocation to religious endowments. Yet with the waning of imperial power, other actors had no real incentive to reprise either of these roles.

The Rise of Bandits, Dacoits, and “Criminal Tribes”

An immediate effect of Mughal decline was a substantial increase in banditry and violence across the north, which proved persistent under colonial rule. As Mughal India began to fragment into competing successor states (Awadh, Bengal, Agra, Benares, Allahabad) “the tussle for power” would often “mean an intervening period of ‘banditry’ and insecurity” which lasted well into the colonial era (Singha 1998: 18; cf. Also Richards and Rao 1980). Early East India Company documents report “a sense that the Bengal districts were being overwhelmed by banditry” (Singha 1998: 32). Among the many struggles waged during
the years of Company Raj, an important one was against so-called 'criminal tribes', or the widespread practice of highway robbery by 'dacoits' who, in the Company’s view, were a dedicated subcaste; “robbers by profession, and even by birth; they are formed into regular communities, and their families subsist by the spoils which they bring home to them”\(^1^6\).

Another alleged ‘criminal tribe’ were the thuggees or thugs, who were viewed by colonial authorities as a hereditary criminal fraternity, engaged in the strangling and robbery of travelers. Finally, in addition to thuggee and dacoity in Bengal and northern India, the British colonial regime faced a major zone of ungovernability in the central provinces of the country; on the shifting and unstable borderland between Maratha and Mughal spheres, now in Madhya Pradesh, the Pindari mercenaries gained revenue through raiding and looting, moving on horseback to evade retribution.

Multiple sources suggest that such lawlessness and violence was not a general phenomenon, but specific to central and northern India following the decline of the Mughal realm. First, we can approximately date the origin of the institution of ‘thuggee’ to this era. In the colonial documentation on banditry by so-called ‘thug’ castes, “accounts of actual attacks do not go further back than the early nineteenth century,” and genealogies of these ‘criminal tribes’ constructed in the 1830s “do not seem to stretch further back than three generations” (Singha 1988: 185). This would place the earliest members of thuggee groups at around the middle of the eighteenth century. Stewart Gordon reaches a similar conclusion, noting that ‘what the British saw as “thug” - “a national fraternity of murderers” consisted of a small core of families, members of which had been murderers for several generations,’ i.e. since the mid-1700s\(^1^7\). Second, we can spatially isolate the ‘violence’ reported by colonial authorities to the former Mughal territories. Reports of criminal violence and lawlessness in precolonial and early colonial India were not universal, but spatially concentrated in central-northern India, in particular in interior Bengal (today’s Bihar and Jharkhand), the Central Provinces (Madhya Pradesh), and Rajasthan.

As the British Court of Directors was to remark in 1832:

\[^1^6\]Committee of Circuit to Council at Fort William, 15 August 1772, Supplement, p. 13.

\[^1^7\]Gordon, ‘Scarf and Sword’ p. 425, 428.
“The enormities of the Thugs are mostly committed in our newly acquired possessions, or in the Territories of the Native Chiefs of the Bundelkhand, Malwa and Rajpootana. The supineness or connivance of those Chiefs prevents them from adopting any effectual means of repression, and enables the Officers of these Governments and many of their Subordinate Jageerdars to shelter the Thugs from whom they receive a portion of the fruits of their crimes”\textsuperscript{18}.

The territories of Bundelkhand and Malwa are located in central India, in contemporary Madhya Pradesh; Rajputana is the basis of contemporary Rajasthan; in the north; and the ‘newly acquired’ territories of the East India Company constituted the remainder of the gangetic heartland and the former Maratha domains of north-central and western India.

So why did banditry become so widespread across the north during this time? One factor was the fragmentation of power between competing Mughal polities. The Mughal Empire had created a system of faujdari to police the major trading routes, but during the eighteenth century “the faujdari network on the highways weakened” and “around 1719 it was reported that many of the faujdars on the royal highway from Delhi to Patna had abandoned their posts or were without contingents”\textsuperscript{19}. As a result, opportunities for banditry increased, and many individuals “turned to plunder because of political anarchy” (Singha 1998: 178).

However, beyond the demise of Mughal military institutions, it is clear that internal attributes of north Indian states, and in particular their patrimonial nature and perpetual revenue shortfall, were responsible for the rise. British officials alleged that far from clamping down on banditry, “Indian rulers and zamindars merely mulcted bandit gangs and released them if their own revenues were not affected, and they even sheltered them for a share of the plunder” (Singha 1998: 25). This practice seems to have been prevalent across the post-Mughal states of the north: Balwant Singh, the zamindar of Benares, Shuja-ud-daulah, the ruler of Awadh and the revenue farmers becoming powerful in his territory, all recruited bands of hunters such as the Badhaks or armed camps of cattle-dealers such as the Banjaras.

\textsuperscript{18}\textsuperscript{18}Letter from COD, 28th November 1832, No. 11, para 3, Home, 1833, T&D, Cons B.2, No. 3 NAI; emphasis added.

\textsuperscript{19}\textsuperscript{19}Saiyid Abdullah Khan Qutb-ul-Mulk to Rajah Chhabela Ram, Subehdar of Allahabad, c. 1719, in Satish Chandra (ed.) Balmukund Nama, 1975: 27.
as an inexpensive way of supplementing their forces and claimed a share of their booty as tribute (Singha 1998: 25). Similarly among the princes of central India, the pindaris were entertained as a source of raiding tribute. The commissioner for the Sunderbans, in Bengal, remarked how “there appears to be an established system of revenue derived from robbers which is connived in by all the Durbars throughout Hindustan”\textsuperscript{20}. In some cases, there was no clear dividing line between officials exacting tribute, and ‘dacoits’ or criminal bands, who similarly demanded ransom for “protection” over nearby areas: “many villages, especially in Jessore and Mahmudshahee, pay a regular Malguzaree [land tax] to the chiefs of the dacoits,” exactly as they would to an appointed zamindar\textsuperscript{21}.

Relatedly, eighteenth-century warfare led to the growth of roaming mercenary bands, whose presence was found “especially from Hindustan into the Deccan and Central India” and “highway robbery was probably always an option for such bands as they moved between one employer and another” (Singha 1988: 181). Weaker states of the north may have been especially dependent on the outsourcing of military defense, as they had less capacity to engage in military recruitment internally, and thus became overwhelmed by such bands, whose behavior could not be restrained in the event of late or non-payment; indeed, a norm was established whereby such recruits were entitled to a share of spoils, obviating the need for salaries (Kolff 1998). Even official troops may have partaken in this; in Awadh, for example the “soldiery was in a state of constant disaffection because of massive arrears in pay” (Bayly 1988: 91). Areas of weak state formation also saw an explosion in banditry and highway robbery as they contained more semi-tribal and nomadic groups with the means to make such predation a way of life, and who (like the Bhils of Rajputana) had contracted their services to patrimonial states: “In the flux of state-building and given the wide dispersal of arms, peripatetic communities providing services of various sorts or bands of mendicants might occasionally take to highway robbery” (Singha 1988: 191). This was especially the case among caste-groups and adivasi communities whose way of life involved hunting and gathering, and which were therefore mobile, skilled in the use of weapons, and familiar with the protective terrain of the forest and the hills. Such groups were greatly

\textsuperscript{20}Answers to Cornwallis’ queries, 1789-90, BRC P/52/22, 3 December 1790.

\textsuperscript{21}Extract from progs of Governor in Council, 19th April 1774, Supplement, p. 125.
more present in the north than in the south of India, in part due to the successive efforts of southern rulers to make their territories amenable to cultivation and taxation; the southern kingdom of Mysore, for example, was considered “a garden from end to end” (Bayly 1988: 34) and had in fact largely cleared its forests for reasons of state; eighteenth-century chronicler Kirmani reporting that “the Mysore Sultans cleared off much forest in their wars against the tribesmen of Coorg and the Nayars” (Bayly 1988: 138; quoting Miles, History of the Reign of Tipu Sultan, 1844: 79).

Finally, banditry and violence worsened as patrimonial military regimes transitioned to the pax Britannica of Company rule, and soldiers who had formerly been engaged in military service, in the receipt of regular salaries paid from taxation, instead turned to raiding as a means of maintaining their existence. As the hero of Taylor’s Confessions of a Thug remarks, “I was a soldier by inclination if not by profession, and I thought... we might make as good a thing on it if we went out on expeditions of our own” (Taylor 1839: 336). While the Company had retained many soldiers in its own armies, there was clearly a massive glut of unemployed soldiers in early nineteenth century India, and no provision to facilitate their reintegration into civilian life. Soldiers therefore had to provide for themselves. Singha reports that “members of the bands whom Sleeman classified as “Budhak alias Bagree dacoits” sometimes described themselves as spending eight to ten months in the year hunting,” which implies a strategy of survival rather than mere predation (Singha 1988: 183). In the native states, this malady was partly avoided; native armies were often retained, albeit in more symbolic proportions, and the Imperial Gazetteer remarks that they were frequently redeployed for civilian policing; a legacy that is clear when one considers the significantly higher police numbers in almost all of the native states compared to the areas of direct British rule. By contrast, in Company areas, the “flow between service and banditry” was “more one way and permanent” (Singha 1988: 182).

There is a question mark as to whether the British colonial state ever fully succeeded in taming this violence and lawlessness. During colonial rule, for example, Bengal and Bihar remained unstable in comparison to other regions. In the years of the First World War, “Bengal revolutionaries intensified their armed robberies (“dacoity”), and conspired with German agents to obtain arms and explosives in order to disrupt rail communications” (Stein 1996: 287). Shortly after, in 1922, Bihar was the site of the infamous “Chauri-
Chaura” incident in which villagers had murdered the local police, prompting Gandhi to call off the non-cooperation movement (Amin 1995). Given the continuity in the instability of the north between precolonial India and the same regions today, it is plausible that the anarchy never fully subsided. While during the British colonial era the police evolved “slowly from forces to combat itinerant raiders and armed gangs (thagi or ‘thugs’ and dacoits, especially in the 1840s and 1850s), and from the systems of village and town watch”, nonetheless “as the system of police developed, the oppression, ignorance, lawlessness and incompetence of its lower orders became notorious” (Robb 2002: 160). It is in this context that the Chauri-Chauri incident must be interpreted. Later in the 1920s, Stein reports that “terrorists resurfaced in the late 1920s in Bengal and the Punjab, and appeared for the first time in the United Provinces” (Stein 1996: 305). While public order might have ebbed and subsided, it may be that from the eighteenth century the present day, the rule of law simply never have existed in these regions.

Why Did State Collapse Occur in the North?

If the post-Mughal states of Awadh, Hyderabad, and Bengal represented a decaying patrimonial order, a broader question concerns why their “parent” regime, the Mughal empire, had itself fragmented. The causes of Mughal decline are still widely debated among historians, with earlier arguments regarding the consequences of religious zealotry under Aurangzeb (Hallissey 1977) ceding place to economic arguments emphasizing the growing tendency towards peasant revolt and the empire’s inadequate access to new networks of finance (Habib 1969; Leonard 1979; Bayly 1983). Yet an important explanation also lay in the patrimonial structure of the Mughal polity, which had failed to generate stable mechanisms of leadership succession, failed to reach beyond a narrow, Persian-speaking Muslim elite, and failed to build affective ties beyond the distribution of patronage, making it highly vulnerable to secession or revolt as soon as the flow of benefits from the imperial center began to falter. These pressures were exacerbated by the “pyramid-scheme” nature of the military-patrimonial regime, by which new “entrants” must be continually rewarded with additional land grants which, when there are no more available, led to withdrawal and finally collapse (Fukuyama 2011). Such a cycle is common to military-patrimonial
regimes, and was also true of the other central Asian khanates. But the Safavids in Iran and the Ottomans in Turkey managed to survive by grafting their nomadic, military-patrimonial system atop a pre-existing and longstanding Persian or Byzantine bureaucracy, whereas in the Mughal court, warlords remained dominant (McNeill 1993, Hodgson 1974). It was the continued reliance of the Mughal polity upon patronage which explains its collapse.

Second, the absence of a strong bureaucracy at the center posed a problem for long-term political stability in the Mughal empire. In early autocracies the bureaucracy has often played a key role in securing political continuity, as lead officials in imperial China or Turkey managed leadership succession and the affairs of state during periods of interregnum. As Kokkonen and Sundell (2014) argue, this was also important for early European states, which settled on the mechanism of primogeniture as a means of ensuring stability in the presence of multiple potential contenders and the lack of a permanent state personnel. The absence of any mechanism in the Mughal case, either by primogeniture, election among nobles, or even the emperor’s choice, placed the empire at recurrent risk of instability. The default norm was that upon the death of the emperor - though sometimes, before - the sons of the great Mughal would fight for the throne, often in protracted military campaigns which pitted large fractions of the imperial nobility against one another in strategic alliances. A clear advantage of this system was that the eventual victor would have to prove himself a strong general, and, to boot, would finish in command of a broad supporting alliance, perhaps the most important precondition for governing a polity that itself was little more than a vast network of military patronage. Yet the cost of such a method was loss of continuity, with intermittent civil war, and each successor “refounding” his dynasty (often with a new court and a new capital (variously moved from Agra, to Lahore, back to Agra and then back to Delhi), and finally the risk of an indecisive outcome, as occurred following the death of Aurangzeb’s son Muazzam in 1712. Dying of illness just five years into his reign and having killed his most able contenders, Muazzam left the Mughal throne subject to a succession of short-lived emperors selected through court politics, none of whom had proven their fitness in battle, or had the legitimacy to carry the military nobility to his side. From this point the situation degenerated rapidly. In the 1600s, just four emperors had ruled, each the head of a great dynasty; in 1719 four emperors came to power in the space of a single year alone, each deposed as swiftly as the last. “Disputed successions, imbecilic
contenders, and short reigns resulted in a rapid depletion of imperial resources, leading to administrative chaos and regional secession” (Keay 2000: 361).

Under better circumstances the realm might have been able to survive, but by the eighteenth century, the dependency of the Mughal realm on its system of military patronage, which had enabled it to expand so rapidly and absorb into its midst so many minor rajputs, landowners, and poligars, was at breaking point. The number who depended upon the system of military grants approached a quarter of the population, or twenty-six million retainers; and with each successive emperor, the number of dependents increased. The total number of among the senior nobility in Akbar’s time was 133, which increased to 437 under Shah Jahan, and then to 575 under Aurangzeb (Spear 1965: 10). As long as the empire continued to expand, maintaining a growing network of patronage was sustainable, but by the end of the seventeenth century, the system had clearly reached a limit. Already during the rule of Aurangzeb, a rebellion in the Deccan led by insurgent Marathas saw the first formal secession from the realm. The number of jagirs began to decline, and “Jagirdars now so dreaded ending up jagir-less that they defied orders to transfer from their jagirs and began to regard them as permanent perquisites which could be leased or farmed out at will and passed on to their heirs” (Keay 2000: 362). The decline continued when the Punjab – in the heart of the Mughal imperium - erupted in revolt against Mughal rule, and though “Banda Bahadur [the Sikh leader] ... and along with him seven hundred other Sikhs, were captured and slain in 1715, Sikh hostility continued to subvert the foundations of Mughal power till the province was in total disarray in the middle of the eighteenth century” (Alam, The Crisis of Empire in Mughal North India: Awadh and the Punjab, 1707-48: 134). The chain of compliance upon which the Mughal state had depended was ever thin. Without the prospect of rewards from the center, it now became non-existant. Mansabdars no longer remitted their share to the imperial center; nor did they contribute soldiers and horses to war. The Mughal empire persisted in name only: a symbolic affection given by the Muslims of the realm to their emperor in Delhi through the reading of his name during Friday prayers in mosque.

The reliance of the Mughal polity on patronage is shown by the fact that its tax system, while in principle able to mobilize vast economic resources, was able to capture little of this rent for itself. Maddison (1971) had estimated that the Mughal polity raised up to
half of national income in taxation, while Raychaudhuri (1982) describes the Mughal state as “an insatiable Leviathan” whose “impact on the economy was defined above all by its unlimited appetite for resources.” Yet seen from the point of view of the imperial center, rather than the landed peasant, the resources of the state would have appeared very different. While a vast flow of wealth may have been extracted from the peasant proprietor, this reduced to a trickle by the time it reached the royal court. By the early seventeenth century, mansabdars accounted for 82 per cent of the total annual budget of the empire, leaving the emperor’s share at just 9 per cent (Richards 1995: 63). The overwhelming proportion of the resources that were extracted from the peasantry were therefore simply redistributed among local elites, and the activities they supported. Along the gangetic plain, cities such as Lucknow, Patna, and Allahabad became centres of imperial patronage, homes for sedentary soldiers, Urdu poets, and Islamic scholars, and in their own right thriving centers of opulence, learning and culture for the privileged elite of the Mughal polity. Yet these were sites of consumption, not of production or public service, and in a way testified to the scale of the Mughal client-state, and its ability to dissipate wealth among its agents rather than centralize and control resources. Unlike the rising coastal cities of Calcutta, Bombay, and Cochin, little was traded in these gangetic cities, and little was produced, save luxuries and ornaments for the nobility. The legacy of Mughal urban society, centered on well-to-do Urdu-speaking communities, could still be seen dimly in colonial-era data on literacy by district, which shows the gangetic towns as still-glowing beacons of urban civilization in an otherwise vast pool of darkness constituted by colonial-era United Provinces and Bihar. Given another century, these last bastions, too, would had faded from the map (Figure 5.5).

Nor was this redistribution especially orderly or systematic. The Mughal regime consisted of multiple nested and often overlapping claims of revenue rights. Tirthankar Roy (2013b) has usefully delineated five layers to the tax structure of the precolonial Indian regime, running down from the emperor, to warlords (amirs or taluqdars), secondary landlords (zamindars), primary landlords (village headmen), and finally, the tenants themselves, who would produce the agricultural surplus that was to be contested by the others. Each layer in the revenue chain had competing claims to compensation, with the right to individual plots of land being “a combination of four types of rights - the state’s share of the taxes to finance wars, the secondary landlord’s share to repay the cost of running the revenue
Figure 5.5: Literacy Rates in 1900, and Cities of the Gangetic North. While literacy rates were low across the north of India in 1900, and especially in comparison to the south and the coastal areas, nonetheless the islands of the urban civilization of the Mughal imperium remain evident from the data. These can be seen from the towns and cities that snake the gangetic basin, annotated above.

system, the primary landlord’s share to repay the cost of extending cultivation and make sure that land is well cultivated, and the peasant’s share to repay the labour expended” (Roy 2013b: 52). Needless to say, with so many overlapping claims to the produce of the land, together with concessions made for religious sites or or ad hoc patronage, coordination would often prove difficult, and conflict - whether in the form of peasant revolt, warlordism, or kingly repression - formed a constant feature. Each layer had an interest to maximize its own claim at the expense of the others, and to a substantial degree the institutional evolution of different parts of the country reflected the different paths by which this conflict could be resolved. In the rising challenger regimes of the south of the country, the sovereign worked with the peasant proprietor to exclude the intermediary layers, establishing individual-cultivator (ryotwari) system of taxation. In the north of India, in the former Mughal domains, it was the warlords and zamindars who prevailed at the expense both of the emperor at one end and the peasant at the other, for “the military aristocracy was far larger in scale, more complex and divided, and more distant from the countryside in imperial northern India” (Roy 2013b).
Explaining Divergent Trajectories: North and South

The contrasting accounts of state failure and state formation in precolonial north and south India raise obvious questions. Why did the pressures of warfare not lead northern rulers toward a similar push to reduce the powers of “feudal” lords in favor of centralized revenue assessment and collection? Why did rulers in northern India choose instead to reduce grants to judicial offices, and accede in the sale of state functions? Why did so many post-Mughal rulers accede to having the East India Company collect their revenues or provide military defense, instead of undertaking these activities internally? And why did they generally collaborate with British overtures rather than wage war alongside Mysore or the Marathas?

These questions perhaps have a common answer, which is that if we are to understand the diverging policy trajectories of northern and southern India, then it is necessary to take into account the coalitional structure that existed internal to Indian precolonial regimes. The northern polities, as successors to the Mughal Empire, had no incentive and perhaps little capacity to challenge the privileges of the Muslim and upper-caste Hindu gentry upon whom the polity depended. That the pressures of warfare were real enough is revealed by the steady incursions of the Afghans, Marathas, and the East India Company into their domains. Yet there were limited attempts at military or political reform. The basis of the regime in the south was its royal court, brahmin officials, and distantly below that, the village headmen or deshmukh of the Deccan; they shared an interest in repelling external threats, having just spent a century overturning the Mughal yoke. It is in this context that we can understand why “the Sikh and Maratha movements had begun as assaults on the town-dwelling Mughal elites, and chroniclers portray several instances of town riots against Mughal officials and wealthy people” (Bayly 1988: 177).

By contrast, the northern regions at the heart of the Mughal state had no insurgency basis. They were the remnants of a decaying power, dominated by its clients; fear of increased revenue demand gave large estate holders reason to resist Company rule, but until their last gasp in 1857, little means or popular legitimacy to mobilize support for their cause. Though the Mughals had ruled for generations, the middle-layer of urban Muslim converts who governed the gangetic heartland remained the remnants of a foreign
imposition, without deep support in the villages. The core elite of the Mughal state was overwhelmingly foreign, not only in language and religion, but also by origin; even as late as the 1640s, Shah Jahan’s historian, Abdul Hamid, reported that, only a fifth of higher Mansabdars were Hindu (mainly Rajputs) and less than a sixth were Indian converts to Islam; the remaining two-thirds being from Persia, Central Asia, or Afghanistan (Robb 2002: 91). As such, they could not mobilize the peasantry on a religious basis, as could the Sikhs or Marathas; yet increasingly, nor could they match the reliable material rewards of Company service.

A coalitional theory can explain the stagnation of northern states vis-a-vis the reformist programmes of southern regimes. The Sikh Empire, the Maratha Confederacy, and Mysore under Haider and Tipu were “breakthrough” regimes in which a modernizing autocracy allied with the broad peasantry in order to purge the middle layer of the ancien regime and mobilize against external threats. Their dynamics can be compared to those identified by Barrington Moore (1966) for the “breakthrough” coup of Meiji-era Japan, as also described by Samuel Huntington (1968) of Ataturk’s Turkey or the Shah’s Iran. This logic demanded a purging of the old hangers-on (the samurai and daimyo in Japan; the zamindar, qazi, and mufti of the Mughal state) and the reallocation of resources to the military and to the peasantry.

The clearest example of breakthrough regimes in this sense were Mysore and Travancore, both of which involved a lone autocrat establishing a new regime by rooting out the vestiges of the former aristocracy, and seeking instead to forge a direct link with the rural peasantry. In Travancore, the regime’s founder Marthanda Varma had begun in 1729 by executing the heads of the eight largest estates of the realm, who sat on a governing council called the Ettuveetil Pillamar (the “Lords of the Eight Houses”): having displaced the key Nair chiefs, Varma formed an alliance with the religious establishment instead, founding a new administration by appointing Tamil brahmins to key positions, and raising money for his accelerating military expenses by an elaborate system of fines, tariffs on trade, and state monopolies, rather than the conventional route of taxing the peasantry.

Though the Mysore case is unusual, as the role occupied by Haider Ali and Tipu was only as chief regent to the king, who continued to remain the nominal head of state even as all executive power was concentrated in their rule.
In Mysore, the kingdom’s founder Chikka Deva Raja had in a similar manner executed all 400 local chiefs upon his rise to accession in 1673, making himself absolute monarch (Nagaraj 2003), though it was under Haider Ali, who served as Sarvadhikari (chief regent to the king) from 1761 to 1782, that Mysore fully squeezed the rentier class. That this occurred with the acquiescence and support of the peasantry is in no doubt from the numerous visitors to Mysore from this period, such as the Persian writer Kirmani, who describes how the “peasantry were pleased by favourable assurances and agreements and by keeping a constant eye on the collectors of revenue, checked with the aid of the confidential information supplied by Harikars (information officers) and the Akhbaranawaz (news-writers)” (.534, quoting Kirmani 489 without quotation marks), Wilks, who found that while “official men had cause to tremble; but the mass of the population felt that the vigour of the Government compensated for many ills” (Wilks in p. 403, Mysore Gazetteer II 2542-2545), and Fullarton (pp. 62-65) who reported that while the “renters, the tax-gatherers and other officers of revenue fulfilled their duty with fear and trembling, for the slightest defalcation was punished with chaubuck or with death,” this meant that “the husbandman, the manufacturer and the merchant prospered in every part of his dominions [where] cultivation increased, new manufactures were established, and wealth flowed into the kingdom”.

At a further extreme were cases such as the Sikhs in Punjab, who did not so much form an alliance between a new autocrat and the lower strata as a purely ‘bottom-up’ act of social mobilisation. A millenarian religious movement, the Sikhs formed an egalitarian brotherhood of misls (cells), offering Hindu jats and other agrarian groups not only an escape from the caste system, but also a means of organized resistance to their Mughal overlords. Richards (1993: 258) notes that their message “focussed long-standing peasant and lower-caste grievances against the regime and its allies, the qazis and other prosperous Muslim gentry who were grant-holders under the [Mughal] regime,” and Grewal (1990) characterizes their attraction simply as “a religious appeal rooted in class hatred”. By consequence of this broad-based, anti-elite social mobilisation, the administrative elite of the Sikhs under Ranjit Singh was unusually egalitarian: their ranks included not only merchants and soldiers but “former barbers, water-carriers, and scavengers”, and individuals from such downtrodden social backgrounds were also able to enjoy “jagirs [land grants] like the other
members of the ruling class” (Grewal 1990). Having removed the Mughal aristocracy, the
Sikhs were able to not only maintain and renovate its apparatus of state, but to do so on
a more efficient basis as in the Punjab “rates of assessment in the early nineteenth century
were lower” following a significant reduction in rural taxes (Grewal 1990).

Somewhere between these two polarities - an alliance between new autocrats and the
peasantry and a pure social mobilization from below - lies the case of the Marathas, a
group that, like the Sikhs, began as a brotherhood of warrior groups that mobilized in
resistance to Mughal rule, but like Travancore, mobilized brahmins into public service to
build a bureaucratic apparatus. Organized under Shivaji and his successors, the Maratha
warrior elite forged a direct link from the Deccan village to the new state, displacing in the
process both the region’s transient Muslim governor and the indigenous Maratha gentry of
deshmukhs that had ruled the Deccan the medieval period, and which Islamic states had
cooprtated in this role. Whereas “when the Muslims invaded the Deccan in the 13th century,
they found [deshkukhs] solidity entrenched in rural society, and their authoriy was accepted
without demur” (Kumar 1965) by the late eighteenth century and a century of Maratha rule,
the deshmukh “had long ceased to hold any official status and had been relegated to a more
or less ornamental position since the days of Shivaji” (Edwardes .387). This was because of
their cooptation by the Deccan sultanates, and later the Mughals: as Gordon (1993) states,
“Shivaji’s was not a revolt against deshmukhs, but a polity that attempted to integrate them”
after Islamic regimes from the north had been “obliged to purchase the cooperation of the
deshmukhs by recognising their traditional status” (Kumar 1965). Instead of operate via
local nobles, therefore, the Maratha polity at first developed an administrative regime which
connected to the village chiefs or patels; later, it sought even more immediate control within
the village by appointing its own agents locally.

Whereas on the peripheries of the Mughal imperium - the Sikhs in the northwest,
Marathas in the southwest, and Mysoreans in the south - polities were born from the ex-
clusion of the old patrimonial elites, and the desire by rising autocrats to refound political
order on a new basis, the coalitional basis of regimes in north India, by contrast, favored a
reversion to patrimonialism. Here the rulers of the post-Mughal era were allied with large
tax-collectors-turned-landlords and with the urban strata of Muslim officials whose roles
had become rent-seeking and extractive. Within such a system, there was little or no in-
centive to reform; even in the final decades of indigenous rule in polities such as Awadh, Bengal, or Agra there seems to have been few attempts to reform the state. Such a coalition also facilitated British entry into such areas. As Bayly (1988) has argued, British officials were able to capture and govern so many parts of the former Mughal Empire because they co-opted the same local interests - merchants, moneylenders, and commercial agriculture - that had been fostered and yet neglected under post-Mughal regimes. In this way, “Company Raj” did fulfill the role of a modernizing coalition between commercial interests and the state, and early attempts to secure highways, reform the judicial system, and reduce corruption were supported by the urban sector in cities such as Calcutta and Surat (Stein 1984). If a century later the economic benefits of these institutional reforms did not seem very apparent, it is largely because fiscal policy was so completely skewed towards British interests: the “drain” of wealth imposed by a large dividend for the Board and then parliament, and the overwhelming neglect of public goods in favour of spending on maintaining a needless (from the Indian perspective) military juggernaut, directed to furthering British interests in Burma or Afghanistan.

With the advent of Company rule, the old elites of the Mughals’ gangetic heartland - the Islamic justice, the tax-collector, the mercenary - were suddenly displaced. British legal reform ended at a stroke the judicial responsibilities and perquisites of the kazi and the end of warfare displaced the mercenary corps. Faced with a sudden loss of status and sustenance, the officials of the ancien regime exploded into revolt, bringing with them a number of their former subjects who found themselves missing the permissive taxation and relative stability of the status quo ante. These included many who were not the obvious beneficiaries of the Mughal successor states, such as Jat farmers around Delhi and high-caste peasant communities in Awadh (Bayly 1988:189). Yet this coalition of interests was balanced by other interests opposing such a return, including the soldiery Punjab, which remained “a solid base for the British throughout the revolt”, and the “new professional classes of Calcutta and the district towns”, who were “little inclined to support a movement which they saw as a typical zamindar revolt in a backward area of the country” (Bayly 1988: 184).
Conclusion

In the period preceding British ascendency over India, there were large regional differences in the degree of state formation and political centralization, and dynamic patterns of change as some states were continuing to centralize while others were undergoing a form of political decay. This chapter has sought to understand how such patterns came into existence. I have shown that following the breakup of the Mughal empire, “challenger” regimes across the Mughal periphery engaged in extensive defensive modernization, primarily under the pressure of interstate warfare, both against the Mughals, each other, and increasingly with European trading companies seeking territory on the subcontinent. On the other hand, successor polities to the Mughal empire itself, such as Bengal and Awadh, were weakening under this pressure and undergoing a steady “repatrimonialization” of the state. As challenger regimes faced increasing costs of warfare due to increases in the scale of armies and the adoption of gunpowder weaponry, this augmented the fiscal pressure on the state, and eventually necessitated reforms so as to improve revenue collection – including the introduction of mechanisms of policing and surveillance over officers of state, and the direct recruitment of revenue collectors.

I argue that the link between military-fiscal pressure and defensive modernization, however, is a conditional hypothesis: elites invested in military capacity where the coalitional structure of the regime facilitated the exclusion of the warlord and landowning classes, and by consequence the centralization of fiscal and military organization. By contrast, where the structure of rightsholding left warlords as powerful actors, political centralization did not occur, and in many cases a process of “repatrimonialism” could be observed.
Chapter 6

The Puzzle of Persistent States Legacies Under Colonial Rule

The foregoing chapter has shown the trajectories of state formation in India’s north and south during the precolonial period. Yet as we have seen, these trajectories were dynamic, not static. This leaves us to explain why the legacy of state construction of the precolonial past persisted during the colonial era, rather than continuing to evolve under the aegis of colonial rule. To this puzzle there are two elements. The first is a well-established problem in political science: the explanation for institutional persistence, or, why it is that certain institutions tend to continue despite changes in policies designed to lead to their alteration. Answers to this problem tend to offer one of several possible solutions, some of which were broached in chapter 2: that the endowments of politically relevant actors provide enduring incentives to frustrate implementation of policies adverse to their interests; that societal norms have consolidated which are adverse to systemic change; or that changes in policy are dependent on the state of the administrative infrastructure, which tends to persist from one time period to the next. The second issue to address is why the colonial era led to a freezing of pre-existing state legacies, rather than their dynamic evolution in response to colonial-era policies such as patterns of direct and indirect rule, attempts at land tenure reform, or patterns of resistance and mobilization to the colonial project. To this issue I argue that the post-1858 colonial state was a fundamentally conservative project, which under the surface preserved much of the local institutional inheritance of India’s precolonial
states, including within the administrative hierarchy of the colonial regime itself.

The rest of this chapter examines patterns of institutional persistence from the pre-colonial to the colonial era. First, I show the persistence of precolonial institutions in the colonial regime by examining a paired comparison from among the areas of directly-ruled India, the respective governance of the Bengal and Bombay presidencies. Both Bengal and Bombay presidencies were coastal regions governed from a major trading hub (Calcutta and Bombay, respectively), subject to direct British rule, and operated upon the basis of the same laws and policies following the formal birth of the colonial regime in 1858. However, each had radically different outcomes in their degree of land taxation, rule of law, and public goods delivery in the colonial era, a difference that I trace to a legacy of persistent precolonial institutions, specifically, the continuation of the weak state capacity of the Bengal sultanate and the strong administrative structures passed from the Maratha confederacy to the British empire. Second, I then examine the specific forms of general institutional persistence from the precolonial to the colonial state across India more broadly, covering the categories of administrative persistence, continuities in land and revenue administration, and continuities in local-level resistance to centralized state authority.

A Paired Comparison: Institutional Persistence in the Bengal and Bombay Presidencies

From 1757 to 1858, the British East India Company consolidated its power over the Indian subcontinent. By the end of this period, two-thirds of modern India was under its direct rule, while a remaining third continued to be governed by indigenous rulers, henceforth officially preserved as “princely states” within the architecture of colonial government. A large number of studies focusing upon institutional legacies in South Asia, have made the difference between areas under direct and indirect British rule the basis of their investigation (Iyer 2010, Mukherjee 2013, Verghese 2016). Therefore, in order to examine how precolonial institutional legacies influenced variation in patterns of governance under the colonial regime, in this section I compare two regions of India that were both subject to direct British rule in the period before 1858: the imperial presidencies of Bengal and Bombay. During the half-century from 1765 to 1818, both regions were conquered by the East
India Company, and in both cases, an indigenous regime (the Bengal Sultanate and the Maratha Confederacy, respectively) was displaced and its government formally replaced by that of a British governor, responsible to the governors of the East India Company and eventually to the British crown. However, by the high phase of British colonialism from the late nineteenth to the early twentieth centuries, large differences in rule of law and public goods delivery were evident between the two regions, and these differences can be traced to precolonial patterns of governance and administration.

The Bombay Presidency

The areas of former Maratha control in Maharashtra and Gujarat, together with the province of Sindh in Pakistan, fell under the direct rule of the Bombay Presidency. In many ways, the trajectories of colonial-era Bengal and Bombay presidencies offer a valuable paired comparison. In both Bombay and Bengal, the East India Company had established itself as a trading company at a port town - Bombay or Calcutta - and in both cases, territories were eventually brought under British control following invasion and defeat of an indigenous ruler (in 1757-65 and 1775-1818). Afterwards, both were subject to direct rule under the administration of a colonial governor. And yet, a century later, the Bombay presidency was among the most socially and economically advanced regions of the subcontinent, and widely seen as the most effective administration in colonial India, while Bengal remained lawless and poor. In Bombay, famine and illiteracy were less common, and crime less widespread: according to Haye (1853), bandits “infested every part of India except the Konkan [the former peshwa’s domain] in the Bombay Presidency”. Throughout the nineteenth century, the Bombay presidency raised significantly more revenues in taxation, even accounting for the state’s higher per capita income, and by consequence spent more of these revenues on public works, education, and famine relief. In 1892, thirty-four years after the transfer of sovereignty to the British crown, the Bombay presidency was collecting over two and a half times the level of taxation per capita as Bengal, on the basis of a level of agricultural output that was only one-quarter higher, and collecting three times as much revenue per capita as the United Provinces, though producing only two-thirds higher a level of primary produce.
This success by the Bombay presidency depended, to a significant degree, on the maintenance and improvement of the system of government inherited from the region’s pre-colonial Maratha rulers, which was adopted by colonial officials with minimal modifications. As detailed in chapters 3 and 5, the core domains of the Maratha confederacy were among the most administratively developed in precolonial India, including the most effective mechanisms of justice, surveillance, and significantly higher levels of per capita taxation than any other part of the subcontinent. Upon British conquest of the territory, minimal disruption was made to this legacy: “The policy of Mountstuart Elphinstone, who was appointed commissioner for the settlement of the Deccan in 1817 and became governor of Bombay two years later, was to interfere as little as possible with the system which he found existing in the conquered territory” (Edwardes 1932). This included the administration of justice in which “the Maratha arrangements for civil justice were maintained more or less unaltered,” and also the entire system of revenue collection, which was simply continued as found, and which Elphinstone wished “to preserve as far as possible unimpaired the practice of the Maratha Government”. The Company made cosmetic changes, mainly in cleaning up the assessment process through new surveys and streamlining the ryotwari system of the Marathas, which had in parts become corrupted during the final years of Maratha rule, and changing some, but not all, of the names given to officers of state.

This continuity is very clear with respect to fiscal administration. Butterworth in 1932 remarked upon how “the general system of revenue administration in force in 1914, and also at the present date, is practically the same as at the date of the assumption of the government of India by the crown” (Butterworth 1932). Under the Marathas, the chief minister (peshwa) appointed regional subsurdars (governors), who in turn oversaw a network of officers (kamasvidars) responsible for collecting taxes via their own auxiliaries (mamlatdars), who was negotiated directly and audited villages and the state. Under the face of reform, much of the Maratha system was preserved. Most significantly, the role of the Maratha mamlatdar was maintained, both in name and in function. The only shift was that the size of the districts under their responsibility were standardised to a more regular scale (Kumar 1965). And just as the mamlatdars had done under the peshwa, they were charged with revenue collection. Just as before, the mamlatdar saw over a team of sub-officials tasked with district tax collection; just as before, these negotiated with the village
heads to secure the revenue payment; just as before, he passed this sum to his superiors. The only major difference now was that he was responsible to a colonial government in Bombay, rather than an indigenous government in Pune. As Gordon (1993) remarks, after describing in detail the precolonial system of Maratha tax administration: “if it sounds suspiciously like the British District Officer of the nineteenth century,” it is because “his role was patterned after the Maratha kamasvidar”. Otherwise, the machinery of Maratha administration simply continued, under the same personnel.

Because the Bombay presidency inherited and built upon the Maratha system of revenue administration, this offered an easier path towards exercising infrastructural power. Rather than construct a new administrative apparatus for the provision of justice, famine relief, or education, the Bombay government could pass these obligations to the fiscal administration. In this way “the various administrative appointments and establishments created between 1858 and 1918, in response to the progress and requirements of the people of the presidency, were grafted upon, added to, or linked more or less closely for administrative purposes with the framework of the revenue administration, which had been constructed, tested, and improved during the first half of the nineteenth century” (Edwardes 1932). By working with the revenue administration, the presidency could achieve far-ranging mobilization of revenue and delivery of public goods and services, in a manner that was simply not possible in other states. As the Montagu-Chelmsford Report argued, in Bombay the revenue administration was “so close-knit, so well established, and so thoroughly understood by the people, that it simultaneously discharges easily and efficiently an immense number of other duties”, including “registration, alteration, and partition of holdings; the settlement of disputes; the management of indebted estates; loans to agriculturalists; and above all, the famine relief” (Report on Indian Constitutional Reforms, 1918). Thus in Maharashtra, unlike in Bengal, the district officer was “always there in the background to lend his support, or, if need be, to mediate between a specialised service and the people” (Report on Indian Constitutional Reforms, 1918).

**The Bengal Presidency**

The first region to fall under Company control was Bengal, over which the Company gained expanding military control from 1757 and official revenue and administrative rights
in 1765. The literature on colonial Bengal is extensive (e.g. Ray 1979; Salim 1975), and I will sketch only the broadest of outlines here. In short, whereas the East India Company only began to assume control over much of southern, western, and northwestern India in the early to mid-nineteenth century, Bengal had fallen under the control of the Company an entire century earlier. This gave the Company time to experiment with a wider range of institutional reforms as well as to train a larger cadre of Indians into its administrative service. Moreover, the capital of colonial Bengal, Calcutta, had been the center for the East India Company’s operations: this exposed the city in particular to a significant influence of European trade, culture and religious proselytism. After an initial period in which the Company largely continued the administrative institutions of the previous rulers, from 1772 a series of alterations to taxation and the administration of justice were introduced, leading the way to a more comprehensive reform, in the form of the ‘permanent settlement’, in 1793. From the steady flow of documents detailing laws, judicial outcomes and reforms of this era, it is possible to read into late eighteenth and nineteenth century Bengal as a story of colonial state formation, and ’contemporary Englishmen sometimes described the changes that took place in Bengal as a “revolution’’ (Marshall 1987). Yet there was also a great deal of persistence between the pre-colonial and colonial administration of Bengal, in particular at the local level, and colonial-era decisions were largely responses to dealing with dilemmas that these precolonial institutions had posed.

By the mid-nineteenth century, it was clear that Bengal was confronting persistent problems inherited from late precolonial era, namely rural insurrection, robbery (dacoity), and very ineffective collection of taxes. This reality is expressed through the narratives of colonial officers and contemporary visitors to the province, who portray not the unflinching advance of a disciplined, colonial state, but rather, a region that remained mired in criminality, corruption, and the absence of any effective governance throughout large tracts of its territory. Lord Minto, who completed a survey of the region, remarked how in 1810 “a monstrous and disorganised state of society existed at the very seat of that government [in Calcutta] to which the country might justly look for safety and protection.” Lovett characterises the region as “administratively starved” and among its problems the fact that “its communications were bad; its government buildings were mean and inadequate; its police stations were few. It contained no troops and no mounted police” and “the whole of
the Lower Provinces needed a larger administrative staff” (Lovett 1932). Similarly, in his India: Its Administration and Progress, Strachey (1888), who had journeyed throughout India in his capacity as viceroy and lieutenant-governor, remarked on how in Bengal “there were almost no roads, or bridges or schools, and there was no proper protection to life or property. The police was worthless, and robberies and violent crimes by gangs of armed men, which were unheard of in other provinces, were common not far from Calcutta”.

The general conclusion from the study of colonial Bengal is that its problems were generated by the colonial state itself, and that the reforms introduced by early colonial governors were failures (Jauregui 2013). Yet such a conclusion neglects the fact that almost all of the problems that could be observed in nineteenth century Bengal were also observable in the eighteenth century. Eventual reforms by the East India Company administration were explicitly motivated in response. The tendency for Bengal to be under-governed - as Lovett (1932) calls it, “under-administered” - was an inheritance from the region’s precolonial elites, who “had neither succeeded in creating a strong state machine nor succeeded in forging close links with the mass of their subjects” (Marshall 1987). Under the Mughal empire, Bengal was a marginal zone, the eastern frontier, and “[f]or long centuries there had been vague confusion varied by the consolidation of some central power strong enough to force payment of revenue and raise military levies where required” (Lovett 1932). The Nawabs, who claimed authority over this zone in the early eighteenth century did little to reverse such patrimonialism, and the decision to outsource tax collection and defense to a foreign trading company, was reflective of their inability to defend their borders or to enforce fiscal compliance internally. This was proven by a number of events that preceded the formal claim of the Bengali diwan by the Company in 1765, notably the Maratha invasions of the 1740s, and an uprising of zamindars in 1759 in Bihar.

Continuity, rather than change, was the real story of British government in Bengal. Such continuity was easy to see at the outset. During its early years the East India Company maintained the governing institutions of the Nawabs, and “the British element was at first little more than a thin veneer” (Marshall 1987). Upon assuming control of the diwan (treasury) of Bengal in 1765, for example, the Company continued to farm tax collection to local revenue officers, to use the pre-existing Mughal revenue records in Murshidabad, and to implement the same rates and rights as the earlier Mughal settlement (Peers 2006).
Indeed formally, the British were simply collectors of revenue, even if in practice they also controlled the civil administration of the state by appointment of the deputy to the Nizam, and also its armed protection. More substantive changes came in 1772, when tax records were moved to Calcutta, district councils introduced, and a board of revenue established to oversee tax collection. These introduced a “revolution from above” in which the bureaucratic apparatus of a European trading company was grafted atop the region’s indigenous institutions, with a new category of “banians,” or personal agents, acting as interpreters between the colonial rulers and the indigenous order. In this process of supplanting the Nawab’s elite, some substantive changes were made; notably, the pensioning off of the faujdars (the Mughal administrative staff, by the eighteenth century already a nominal presence in many areas) and the disarmanent, if not the abolition, of the zamindars. Yet below the staff of the company, the old ways continued. The company “could not do without the services of the merchants and bankers, of the great zamindars, and even, for a few years, of the Muslim grandees” (Marshall 1987).

Finally, the major reform to revenue management, the permanent settlement, would occur 30 years after the Company first assumed responsibility for the revenues of the region. The permanent settlement gave tax collectors (zamindars) rights of saleable property over their concessions, in the hope of creating a system of agrarian property rights, and is the subject of an extensive literature. Generally, it is seen as a failed reform, as many “estates” became neglected, sub-leased, or abandoned. In retrospect, however, the permanent settlement was not so much a failed reform, but a failure to engage in reform, by setting into law the status quo of the late precolonial era. In its early years the Company had made efforts to extend its control further down the administrative hierarchy that it inherited, appointing Europeans to larger zamindars or the jurisdictions of faujdars and amils. The “permanent settlement,” by contrast, marked the end of serious efforts to penetrate to the local level, instead offering zamindars in perpetuity the right of exploitation of their domains. Instead of extending the state apparatus down, the settlement established courts to allow tenants and claimants to bring their case before the state, so to speak bringing governance ‘up’ from the local level. But the consequence was that the colonial state had little information from the local level by which to implement decisions and judgments, and thus, lacked effective capacity to enforce change on the ground. Courts could not secure prosecutions as long
as witnesses could be intimidated, false witnesses purchased, and money could purchase a better representation for the zamindar, in the form of their vakil. Extensive protections granted in law for tenants against their lords (zamindars) were simply ignored, and the former were exploited ruthlessly: for example, when tax rates were lowered in nineteenth century Bengal from their previously elevated levels, such reductions were “immediately followed by new taxes and new impositions” by zamindars on their tenants, who were made no better off (Amini Report, 1778. Ramsbotham, Studies, p. 188). Reports were written by colonial officials regarding the scale of the problem, but as long as peasants were not represented politically, and rural unrest did not become revolutionary, there were few incentives to alter the situation. All this is true. Yet it reflects a continuation of the trends that began during the late Mughal years.

If the East India Company in its first decades was engaged in a flurry of judicial, police, and revenue reforms, these were often not so much a product of modernizing zeal, as a reaction to recurrent problems it encountered after assuming control of the province. These problems primarily concerned two related domains, which were lack of fiscal capacity, and the absence of a monopoly of coercive power on the part of the nominal rulers. Because the zamindars were not appointed but rather “claimed” their right as intermediary between the tenant and the state, zamindar status was not secure, but contested. The zamindar was not simply a “tax farmer”, but also a warlord, requiring a constant state of martial readiness to protect his stake from others who might easily threaten it. Zamindars maintained powerful forces in order to protect such claims: The zamindar of Burdwan, for example, had over 30,000 militia in his service (Ray 1979). Such a state of mobilization was necessary, as while in theory provision of law and order were the duty of faujdars, government officials who in Mughal times had been appointed from the Mughal court, in practice, by the eighteenth century the faujdars went unpaid, and in Bengal their jurisdiction covered only selective tracts of the country on the frontiers and in the towns. The structure of the “zamindariat” was continually shifting in response to local power struggles, and “by the end of the Nawabs’ rule 60 per cent of the revenue came from fifteen zamindars” (Marshall 1987) as larger zamindars managed to assert their claims over smaller rivals. Some would have entered the service of their more powerful rivals; though others might have found it better to seek revenue in other ways, such as through highway robbery, theft, and extortion.
Similarly, the fiscal administration of precolonial Bengal was also fragmentary and weak, and this weakness persisted into the colonial era. Some historians have tended to see Bengal as a “tax state”, on account of the vast revenues that it commanded and the dependency of the Mughal empire, at least in its later years, on the regular payment from Bengal. Yet this is an illusion resulting from Bengal’s vast demography, as with a population of 40 million, the total tax revenue of the region would inevitably exceed that of Awadh with 6.5 million, Mysore with 8 million, or the Maratha peshwa with 5 million. For any potential empire, this made control of Bengal a rewarding prize, but the scale of Bengal’s annual total tax collection says little regarding the efficiency or effectiveness of its collection. This was in fact decentralized and balkanized, and left to fragmentary and warlike zamindars. Though much may have been extracted from the peasant by many competing claimants, only a trickle found its way to the treasury: when considered in terms of per capita flows to the royal court or to the Company, the extractive capacity of both precolonial and colonial-era Bengal were very weak.

The basis of separating the paired comparison of colonial-era Bengal and Bombay was that the Maratha territories had been the core state of India’s dominant eighteenth-century power, notable for its administrative and military capacity; while Bengal was a decaying province on the margins of the Mughal imperium, plenty in agarian output, but military weak. While Bengal might be taken as an example of how colonial-era policies led to weak administration, strictly speaking, both Bengal and Bombay presidencies operated in the same manner at first: by working within the structure of the political institutions that existed ex ante. Both in Bengal from 1757-1772, and in Bombay after 1818, governors began with reforms that were largely cosmetic. It is true that in Bengal, this was followed by a second wave of more radical reforms 30 years into the colonial era, but this was not so much the product of inherent zeal as a response to the problems of public administration that resulted from the system in place. In the case of late-eighteenth century Bengal, this meant inadequate supply of taxation to support the Company’s permanent war footing, and the spread of violence and criminality (dacoity) through its rural territories as a result of a longstanding armed conflict between fiscal claimants (‘zamindars,’ actual, former, or aspirational) and the tendency for many of these to turn to brigandage. The solution that
the Company found was to take the existing zamindars and make their rights legally enforceable, and to establish courts that would adjudicate the rights newly created. Whether this failed or succeeded, we cannot learn by comparing Bengal to the Deccan, where there simply were no zamindars, and by consequence such a policy could not have been implemented. The Maratha polity had long ago done away with the Deccan nobility (deshmukhs) and replaced them with centrally appointed officials. In both regions, colonial administrators began from the starting point of dealing with the institutions they found in place. But they arrived at different solutions, as a result of different problems these generated.

**Categories of Persistence**

The comparison of early colonial governance in the Bengal and Bombay presidencies can lead us to an appreciation of how institutional legacies persisted more generally across colonial India. Considering not only these two regions but India as a whole, we can see evidence of continuity from the precolonial to the colonial administration across a variety of domains: in the structures of colonial government; in the management of land rights and the revenue system; and in the structure of organized resistance to state authority, in the form of banditry, resistance from poligars, and so-called “criminal tribes,” which centralized precolonial regimes such as Mysore or Travancore had largely worked to eliminate.

1) *Pure Administrative Persistence*

First, the persistence of administrative infrastructure is particularly relevant in the case of precolonial and colonial-era India, as most historians now reject the notion, common to the imperialist narrative, of a “break” separating these two eras. Under the British empire “Indian institutions, practices and personnel were to form the basis upon which administration was founded” and “distinctive provincial patterns in their administrative, revenue and judicial practices resulted in a variability which reduced the degree of centralized administrative authority that could be achieved over all of British India” (Stein 1996: 207, 242). It could hardly have been otherwise: A trading company with so few administrative staff could not possibly develop its own fiscal and administrative institutions rather than manage, or at least, graft itself atop of, the structures in place. Even at the height of the British
Empire in 1872, a mere one thousand members of the Indian Civil Service nominally ruled a population of 240 million. Because India’s precolonial institutions were diverse from one region of the country to another, this meant sustained variation in fiscal, administrative, and legal systems. The continuity between precolonial and colonial institutional forms was clear. As Robb (2002) remarks,

Many of the same titles and offices persisted, as did many of the alliances with local mercantile power, much of the emphasis upon armed force, and much of the reliance on personal ties and interests among the ruling elites, both European and Indian. The high officers of the Company, and its historians, presented the establishment of its rule increasingly as a decisive break... In its everyday affairs, however, it was necessarily and often deliberately gradual and partial in its innovations. Both Mughal and British empires intruded on and weakened local and regional entities, but did not replace them (Robb 2002: 115).

The early administration of colonial India therefore involved not the creation of an Indian “government,” but rather, a disparate process of “governance” in which a small number of British officials working in different parts of the country sought as best as possible to make sense of the society they found under their charge. In doing so, they applied ad hoc solutions based on temporary expedience, making the East India Company “little more than a collection of regional powers” (Robb 2002: 149), and a regional subadministration was built into its very design, which divided its rule over India into separate presidencies, each with responsibility for a different portion of the country, and its own system of dealing with local rulers and the local agents of revenue policy. Until the Regulating Act of 1773, these presidencies acted independently to the point of having an independent right to declare war and sign treaties; only thereafter were they made subject to the Bengal Presidency’s ultimate control.

As each part of India came under company rule, governors found they had to enact decisions on the ground regarding the organization of the administration, the nature of land tenure, and judicial reform. Because there were different outcomes in each case, and Company rule was extended under the governorship of different individuals, in seeking to explain the differences a great deal of emphasis has been placed on the biographies of gover-
nors and their differing “philosophies” of government; in particular, their opinions regarding the extent to which the British administration in India should attempt to reform Indian society or leave in place or even protect indigenous practices and institutions (Mukherjee 2005; Mill 1858; Maine 1871). Until the revisionist social history of the 1970s and 1980s sought to detail more regarding the socioeconomic structure of eighteenth-century India, far less was written about the circumstances in which administrative decisions were made, and how governors simply responded to differing legacies of precolonial administration, pre-existing structures of land holding, and the potential of local elites to resist attempts at reform. The new social history of early colonial India has largely discredited the notion that colonial administrators acted with dictatorial elan, or made more than gradual, incremental reforms to the territories they governed (Bayly 1988). As Fagan (1932) writes, “up to 1857 the development of administration had been in the main a process of improvisation, which responded to the varying conditions of the territories successively acquired rather than to the a priori requirements of any precisely defined system”.

As a result, we can trace how during the early colonial era precolonial institutions were absorbed organically into the colonial state, and how the key attributes of India’s precolonial regimes - their systems of taxation, hierarchies of rural life, or mechanisms of policing and judiciary - were replicated under the aegis of distinct colonial administrations. Because the Company extended its power over India in chronologically separate campaigns, and because these were then subject to distinct “presidencies” operating initially as quasi-independent powers, we can treat each of these as separate cases.

*Continuity in Land Administration*

Second, just as there was persistence from the precolonial to the colonial era with respect to local administration and justice, a key form of continuity between the precolonial and the colonial era was in the system of revenue administration, and in particular the division between zamindari and mahalwari areas of the north, in which collection was subcontracted to landlords and village chiefs respectively, and ryotwari areas of the south and west, in which individual cultivators were assessed and taxed directly. Because taxation was so central to the colonial state, a vast amount of nineteenth-century administrative literature deals with the respective benefits and disadvantages of each method, a debate that was often
more reflective of English eighteenth and nineteenth century discussions regarding property rights, land reform, and the relationship between commercial agriculture, small proprietors, and the traditional landed estate, than the indigenous conditions of rural India. The volume of these discussions has generated the false conception that the distinction between individual and landlord cultivation was invented by British colonial officials, or that changes to the indigenous system of land administration could be widely imposed (Bannerjee and Iyer 2005). Yet Read and Munro, the chief advocates of individual-cultivator taxation, did not adopt the ryotwari system as a result of ideology, but rather because earlier reforms by indigenous south Indian polities left it as the only option. “By 1820 it was this [ryotwari] system rather than settlements with larger magnates which had become the norm for the new Madras presidency. The reason for this was not so much the triumph of Munro’s deeply held ideal of peasant individualism as an acknowledgement of the rapid changes which had overtaken rural society in the south during the previous century” (Bayly 1988: 68). Thus “near the hearts of eighteenth-century principalities rulers had already warred down intermediary chieftains and magnates in order to deal with village elites... The famous ryotwari systems of Malcolm and Munro were developed in the Baramahal territories of Mysore where Tipu Sultan had already reduced the power of the warlords and come to terms with village headmen” (Bayly 1988: 106).

Perhaps the best proof of this is the off-axis case: areas in which administrators like Munro had designed the fiscal system, but in which precolonial land tenure patterns tilted towards the presence of large landowners. Several examples were to be found among the mirasidar proprietors of Tanjore and Tinnevelly in the far south. Yet rather than submit their claims to the colonial state, here landlords “put up a clever and well-orchestrated opposition to Sir Thomas Munro’s ryotwari regulations in the areas which they controlled during the 1820s and 30s. The result was that they became recognised as ‘ancient lords of the land’ in those districts” (Bayly 1988: 115). This provides a confirmatory observation to the hypothesis that it was precolonial state institutions, in particular, the system of land grants, which patterned decisionmaking by colonial officers, and not the ideology of colonial officers which determined patterns of landholding. By contrast, earlier attempts to impose “landlord” systems on the individual cultivator areas had largely failed, due to the absence of a rural gentry; thus in the Madras Presidency, “attempts to find zamindars like
those of Bengal” proved to be “short-lived” (Bayly 1988: 107). This is perhaps obvious, for “the ‘tilt towards the landlords’ was only possible in areas where large magnates existed... Elsewhere the British sought to improve their relations with the vast body of cultivators” (Bayly 1988: 198):

“In their search for revenue the regional and petty states almost always tried to gain a closer control over the headmen or their village propretors. Sometimes direct representatives of the state acquired revenue-collecting rights at the village level and eroded the power and perquisites of the headmen families. This was particularly true where the state was strong as in the Maratha territories around Poona between 1751 and 1818, or in Mysore where Haider Ali and Tipu Sultan tried to eliminate all intermediaries between themselves and the peasant farmers, subjecting them to new demands for the punctual payment of revenue. Elsewhere the absence of powerful authorities outside the village enhanced the authority of the elites as protectors and petty rulers of the villages” (Bayly 1988: 28).

Another “off-axis” case are those examples where precolonial institutions had operated without zamindars, and zamindar rights of saleable estate were either assigned to village chiefs (poligars) or to intermediaries who were found to play this role (muttah). This rather odd experiment entailed “forcing zamindars on districts where no zamindars were found” (Gywnn 1932), and there are very few instances in the whole of India. Yet it did occur in some parts of the Madras presidency, only to prove an abject failure where it was tried. Village chiefs did not see their relationship to the village as potentially alienable, or their people as tenants, while artificial zamindars commanded no traditional legitimacy over their ‘tenants’ and had no coercive means to secure it. In such cases, then, “the zamindari settlement was not a success, even where it had for basis the old estates of poligars; as to the artificial estates, or muttahs, they came tumbling down almost as soon as they were set up”. Eventually, the experiment was abandoned. In this way “the whole of the Guntur collectorate and much of the Masulipatam collectorate” passed “from zamindari to ryotwari between 1835 and 1849” (Butterworth 1932).

Otherwise, ryotwari structures simply followed the groove carved by precolonial rulers. This was the case in the Bombay Presidency, which simply developed upon the reforms
already set in place by the Marathas: “the ryotwari had existed in Bombay and Madras from ancient times” (Edwardes 1932). As Kumar (1965) notes,

“Under the earlier Peshwas, once the tax to be paid by a village had been settled between the village headman and the revenue officials of the state, the internal distribution of this tax was left to the jathas of the village, and was accomplished on the basis of the customary or rivaj rates of the community. After the fateful defeat of Panipat, however, when the revenues of Mahatashtra stood in desperate need of augmentation, Madhav Rao Peshwa introduced the kamal survey, which anticipated the ryotwari systems of Sir Thomas Munro.” (Kumar 1965)

Such a system managed to “undermine the autonomy of the jathas in the financial administration of the village, by obliging the kunbis to pay their dues directly to the state, instead of through the village community”. In this sense, the “reform” of revenue administration in Maharashtra under colonial rule – the famed ryotwari settlement – was somewhat of a fiction. The transition from mahalwari (village-based) to ryotwari (tenant-based) taxation was already well underway; Munro and his administrators completed the task that the earlier Maratha government had inaugurated.

What about the case of the zamindars? It is true that the zamindar of precolonial times was an ill-defined category, and referred to individuals with differing size “estates”, some of whom were farmers, and others who were tax collectors. In some areas, commercialization of zamindari rights had already begun, and these had become a form of property, which could be bought, sold, and inherited, making the zamindar something intermediary between an owner of the territory, a landlord, and a tax official. The financialization of eighteenth century India had made it possible to borrow loans in order to purchase such rights, and then repay these from the income generated. In these cases zamindars were “tax farmers” rather than corrupted tax officials (Siddiqi 1970). In others, zamindars were military entrepreneurs who had sufficient means to extract from the populus of a given area, but not enough to defend against more organized central states, to whom they therefore paid taxes as a form of protection money. In such cases, the security of zamindari rights may have been too weak and personality-dependent to enable a market in such claims to emerge,
and zamindars operated in the military economy described by Kolff (1998). All that united these different actors was their relation to the state, for “what they had in common was the right to collect taxation and the duty to pass part of it on to the government” (Marshall 1987). In asserting their local right over tax transactions, zamindars were effectively local entrepreneurs who claimed the right to serve a more powerful patron. However in many parts of India in the late precolonial era, zamindars in this sense simply did not exist. Because zamindars did not exist, the option of devolving fiscal administration to zamindars was similarly fictitious.

The alternative to zamindari tax collection was that villages, sometimes acting on behalf of their individual cultivators, would pay taxes directly to an agent of the state, without need for an intermediary. The latter solution was a great deal more stable, and entailed significantly less potential for violence between competing claimants. However, it also presupposed the existence of a state that had assigned administrative staff to this role, with the resources to regularly audit the land and maintain contact directly with villages over a large area.

Finally, even in cases where colonial administrators announced a “change” from one tax system to another - an event which was not entirely common - we must be sceptical as to whether this reflected a change in actual chain of tenure, or simply an alteration of nomenclature. Once colonial administrators had declared a region to be under one form of administration, individuals had to be found to serve as “zamindars” or “ryots”. Yet often this meant simply placing a label on what was already the case, and what was termed “ryotwari” in one area, could just as easily be “zamindari” in another. For example, as Roy (2013b) observes, under a ryotari settlement, “the holder of the ownership could end up being an actual peasant, in the true spirit of ryotwari, or a prominent peasant (mirasdar in south India, bhomia in Rajasthan), a primary landlord (malguzar in central India, talukdar in Gujarat) or a peasant kinship group (mahalwari or bhaiachara in parts of north India).” In short, the ryotwar could well be a zamindar, making nonsense of the ryotwari/zamindari distinction. As a matter of fact, because small peasant proprietorship was already more widespread in south India than in the north, when colonial administrators did announce a ryotwari tax system there, they could pass ownership rights to the existing tenure arrangement, notwithstanding a few instances where the holders of these rights would in fact be
landlords. And because large landlords were prevalent in the north, when a zamindar settlement was introduced, colonial administrators could simply pass ownership to these figures. As Roy concludes “path-dependence between colonial property regimes and pre-colonial ones was thus maintained, even though the continuity was not perfect” (Roy 2013b: 55).

Even if the origins of individual-cultivator and landlord-based methods of taxation were rooted further in the precolonial past than imperial historians or contemporary economists imagine, nonetheless their economic effects may have been very real. On average, in the mid-nineteenth century agricultural output per capita had been higher in the north than the south; though this is largely due to the poverty of Madras, as the former Maratha territories of Bombay province were among the the most productive in India at that time. Nonetheless agriculture across north India stagnated throughout the nineteenth and early twentieth centuries, such that per capita output was scarcely higher a century after the 1857 revolt. There are a number of possible reasons for this malaise. The system of landholder rights, namely the contrast between individual proprietor (ryotwari) and landlord or village (zamindari or mahalwari) cultivation seems a probable cause (Bannerjee and Iyer 2005); ryotwari areas such as Madras had stagnated in the early nineteenth century due to punitive rates of taxation, but grew quickly once rates were reduced. Other contributors may include the continuing instability and insecurity of the north, as evidenced by ongoing dacoity and banditry, and the fact that after 1857 the gangetic plain lost an important source of revenue as soldiers were instead recruited from the Punjab.

**Resistances to Colonial State-Building**

Third, the British Empire not only inherited the state apparatus of the pre-existing regimes they found in place, but also the degree to which precolonial states and regimes had pacified resistant social actors, whether in the form of local zamindars and warlords (as discussed in the previous sections), chieftains (“poligars”), or other forms of organized criminality and banditry that the colonial state came variously to refer to as “criminal tribes” and “dacoits.” As Bayly observes, “[t]he most widespread dissidence was encountered when the British attempted to control and tax territories which had never really come under the direct rule of their eighteenth-century predecessors” (Bayly 1988: 172). Because
Figure 6.1: Agricultural Output per Capita (Rs.), Directly-Ruled North and South India, 1868-1955. Estimates of agricultural output per rural inhabitant are taken from the Cambridge Economic History of India, Table 4A.12. 1868 and 1940 estimates are as given; 1955 figures are from the National Council of Applied Economic Research (NCAER) publication, Agricultural Income by States, 1960-1 and deflated accordingly.
the relationship between the state and the citizen is not a mutual contract, the natural incentive of social actors is to resist state efforts to extract resources and otherwise impose its will over a region, and as we have seen precolonial regimes were in a more or less constant battle to establish their authority in this way. By consequence, the weakness of precolonial states in many regions made the imposition of the demands of the colonial state all the more difficult, as authorities found that their fiscal and judicial reforms encroached on local interests in way not paralleled in other areas where precolonial regimes had already warred down local interests and resistance. In particular, in the former Mughal territories of the north, “since the British inherited the expansive but fragile system of Mughal revenue management, the Company soon found itself in conflict also with the Hindu warrior lords of the countryside” (Bayly 1988: 48). These conflicts led to continual upheavals and resistance to colonial rule. As Christopher Bayly has argued in his study of class relations in precolonial and early colonial India,

“Among the myths which became current in the wake of the rebellion of 1857-8 was the idea that it was a unique event, something that had to be explained in terms of the peculiar folly of the revenue policy of the government of the North-Western Provinces or the foolhardiness of the annexation of Awadh in 1856. In fact armed revolt was endemic” (Bayly 1988: 170).

Rebellion was therefore tied to the gaps and lacunae in capacity of precolonial states. The southern regimes of the Marathas or Mysore had already carried through a thorough extension of state power into their regions, such that “near the hearts of eighteenth-century principalities rulers had already warred down intermediary chieftains and magnates in order to deal with village elites” and “had already reduced the power of the warlords and come to terms with village headmen” (Bayly 1988: 106), by contrast, in northern, central, and northeastern India, the company “encountered resistance in those parts of the northern plains where the Mughals and the eighteenth-century successor states had never really imposed their authority” (Bayly 1988: 173). Bayly thus reports that “the Haryana region near Delhi was the scene of perpetual small-scale rioting as nomadic and wandering groups seized the cattle and silver of nearby farmers” (Bayly 1988: 175), while under British rule there was significant uprest in adivasi (tribal) areas, notably the rebellions of the Santhal
and other ‘Kheroal’ (egg-born) people of 1855 and 1870 in northeastern and northwestern India respectively, and the millenarianist rebellion in Gujarat by the Naikda forest people in 1868 (Stein 1996: 269). “Resistance culminated in the 1899 uprising of the Munda peoples of the forested hills around Ranchi, in eastern India” (Stein 1996: 271). Similarly, in the gangetic heartland, “the northern and southern fringes of Awadh caused continuous trouble for the Lucknow authorities and for adjacent British collectors, particularly during the tense 1830s. On the southern fringes of the plains the rajas and clansmen of the central Indian hills opposed both the British and the attempts of the local states of Gwalior and Rewah to coerce and tax them on a regular basis. In 1842 there was a serious revolt among the Bundela Rajput chieftains which disrupted trade and agriculture in the region for some years” (Bayly 1988: 173). After the defeat of Mysore, insurrection in the south of India was not unknown, but seems to have been comparatively rare. The exceptions here prove the rule: in the few cases where rebellion was noted in the south, it was in those fringes of the former Maratha or Mysore domains where princely rule had been fragile: among the Nayars of the Malabar coast or in southern Maharashtra towards Goa, or among the poligars (pailaiyyakar) of Tamil Nadu.

After the 1857 revolt, colonial tax policy became more lenient; “the colonial state, fearful for social order, had begun to lose the nerve to tax the countryside” (Bayly 1988: 195). As a result, “gradually land revenue diminished as a proportion of government income.” A consequence was that even in the British colonial era, patterns of precolonial state formation continued to determine the scope and efficacy of the colonial state; specifically the degree to which taxes could be imposed and revenues could be allocated locally for the purpose of public works.

The Outcome of Persistence: Local-Level State Capacity Under the British Empire: 1858-1947

For the ninety years from 1858 until the stroke of independence in 1947, India would be governed as a single entity under the British crown, and it was during this period that most of the reforms associated with the colonial era occurred, such as the rise of the Indian Civil
Service, the railways, and the shift of administration to Delhi. To what degree did this new era constitute a tabula rasa, and to what degree did regional patterns set in the eighteenth century persist?

We are able to answer this question with some empirical accuracy, as British colonial administration generated a rich source of data from across the continent as official statistics were compiled from across 349 colonial districts, allowing for analysis of regional on a wide range of measures of fiscal capacity, delivery of public goods, and the rule of law. For example, if we compare a map of land taxation in India in 1900 (Figure 6.2) with the map of precolonial state centralization at the start of chapter 3, then one fact is striking: that patterns of taxation in colonial India almost entirely match the patterns of state strength found in the preceding century. That this is so should not be surprising; colonial authorities worked with the institutions that they inherited, including both administrative personnel but also the degree to which precolonial states had succeeded in warring down local chieftains and landlords and establishing the right of their officials to collect revenue directly from tenants.

*Comparison of Princely States and Presidencies*
Table 6.1 shows levels of taxation across the various presidencies and princely states of British-ruled India at the start of the twentieth century. What is striking is that the areas of greatest political development in the late eighteenth century, are precisely those in which the extractive capacity of the colonial state was highest a century later. The various former Maratha territories lead the list, evidence of the fiscal and bureaucratic reforms of the eighteenth-century peshwas: in Indore and Gwalior, taxation per capita stands in the range of 5-6 Rs. per capita, some three to four times higher than the 1.39 Rs. per capita estimated in the surrounding regions of the Central Provinces; even in Bombay Province, which fell under direct British rule, the rate of 3.2 Rs. per capita is two-and-a-half times greater. They are followed by the “Afghan” states of Junagadh and Bhopal, states whose princes were scions of Afghan invaders of the sixteenth and seventeenth centuries; here taxation per capita is in a similar range (3.75 to 6.7 Rs). These perhaps provide a snapshot of what might have been had the Mughal areas managed to maintain or consolidate their polity unity. Yet what is generally noticeable is that all of the centralizing states of southern India top this list: both the Mysore Kingdom and Madras Presidency (the latter about fifty percent composed of former Mysore territories that were ceded to the Company); Hyderabad; and Travancore. These figures, moreover, reflect differences in fiscal efficacy rather than differences in economic productivity; agricultural output was higher in the Central Provinces, for example, than most other regions of the country, yet taxation per capita was a mere third of the level in much of the south of India. The simplest explanation is pure persistence. As Gordon (1993) remarks in the context of the Maratha kingdom, its system of “tax collection, record keeping, and administration carried on well into the colonial period and beyond”.

A similar picture emerges if we look at the provision of policing. Table 6.2 shows the density of police officers in provinces and states of India in 1904, expressed in terms of the number of population per police officer. The average in southern India is much lower (indicating greater police numbers) than in the north; at 1 officer per 815 population versus 1 per 1304. However these differences may also reflect a legacy of direct versus indirect rule, as it is the native states which show much greater density of policing than the areas under British administration, and the principle native states which are disproportionately concentrated in southern regions. As mentioned in the previous section, the preservation of
<table>
<thead>
<tr>
<th>Category</th>
<th>State</th>
<th>Tax Level Per Capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-Maratha States</td>
<td>Indore</td>
<td>6.35</td>
</tr>
<tr>
<td></td>
<td>Gwalior</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td>Bombay Presidency</td>
<td>3.21</td>
</tr>
<tr>
<td>Afghan States</td>
<td>Junagadh</td>
<td>6.70</td>
</tr>
<tr>
<td></td>
<td>Bhopal</td>
<td>3.75</td>
</tr>
<tr>
<td>Mysore Kingdom and Territories</td>
<td>Mysore</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>Madras</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>Travancore</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>Sikkim</td>
<td>2.68</td>
</tr>
<tr>
<td>Mughal Successor States</td>
<td>Hyderabad</td>
<td>4.21</td>
</tr>
<tr>
<td></td>
<td>Kashmir</td>
<td>3.20</td>
</tr>
<tr>
<td></td>
<td>Udaipur</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>Jaipur</td>
<td>2.44</td>
</tr>
<tr>
<td></td>
<td>Punjab</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>Bastar</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>Bengal Presidency</td>
<td>1.41</td>
</tr>
<tr>
<td>Tributary Regimes</td>
<td>Manipur</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>Central Provinces</td>
<td>1.39</td>
</tr>
</tbody>
</table>

Table 6.1: Tax Revenue across Units of Colonial India, Per Capita (Rs.), 1904 Post-Maratha States. Imperial Gazetteer (1905-6).
local militia by princely rulers is an important factor behind this discrepancy.

**Persistence to the Post-Colonial Era**

However imperfectly, by the dawn of independence the British colonial state had begun to consolidate its governance over directly-ruled India. The ranks of the Indian Civil Service had begun to grow, and the middle layer of administrative officialdom, so sorely lacking for a century or more, began to fill again the towns and cities of the Ganges plain. The arbitrary use of discretion, once widespread, was no longer tolerated: the Jalliwallah Bagh massacre of 1919, in which a British officer fired on and killed innocent civilians in
the city of Amritsar, became a major scandal, unlike the countless such incidents that could be documented a century earlier. Above all, the British colonial state had begun to achieve its goals on two dimensions, which were the consolidation of the judicial system, and the reform of taxation.

However, if the colonial state erected a new administrative trellis in the areas under its direct rule, its foundations did not run deep. Such fragility is demonstrated by the decline of state capacity across regions of northern India in the decades after 1947, in contrast to consolidation across regions that succeeded the former challenger states. This was not simply a case of direct against indirect rule: several areas under direct rule, such as Tamil Nadu, saw improvement after independence, while others that were largely under indirect governance, such as Rajasthan, instead underwent deterioration. Rather such patterns must be understood as a more distant legacy of the manner in which precolonial polities generated patterns of fiscal compliance, public service delivery, and rule of law that persisted under the aegis of the colonial regime, and that became newly salient following the retraction of colonial officers and the sudden transfer of responsibility to local elites.

*Reversion of State Capacity Following Independence*

In the post-independence era, state capacity has declined across much of the center-northern regions of India where the British had long struggled to build a colonial state. This began with an immediate loss of personnel, as senior positions in the Indian Civil Service (ICS) had been occupied by British colonial officers, who were now recalled to London; the “spoil” of appointing their successors set a bad precedent for the future politicization of such posts.

One of the attributes of this trend has been the outright politicization of the courts, constabulary, and the civil service after the advent of Congress:

“The democratic mandate exacerbated the erosion of judicial and administrative independence, as elected ministers felt justified in using the apparatus of the state as their personal fief; this attitude is summarized well by K.P. Sharma, a police officer from Madhya Pradesh, who notes in his memoirs: ‘It is we who elect the MLAs. The MLAs elect the Chief Minister. Therefore the Chief Minister and the ministers... are our servants. And the government servants are the servants of our servants whether it is the Collector or the SP [Superintendent
of Police]... Hence the servants of our servants can never be our masters!’ ”

This “democratization” of the state – or clientelization of the state, as public offices became subject to the patrimonial discretion of elected rulers – was far worse in areas that had been undergoverned at the point of British suzerainty, and in which colonial rulers had simply constructed new institutions rather than build upon those already in place. This of course was the case in much of the Gangetic north, from Bihar to Uttar Pradesh; it was also the case across the broad central provinces where Bhonsle princes had exerted but little control over their myriad tributaries.

One indicator of this erosion of state capacity is the neglect given to policing and the protection of property rights in democratic India. While the provision of law and order is often considered one of the most “basic” public goods, there is no reason why it would necessarily rank as the first priority in attracting political support: for one thing, there are few votes in enforcing property rights in a polity in which the electorate itself is largely without property to begin with. The poor do not use India’s courts because they cannot afford to, while the powerful do not use the courts because they do not have to. For another, politicians who themselves break the law may deliberately wish to maintain a weakened police and judicial branch, so as to better maintain influence over the latter through their personal patronage and discretion. Neither at the level of electoral politics, nor elite politics, was there an incentive to maintain the independent judicial institutions of the colonial era.

India’s democratic politicians have consistently neglected spending on universal public goods, in favor of spending on direct sources of patronage such as jobs, energy and fertilizer subsidies, and outright provision of food or festivities (Keefer and Khemani 2004, Piliavsky 2012). Since the 1960s, law and order has been especially de-emphasized as a priority for elected leaders. This is evident from the decline of per capita police numbers across many states (Table 6.3). Bengal was in many ways the laboratory for the establishment of British law in India, yet in West Bengal the ratio of police officers has more than halved in the past generation, from one per 741 inhabitants in the 1970s, to just one per 1,511 population today. In Uttar Pradesh, this ratio has fallen from one per 943 people to one per 1,284 over the same period, and in Madhya Pradesh, from one per 826 to just one per 1,353
population. It is perhaps Bihar, however, that stands out as having had the most grievous fall in police capacity: from an already slight one officer per 1,111 inhabitants in the 1970s, to just one for every 1,915 persons, in the most recent data. Together with the neglect of police numbers, there has also been a neglect in police training and pay: police have often been last in the political “feeding order”, and by 1979, unhappiness over pay, conditions, and political interference led to a wave of police strikes – in Punjab, Haryana, Chandigarh, Himachal Pradesh, Andhra Pradesh, Bihar, Delhi, Gujarat, Kerala, Maharashtra, Orissa, Tamil Nadu, and Rajasthan (Wilkinson 2004: 80).

A similar picture is to be found in the efficacy of the judicial system. A major pre-occupation of British colonial authorities was the enforcement of the courts. One of the first policies of the English East India Company was to establish a court system capable of enforcing tax collection and protecting the commercial interests of investors in urban manufacturies: under Hastings (governor general from 1813-1823) eighteen new courts were established in Bengal and their authority was extended from the cities to rural areas, the court in Calcutta established as supreme court, and a major effort was made to codify Islamic and Hindu law into a single corpus of legal tradition. This became the model for British direct rule across India and was copied in Madras, Bombay, and the Punjab. In many respects, British administration was a “judicial empire”, operating via the judicial rather than via the executive and administrative branches of government. Compared to its top-heavy Mughal predecessors, the density of civil service staff was low, and in Bengal and Bihar, the permanent settlement had in any event reduced the need for direct administrative involvement in tax collection; “civil servants who grew up in the Bengali tradition normally disapproved of all measures which executive intervention and tended to rely on the workings of the courts” (1986: 242). The district officer operated more as a district magistrate than as an executive of local government. Beyond the nightwatchman functions of law and public order, these functions fell instead to local and informal systems of authority.

Under colonial institutions, courts were expected to prosecute and convict criminal cases within four months of charges being filed, but since independence few states have managed to maintain anything close to this target. In Bihar and Jharkhand, already just 35 per cent of court cases were settled in less than a year in the 1960s; today, this has halved to
<table>
<thead>
<tr>
<th>State</th>
<th>1970s</th>
<th>2010</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jammu and Kashmir</td>
<td>369</td>
<td>258</td>
<td>+43.1%</td>
</tr>
<tr>
<td>Kerala</td>
<td>1124</td>
<td>851</td>
<td>+32%</td>
</tr>
<tr>
<td>Uttarakhand</td>
<td>943</td>
<td>710</td>
<td>+32.9%</td>
</tr>
<tr>
<td>Goa</td>
<td>450</td>
<td>347</td>
<td>+29.7%</td>
</tr>
<tr>
<td>Arunachal Pradesh</td>
<td>481</td>
<td>388</td>
<td>+23.8%</td>
</tr>
<tr>
<td>Haryana</td>
<td>719</td>
<td>590</td>
<td>+22%</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>1031</td>
<td>882</td>
<td>+16.9%</td>
</tr>
<tr>
<td>Karnataka</td>
<td>1010</td>
<td>919</td>
<td>+9.9%</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>671</td>
<td>664</td>
<td>+1%</td>
</tr>
<tr>
<td>Punjab</td>
<td>578</td>
<td>578</td>
<td>0.0%</td>
</tr>
<tr>
<td>Delhi</td>
<td>250</td>
<td>251</td>
<td>-0.5%</td>
</tr>
<tr>
<td>Tripura</td>
<td>338</td>
<td>344</td>
<td>-1.8%</td>
</tr>
<tr>
<td>Andhra Pradesh</td>
<td>1124</td>
<td>1215</td>
<td>-7.5%</td>
</tr>
<tr>
<td>Mizoram</td>
<td>246</td>
<td>269</td>
<td>-8.3%</td>
</tr>
<tr>
<td>Chhattisgarh</td>
<td>826</td>
<td>925</td>
<td>-10.7%</td>
</tr>
<tr>
<td>Himachal Pradesh</td>
<td>546</td>
<td>708</td>
<td>-22.8%</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>943</td>
<td>1284</td>
<td>-26.5%</td>
</tr>
<tr>
<td>Manipur</td>
<td>190</td>
<td>265</td>
<td>-28.1%</td>
</tr>
<tr>
<td>Rajasthan</td>
<td>746</td>
<td>1069</td>
<td>-30.2%</td>
</tr>
<tr>
<td>Odisha</td>
<td>962</td>
<td>1423</td>
<td>-32.4%</td>
</tr>
<tr>
<td>Madhya Pradesh</td>
<td>826</td>
<td>1353</td>
<td>-39.0%</td>
</tr>
<tr>
<td>Meghalaya</td>
<td>292</td>
<td>492</td>
<td>-40.6%</td>
</tr>
<tr>
<td>Bihar</td>
<td>1111</td>
<td>1915</td>
<td>-42.0%</td>
</tr>
<tr>
<td>Gujarat</td>
<td>649</td>
<td>1148</td>
<td>-43.4%</td>
</tr>
<tr>
<td>Assam</td>
<td>599</td>
<td>1086</td>
<td>-44.9%</td>
</tr>
<tr>
<td>Sikkim</td>
<td>204</td>
<td>370</td>
<td>-45.0%</td>
</tr>
<tr>
<td>West Bengal</td>
<td>741</td>
<td>1511</td>
<td>-51.0%</td>
</tr>
<tr>
<td>Nagaland</td>
<td>83</td>
<td>351</td>
<td>-76.4%</td>
</tr>
</tbody>
</table>

Table 6.3: Police Intensity (Inhabitants per Officer), 1970s to 2010.
just 16.2 per cent of cases in Jharkhand, and 14.5 per cent of cases in Bihar. In Maharashtra, as few as 21 per cent of cases were dealt with in under one year the 1960s; yet just 11 per cent are today. Perhaps the most grievous fall is in Uttar Pradesh, however, where in 1966, just 13 per cent of cases were settled in less than a year, a level that was already among the lowest in India; by 2012, however, this has fallen to just 3.4 per cent, or one in every 29 cases brought. These changes are shown in Table 6.4 below.

As the colonial state has passed into history, more and more have patterns of regional governance across India reverted to the pattern in place before the colonial state existed. In the areas of the subcontinent where centralized “challenger” regimes ruled in the eighteenth
century, notably Karnataka (Mysore), Maharashtra (the Maratha Peshwa), and Kerala (Travancore), cohesive elites and state bureaucracies have promoted developmental agendas for growth and development (Kohli 1987). Yet in the areas where the decline of the Mughal imperium left fragmentary, feudatory regimes, contemporary governance resembles a pattern found elsewhere; from “patrimonial,” to colonial, to “neopatrimonial” rule (Bratton and van der Walle 1994).

Conclusion

This chapter has sought to show the mechanisms by which precolonial institutions persisted through to the colonial era. The experience of colonialism in India is widely portrayed as an exogenous event, in which the indigenous institutions of a major world region were suddenly replaced by those of a colonial power. Yet British domination of India was not a sudden ‘accident,’ a decision made “in a fit of absence of mind” - as famously expressed by Victorian historian John Seeley - but a costly, patient, and determined enterprise requiring more than a century of wars, alliances, negotiations, and the cumulative mobilization of hundreds of thousands of soldiers into battle. As the subcontinent came gradually under British rule, starting with the coastal ports of Bombay, Madras, and Calcutta, moving on to the administration of Bengal from 1764, Madras as a result of the seven years war of 1756-63, Mysore in 1780-1799, the Deccan in 1818 and the Northwest in 1848-56, indigenous institutions were often preserved, in particular at the local level, while the colonial state operated as a relatively thin skein that was superimposed.

Continuity occurred due to several reasons. The first is the persistence of the administrative infrastructure of precolonial regimes under the aegis of colonial administrations, a fact that was clear in the case of the continuity of Maratha offices under the Bombay presidency and the continued reliance of northern districts on zamindars for the collection of revenue. The second is the role of organized resistance to colonial reform, a fact that became clear following the uprising of 1857, and left the colonial regime unwilling to attempt further reforms, and to reverse many of those it had attempted. Third, colonial institutions were unable to alter entrenched patterns of conduct at a local level, even where formal in-
stitutions were altered. The clearest examples of this were the persistence of exploitative landlordism in the North, and the resistance of local communities in former zamindar areas to raise levies for collective provision even after decentralization reforms in the late nineteenth century.

While social scientists have begun to explore the colonial origins of local-level variation in governance across India today, there remains a neglect to the legacies left by precolonial regimes, and a tendency to see causally endogenous factors such as caste fractionalization, strength of collective identity, differences in colonial policy, or the structures of civil society as exogenous (Banerjee and Iyer 2005). This chapter has used historical work on regional variation in governance during the colonial era and official statistics from colonial archives to show that precolonial structures persisted into the colonial era (Bayly 1988, Ramusack 2003, Stein 1996). By showing that the patterns of governance in the late nineteenth century reflected geographic discrepancies in state centralization from before the colonial era, this chapter points to a deeper historical point of divergence.
Chapter 7

The Global Perspective

Having demonstrated how historical state legacies foster higher compliance and state capacity across regions of India, I will now examine the validity of this relationship at a global, cross-country level. Do countries with early state legacies exhibit greater capacity today? And can we see a similar within-country relationship in other countries and cases? In answering these questions, I proceed in two stages. In this chapter, I examine the global cross-country data, in order to investigate the more general relationship between contemporary state capacity and historical legacies of state formation. In the next chapter, I will then engage in subnational analysis of trends in state capacity in a single large multi-ethnic polity, the Soviet Union, and in particular its Russian successor.

In the rest of this chapter, I will show that early state formation is among the best indicators of state capacity at the cross-country level today, ahead of alternative explanations such as economic development or colonial legacies. I begin by showing that measures of cumulative state presence up until 1750 or 1500 are significantly associated with variation in contemporary state capacity, before showing that precolonial state formation, measured using the State Antiquity Index, is also predictive of improvements in postcolonial state capacity in recent decades. This suggests that in the postcolonial era, earlier legacies of state formation have reasserted themselves over the legacies of institutional development occurring under colonial rule.
Defining and Measuring State Capacity in Global Comparison

Before we can compare countries on their degree of contemporary state capacity on a worldwide basis, it is first necessary to define and operationalize the concept in a manner suitable for cross-country analysis. State capacity can be understood as the institutional infrastructure that allows elites to implement policies and decisions, ranging from taxation (fiscal capacity), to legislation (judicial, surveillance and policing capacities), to the central administration and formulation of policy (bureaucratic capacity). In selecting indicators to measure each of these areas, I have also given attention to avoiding ‘perception-based’ measures and instead selecting where possible only actionable measures based on underlying data (Knack 2004). These indicators are summarized in Table 7.1.

First, perhaps at the core of any definition of the state is Weber’s distinction between bureaucracy and patrimonialism. In Economy and Society Weber (1904) provided a list of ten criteria of bureaucratic development, including selection on the basis of technical qualifications, remuneration by fixed salaries, and the possibility of career progression. In these ways, discretion over hiring, remuneration (e.g. via abuse of prebends), and work behavior are taken out of the hands of the individual bureaucrat, and instead inscribed in a series of fixed rules that govern their actions. Thus an initial index of Weberian bureaucracy can be constructed from a set of six items that specifically concern the recruitment, training, and efficacy of the central civil service: an assessment of meritocratic recruitment; civil service career opportunities; and bureaucratic salary and compensation, taken from the Evans and Rauch (1999) dataset; a measure of “professional public administration” taken from Teorell, Dahlstroem and Dahlberg (2011), which combines meritocratic recruitment, nepotism, and bureaucratic autonomy; and World Bank ratings on the quality of public administration and the quality of project budgetary management from the Country Policy and Institutional Assessment (CPIA) exercise (World Bank 2013). All indicators are drawn from the Quality of Government database at Gothenburg University (Teorell et al. 2013).

Second, fiscal capacity is often considered at the core of what Mann (1986) has termed the “infrastructural power” of the state, and concerns its ability to raise resources through taxation. For fiscal capacity, two indicators are selected - the CPIA rating for “efficiency of revenue mobilization,” and the country credit rating (Teorell et al. 2013). Though both
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Indicators</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weberian Bureaucracy</td>
<td>Meritocratic recruitment</td>
<td>Evans and Rauch (1999)</td>
</tr>
<tr>
<td></td>
<td>Civil service career opportunities</td>
<td>Evans and Rauch (1999)</td>
</tr>
<tr>
<td></td>
<td>Remuneration relative to private sector</td>
<td>Evans and Rauch (1999)</td>
</tr>
<tr>
<td></td>
<td>Professional public administration</td>
<td>Teorell, Dahlstroem and Dahlberg (2011)</td>
</tr>
<tr>
<td></td>
<td>CPIA score for quality of public administration</td>
<td>World Bank (2013)</td>
</tr>
<tr>
<td>Fiscal Capacity</td>
<td>Country credit rating</td>
<td>World Economic Forum (2012)</td>
</tr>
<tr>
<td></td>
<td>CPIA score for efficiency of revenue mobilization</td>
<td>World Bank (2013)</td>
</tr>
<tr>
<td>Control of Corruption</td>
<td>Percentage of population, have paid a bribe in any form</td>
<td>Treisman (2007)</td>
</tr>
<tr>
<td></td>
<td>Percentage business leaders, common to pay irregular payments</td>
<td>Treisman (2007)</td>
</tr>
<tr>
<td></td>
<td>Percentage population, asked or expected to pay a bribe</td>
<td>Treisman (2007)</td>
</tr>
<tr>
<td>Rule of Law</td>
<td>Log Homicide Rate (50% wt)</td>
<td>United Nations Office on Drugs and Crime (2013)</td>
</tr>
</tbody>
</table>

Table 7.1: Cross-Country State Capacity Indices. The State Capacity index is equal weighted across four components: Weberian bureaucracy, fiscal capacity, (control of) corruption, and the rule of law. All indicators are proximally taken from the Quality of Government Database (2013), with the exception of homicide rates, taken from the United Nations Office on Drugs and Crime (UNDOC 2013).
broadly correlate with the overall level of tax revenue as a percentage of GDP, that may be a poor indicator because - as Fukuyama (2013) has observed - countries with equal levels of fiscal capacity may still vary in regard to their taxation preferences, while countries enjoying natural resource rents may be able to raise large amounts in tax without building a sophisticated fiscal infrastructure. On the other hand, the CPIA rating is designed to assess the fiscal system directly, including the width of the tax base, the degree of fiscal compliance, and the balance between direct and indirect taxation, and thus comes closer to a true measure of revenue mobilization in the sense implied here. Meanwhile, though the country credit rating reflects a form of perception, it reflects a very specific type of perception - namely, regarding a country’s ability to manage its debts, and to raise the resources required to reduce them suddenly if required. As such, it may be less prone to correlated error than other perception-based measures, and also more reliable: countries whose system of taxation raises large revenues from natural resources may acquire a very low rating if their fiscal system is deemed inadequate to cope with any sudden fall in commodity prices, and countries with a preference for low taxation, such as Switzerland or Japan, can attain a very high rating if their fiscal system is judged capable of an increase in revenue mobilization when required. It might also be added that a great deal is at stake in the assessment of creditworthiness, and that while ratings agencies are subject to various kinds of bias, they are very unlikely to rate countries higher simply because of the openness of their institutions, except insofar as is relevant to their debt management capacity.

Third, corruption is a central aspect of state capacity, because it reflects the control of state principals over its agents. Corruption, defined as the “use of public office for private gain,” covers such categories as embezzlement (the outright theft of public resources), nepotism (the use of power in order to secure jobs and contracts for friends and relatives), bribery (offering or demanding payments in exchange for public provision services) and expropriation (the use of legal privilege in order to acquire the property of others). For (control of) corruption, items from the Treisman (2007) corruption surveys are selected - namely, the proportion of the population that in the last 12 months has paid a bribe in any form, the proportion of business representatives that find it common to pay irregular payments, and the percentage of the population that in the last 12 months have been asked or expected to pay a bribe to government officials. Admittedly, this does leave out important
aspects of corruption such as nepotism, embezzlement and legal corruption, yet these are by definition difficult to measure and detect, and the absence of a good indicator does not necessitate recourse to a poor one. Measures such as the Corruptions Perceptions Index, the Freedom House corruption rating, and the Worldwide Governance Indicator for control of corruption have been avoided, because these are (or are largely composed from) perception-based measures, rather than within-country, “experience-based” indicators that survey actual knowledge of specific acts of corruption. Given the objective of analyzing whether and how dimensions of governance relate to one another, it is necessary to avoid “halo effects” that might produce correlated bias among our indicators and overstate the degree of covariance between dimensions.

Finally, rule of law is an important aspect of state capacity as it concerns the ability of the state to legislate, govern, and command an effective monopoly of coercion over its territory and subjects. The concept of the rule of law encompasses two related, though not commensurable, sub-concepts - public order and equality before the law - and the indicator used is an equal-weighted composite of these. As a measure of public order, the best general proxy is the log homicide rate, taken from the United Nations Office on Drugs and Crime: unlike in the case of many other crimes, the evidence of murder is difficult to hide, and deaths from violence are relatively easy to classify (UNODC 2013). As indicators of equality before the law, the Cingranelli-Richards measure of independence of the judiciary and the World Economic Forum survey of judicial independence are selected - though with reservations. Both are unsatisfactory and ideally we would want “harder” comparative data such as the average length of court cases, the cost of justice, and disparities in conviction rates between poor and wealthy defendants. Nonetheless, they are relatively specific in what they claim to assess, and as in the case of corruption, “generic” perception-based measures are avoided, for broadly similar reasons: these include the Bertelsmann Foundation Rule of Law index, the Freedom House Rule of Law measure, and also the Worldwide Governance Indicator for Rule of Law (which includes many non-perception-based sources).
Examining the Historical Origins of State Capacity on a Cross-Country Basis

Why do countries vary today in their ability to exert state capacity? A first set of theories, accentuated in this study, emphasizes the role of contingent historical processes and events in the accumulation of state capacity, leading to durable legacies over time (Ertman 1997, Tilly 1992, Fukuyama 2011). Tilly (1992) and Skocpol (1979) suggest that events such as war and revolution may lead to incremental improvements in state capacity over time, subject to a “ratchet effect”: gains in state capacity are not given up because of the benefits that accrue to elites from the new state infrastructure, or because states create vested interests (such as a bureaucratic class) which retains an interest in maintaining and expanding their capabilities. A similar perspective can be found in the historical work of Max Weber (1922), as well as in contemporary studies of bureaucratic autonomy which argue that state elites can be relatively autonomous from social pressures and guided by a logic that is independent of dominant economic interests, but rather guided by the logic of statehood (Evans 1995). The key determinants of state formation, in this view, are time - which increases the possibility of state capacity augmenting events, such as interstate wars that require total fiscal and manpower mobilization - and the luck not to be dominated or destroyed by one’s neighbors.

Against this argument, however, we must consider several alternative hypotheses. For example, modernization arguments have sustained that long-run socioeconomic development assists states in building their central coercive capacity. Technological changes - in particular, improvements in transportation and communication - can reduce the costs for central elites of monitoring, surveillance, and enforcement. Advances in weaponry significantly increase the cost of warfare, requiring more sophisticated fiscal infrastructure, and leading to increased demands of the state upon society (Roberts 1967, Parker 1996)\(^1\). Starting from the advent of printing, advances in recording technologies enable record-keeping, necessary for central systems of law and taxation, while improvements in transportation

\(^1\)The “gunpowder hypothesis” has been given a sharp rebuke by military historian Azar Gat (2008), who contends that gunpowder weaponry did not entail larger costs or lead to the end of military fortifications, but merely a shift from the castle to the trace italienne.
enhanced the ability of central bureaucrats to administer control over dispersed agents of the state (Diamond 1991). Thus economic development - and the ability of elites to command sufficient financial resources in order to invest in new political techniques - might also be expected to raise the level of state capacity.

Another argument that has a long pedigree in discussions of the origins of states is the “reductionism” of state capacity to environmental factors. Many of the classical theories of state formation such as Wittfogel (1957) and Carneiro (1970), suggested such an environmental determinism: Wittfogel argued that the irrigation agriculture of the Nile or Euphrates rivers gave Egyptian and Babylonian elites control over their subject peoples, allowing the construction of the first proto-states, while Carneiro’s “circumscription” theory argued that the surrounding desert made it impossible to flee the accumulation of despotic rule. More contemporary theories have also cited environmental factors: Herbst (2000) has attributed the weak state capacity of African states to low population densities following centuries of slave-raiding, while Scott (2009) attributes the low state capacity of upland Southeast Asia to the region’s mountainous territory, into which oppressed peoples would flee. Similarly, Fearon and Laitin (2003) have suggested that one of the key factors in determining whether states decide to consolidate their monopoly of force is whether rugged terrain raises the cost of doing so. While these viewpoints are diverse, they agree that rivers and plains make it easier for centralized polities to emerge and consolidate their rule, while mountainous territory is inimical to state-building.

Finally, a range of theories from within new institutional economics have revived the earlier perspective of world-systems theory, that it was not geography but rather the intervention of western colonialism that has determined - for the most part, negatively - the state capacity of developing countries (Wallerstein 1974). Acemoglu, Johnson and Robinson (2002) for example suggest that areas ruled but not settled by Europeans were subject to “institutional reversal” as European powers established ‘extractive institutions’ in order to exploit the local population (Engerman and Sokoloff, 1994). Other traditions in political science have also stressed the role of colonial institutions, but with regard to the specific policies of colonial powers. While the British colonial legacy has sometimes been seen as ambiguous due to its tendencies to rely on indirect rule (Gerring et al. 2010) or for introducing tentative democratic institutions (Wilkinson 2013) - the French and Spanish legacy
have largely been seen as negative due to the introduction of Roman law (La Porta et al. 1997). Meanwhile, Japanese rule has been argued to be a cause of bureaucratic and fiscal development (Kohli 2004).

**Operationalizing the Causes of State Capacity**

In order to test for the potential causes of long-run state capacity, each of the hypotheses is operationalized as follows.

1) As a measure of *geographic suitability* for state formation, this paper uses the “terrain ruggedness” variable published by Nunn and Puga (2012). The ruggedness variable is based on the terrain variable of Riley, DeGloria, and Elliot (1999) to quantify topographic heterogeneity in wildlife habitats, and is a useful proxy for whether a landscape provides a useful environment for non-state actors to resist centralized rule. Consistent with the theories of Wittfogel (1957) and Carneiro (1970), countries such as Egypt and Iraq have a low ruggedness score (0.72 and 0.67 respectively, relative to a 1.37 average), while states such as Afghanistan (2.52) and Yemen (2.32) are geographically problematic from the perspective of centralized state-building.

2) The *modernization thesis* that economic development is the cause of accumulated state capacity is more difficult to test: while GDP per capita can tell us the ability of elites to invest in state-making technologies, economic development is obviously endogenous to processes of state-building (Putterman et al. 2002, Kohli 2004). Therefore this paper instruments for economic resources using the per capita level of hydrocarbon (oil and gas) exports. In the postcolonial era, countries such as the Gulf states received a large exogenous boost in terms of resource rents from hydrocarbon exports, and this has allowed them to invest in western consultancy, surveillance technology, and education and training for their administrative bureaucracy. We might therefore expect such states to exhibit higher state capacity as a result of this investment, in accordance with the modernization hypothesis. An established hypothesis in political science is that the energy resources of Gulf States and other hydrocarbon exporters has allowed elites in these countries to invest in state capacity, and therefore build stable autocratic regimes in place of democratic transition (Ross 2001,
3) As a measure of *historical warmaking*, this article takes the proportion of time (in years, months and days) spent engaged in interstate war in the period from 1800-1970, as reported in the Correlates of War dataset. While it would be preferable to include earlier wars, the measure is constrained by the start of the data series in 1800. Only interstate wars are taken, as intrastate conflicts may vitiate rather than help build state capacity.

4) To measure the effect of *colonial legacies*, separate dummy variables are included for a legacy of French, Spanish, British and Japanese rule. Furthermore, in addition, in order to distinguish between ‘settler’ and ‘extractive’ colonies, a variables is also included for the proportion of the population that is of European descent (Acemoglu, Johnson and Robinson 2001).

5) Next, to measure the cumulative effect of *early state formation*, a state antiquity variable is constructed from the latest Putterman (2013) dataset. Putterman and colleagues have coded an aggregate score for each 50-year period since 0 AD for the territory of each of today’s states, based on three criteria: i) whether a state existed at that time; ii) the degree to which the state was indigenous or foreign-ruled; and iii) the degree of territorial contiguity with the contemporary state existing in that area. These are aggregated with a discount rate to weight more recent state legacies more heavily over those situated in the distant past. For the state antiquity variables used in this article, in order to limit endogeneity from modern state capacity to state survival in the modern age, we “cap” the accumulation period at 1500 and 1750, respectively, and report results in separate tables (7.2 and 7.3). In this way, the state history variable reflects the accumulation of contingent historical circumstances during the medieval era: whether a polity survived or was destroyed during the Mongolian invasions, for example, or the subsequent era of Islamic “gunpowder empires” (Streusand 2010).

6) Finally, because the collapse of *communism* is widely seen to have resulted in a widespread loss of state capacity in many post-communist (and above all, post-soviet) societies, a dummy variable is included for whether a country has recently experienced
Results

Results are shown in Tables 7.2 and 7.3. The estimated coefficients indicate that historical institutionalist views of state formation are well-supported by the data. Consistent with the ‘new institutionalist’ view, French and Spanish colonial legacies are associated with significantly lower state capacity, and across most dimensions. British colonialism is neither positive nor negative, though like French and Spanish rule, associated with greater corruption, which may reflect the tendency for post-independence elites to mimic extractive practices among former foreign rulers. Japanese colonialism, as argued by Kohli (2004), is positively and significantly associated with the development of fiscal capacity, and positively, though non-significantly, associated with other dimensions. Consistent with the arguments of Tilly (1992), while warmaking is not associated with bureaucratic development, rule of law, or corruption, it is strongly associated with fiscal capacity: states which have historically fought more wars have significantly better mechanisms for raising taxes and paying their debts. Next, European settlement is significantly positively associated with most aspects of state capacity, supportive of the view that Europeans brought state-building models and institutions with them. Finally, time, and survival, matter: states with long histories for the millennium that precedes the modern age (dated either in 1750 or 1500) are far more likely to have emerged centuries later with Weberian bureaucracies, rule of law, and fiscal capacity.

On the other hand, the findings do not appear supportive of theories rooted in geographic determinism or economic modernization. No association is found between geography and state formation, as more rugged states do not have lower state capacity than those with more governable river, valley and desert areas. And contrary to the modernization argument, nor is any association found between elite control of economic resources (instrumented using per capita energy exports) and the success of elites in developing state capacity: while a number of Gulf autocracies do appear to have invested energy resources in building effective state bureaucracies, outlier cases such as Equatorial Guinea and Angola suggest that this is not a necessary consequence, but at best an interaction with elite
<table>
<thead>
<tr>
<th>Covariate</th>
<th>INDEX</th>
<th>WEBER</th>
<th>FISCAL</th>
<th>CCORR</th>
<th>LAW</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP per capita (instrumented using carbon exports per capita)</td>
<td>0.038</td>
<td>0.067</td>
<td>0.14</td>
<td>-0.029</td>
<td>-0.006</td>
</tr>
<tr>
<td>Rugged Terrain Index (Nunn and Puga)</td>
<td>0.05</td>
<td>0.021</td>
<td>0.012</td>
<td>0.051</td>
<td>0.066</td>
</tr>
<tr>
<td>State Antiquity Index, 0-1750 AD</td>
<td>0.001**</td>
<td>0.002**</td>
<td>0.001*</td>
<td>-0.001</td>
<td>0.002***</td>
</tr>
<tr>
<td>Cumulative Interstate War Duration, 1800-1970</td>
<td>0.013</td>
<td>0.007</td>
<td>0.048**</td>
<td>-0.007</td>
<td>-0.004</td>
</tr>
<tr>
<td>Percent European origin</td>
<td>0.01***</td>
<td>0.006*</td>
<td>0.014***</td>
<td>0.007**</td>
<td>0.009***</td>
</tr>
<tr>
<td>Colonial legacy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French (0/1)</td>
<td>-0.459*</td>
<td>-0.324</td>
<td>-0.392</td>
<td>-1.513***</td>
<td>-0.567**</td>
</tr>
<tr>
<td>Spanish (0/1)</td>
<td>-0.569**</td>
<td>-0.91***</td>
<td>0.088</td>
<td>-0.691*</td>
<td>-1.093***</td>
</tr>
<tr>
<td>British (0/1)</td>
<td>-0.034</td>
<td>0.015</td>
<td>0.257</td>
<td>-0.976**</td>
<td>-0.11</td>
</tr>
<tr>
<td>Japanese (0/1)</td>
<td>0.825</td>
<td>0.88</td>
<td>1.38*</td>
<td>0.553</td>
<td>0.311</td>
</tr>
<tr>
<td>Recent Communist Regime Collapse (0/1)</td>
<td>-0.622***</td>
<td>-0.624**</td>
<td>-0.611*</td>
<td>-0.672***</td>
<td>-0.687***</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.621**</td>
<td>-0.356</td>
<td>-1.159***</td>
<td>0.27</td>
<td>-0.49**</td>
</tr>
</tbody>
</table>

| n                     | 122      | 113      | 115      | 84       | 122       |
| Adj. ²                | 0.46     | 0.26     | 0.45     | 0.43     | 0.56      |

Table 7.2: Covariates of State Capacity. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level.
<table>
<thead>
<tr>
<th></th>
<th>INDEX</th>
<th>WEBER</th>
<th>FISCAL</th>
<th>CCORR</th>
<th>LAW</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP per capita (instrumented using carbon exports per capita)</td>
<td>0.033</td>
<td>0.066</td>
<td>0.137</td>
<td>-0.028</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.055)</td>
<td>(0.083)</td>
<td>(0.081)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>Rugged Terrain Index (Nunn and Puga)</td>
<td>0.047</td>
<td>0.018</td>
<td>0.01</td>
<td>0.053</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.075)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>State Antiquity Index, 0-1500 AD</td>
<td>0.001**</td>
<td>0.002*</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0)</td>
</tr>
<tr>
<td>Cumulative Interstate War Duration, 1800-1970</td>
<td>0.014</td>
<td>0.008</td>
<td>0.049***</td>
<td>-0.007</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.016)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Percent European origin</td>
<td>0.01***</td>
<td>0.005*</td>
<td>0.014***</td>
<td>0.007**</td>
<td>0.009***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Colonial legacy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French (0/1)</td>
<td>-0.457*</td>
<td>-0.326</td>
<td>-0.397</td>
<td>-1.511***</td>
<td>-0.556**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.202)</td>
<td>(0.291)</td>
<td>(0.306)</td>
<td>(0.404)</td>
</tr>
<tr>
<td>Spanish (0/1)</td>
<td>-0.557**</td>
<td>-0.921***</td>
<td>0.083</td>
<td>-0.693*</td>
<td>-1.051***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.192)</td>
<td>(0.269)</td>
<td>(0.286)</td>
<td>(0.264)</td>
</tr>
<tr>
<td>British (0/1)</td>
<td>-0.045</td>
<td>-0.007</td>
<td>0.241</td>
<td>-0.967**</td>
<td>-0.111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.192)</td>
<td>(0.282)</td>
<td>(0.294)</td>
<td>(0.333)</td>
</tr>
<tr>
<td>Japanese (0/1)</td>
<td>0.851</td>
<td>0.909</td>
<td>1.399*</td>
<td>0.543</td>
<td>0.354</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.453)</td>
<td>(0.62)</td>
<td>(0.661)</td>
<td>(0.527)</td>
</tr>
<tr>
<td>Recent Communist Regime Collapse (0/1)</td>
<td>-0.624***</td>
<td>-0.616**</td>
<td>-0.606*</td>
<td>-0.672***</td>
<td>-0.688***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.158)</td>
<td>(0.221)</td>
<td>(0.238)</td>
<td>(0.195)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.621**</td>
<td>-0.231</td>
<td>-1.074***</td>
<td>0.234</td>
<td>-0.45**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.198)</td>
<td>(0.273)</td>
<td>(0.288)</td>
<td>(0.351)</td>
</tr>
</tbody>
</table>

| n | 122 | 113 | 115 | 84 | 122 |
|  Adj. $R^2$ | 0.46 | 0.24 | 0.44 | 0.43 | 0.58 |

Table 7.3: Covariates of State Capacity (Using State Anquity, Capped at 1500. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level.)
Examining the Accumulation of State Capacity Over Time

The results of Tables 7.2 and 7.3 are however static: but what about changes in state capacity over time? While the variables used to construct the four state capacity indices do not allow for time-series aggregation, the combined state capacity index correlates highly ($r = 0.84$) with the International Country Risk Guide (ICRG) measure for “quality of government,” which combines indices for bureaucratic quality, control of corruption, and rule of law, and is therefore conceptually commensurable. A time-series for the ICRG data exists from 1984-2008, with a larger sample size beginning in 1986. Table 7.4 therefore presents a series of regressions upon the change in state capacity from 1986 to 2008.

The control variable for state capacity at time $t_0$ (1986) is significant and negative, and this most likely reflects an artifact of the “upper bound”: countries with maximum state capacity in the 1980s were of course unlikely to see any increase in state capacity from this point. Consistent with the results of Tables 7.2 and 7.3, the collapse of communism is shown to have had a large and deleterious effect on state capacity, and can be shown here longitudinally: ICRG scores fell in most post-communist nations between the mid-1980s and the present day, reflecting the difficulties in maintaining rule of law and preventing widespread embezzlement of state resources during the transition era. Despite the large investments in bureaucratic infrastructure in countries like Qatar and Bahrain, exogenous growth driven by hydrocarbon exports does not appear to have led on average to increased state capacity.

One of the curious results in Table 7.4 is that a strong predictor of rising state capacity in recent decades, is a country’s legacy of state formation in the period before 1750. This association remains robust to a series of alternative specifications, and to use instead of the State Antiquity variable dated to 1500. Why would this association hold true? One explanation may be that in the period following decolonization, the end of the Cold War, and the break-up of the Soviet Union, countries have reverted to their prior indigenous institutional legacies: states in Central Asia and Africa with weak legacies of centralization and bureaucratic development, were more likely to revert to patronialism, whereas polities in East...
<table>
<thead>
<tr>
<th>Dependent variable: Change in ICRG Quality of Government Index, 1986-2008</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colonial legacy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French (0/1)</td>
<td>-0.039</td>
<td>-</td>
<td>-0.039</td>
<td>-</td>
</tr>
<tr>
<td>(0.059)</td>
<td>(0.059)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish (0/1)</td>
<td>-0.148**</td>
<td>-</td>
<td>-0.148**</td>
<td>-</td>
</tr>
<tr>
<td>(0.051)</td>
<td>(0.05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>British (0/1)</td>
<td>0.024</td>
<td>-</td>
<td>0.027</td>
<td>-</td>
</tr>
<tr>
<td>(0.056)</td>
<td>(0.056)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese (0/1)</td>
<td>0.192</td>
<td>-</td>
<td>0.19</td>
<td>-</td>
</tr>
<tr>
<td>(0.117)</td>
<td>(0.117)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recent Communist Regime Collapse (0/1)</td>
<td>-0.224***</td>
<td>-0.194***</td>
<td>-0.226***</td>
<td>-0.196***</td>
</tr>
<tr>
<td>(0.052)</td>
<td>(0.054)</td>
<td>(0.052)</td>
<td>(0.054)</td>
<td></td>
</tr>
<tr>
<td>GDP per capita (instrumented using carbon exports per capita)</td>
<td>0.007</td>
<td>0.009</td>
<td>0.008</td>
<td>0.011</td>
</tr>
<tr>
<td>(0.014)</td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Rugged Terrain Index (Nunn and Puga)</td>
<td>0.022</td>
<td>0.022</td>
<td>0.021</td>
<td>0.02</td>
</tr>
<tr>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>State Antiquity Index, 0-1500 AD</td>
<td>0.272*</td>
<td>0.377**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(0.126)</td>
<td>(0.127)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Antiquity Index, 0-1750 AD</td>
<td>-</td>
<td>-</td>
<td>0.31*</td>
<td>0.417**</td>
</tr>
<tr>
<td>-</td>
<td></td>
<td>(0.138)</td>
<td>(0.14)</td>
<td></td>
</tr>
<tr>
<td>Cumulative Interstate War Duration, 1800-1970</td>
<td>0.001</td>
<td>0.004</td>
<td>0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Percent European origin</td>
<td>0.003***</td>
<td>0.003***</td>
<td>0.003***</td>
<td>0.003***</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0)</td>
<td>(0.001)</td>
<td>(0)</td>
<td></td>
</tr>
<tr>
<td>ICRG Quality of Government Index, 1986</td>
<td>-0.997***</td>
<td>-0.997***</td>
<td>-0.997***</td>
<td>-0.998***</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.362***</td>
<td>0.323***</td>
<td>0.347***</td>
<td>0.306***</td>
</tr>
<tr>
<td>(0.055)</td>
<td>(0.034)</td>
<td>(0.058)</td>
<td>(0.037)</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4: Change in State Capacity, 1986-2008. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level.
Asia and Central-Eastern Europe with strong histories of indigenous administrative development continued on a trajectory towards effective state formation. The greater degree to which sovereignty has been reasserted in the modern age, the greater the importance of earlier indigenous traditions of statehood (Krasner 1999). Further evidence for this viewpoint can be adduced by looking at the relationship between historical state antiquity, and the timing of post-colonial independence: countries that acquired their independence first during decolonization, such as Iraq, India, Ethiopia or Malaysia, were those in which indigenous elites were able to mobilize against colonial powers; whereas independence came much later to colonies in which challenger elites did not exist as a result of historical states (Figure 7.1).

**Are Historical Factors Reducible to Geography?**

Can geographic reductionism be counted out, or has the mechanism been misspecified? An alternative reading the coefficients for geography in Tables 7.2- 7.4 is that historical factors may be proximate, but geographical factors are their ‘deep’ determinants. Perhaps colonial powers were more likely to focus on ‘easy terrain’ targets, for example, or state
antiquity in the period prior to 1750 or 1500 were already less likely to consolidate in
mountainous territory. In order to explore the possibility that geographic variables might
serve in this way as instruments for more recent historical outcomes such as state survival
(state antiquity) or the outcomes of colonialism, a set of potential “first stage” regressions
are conducted on each of the historical variables and reported in Table 7.5. In order to tease
out potential geographic causes, the full range of variables from Nunn and Puga (2012) are
included here on the right-hand side. Note that while the inclusion of both average distance
to the coast and the percentage of territory within 100km of the coast might raise concerns
over collinearity, the correlation between these two variables is only moderate ($r = 0.68$).

Table 7.5 offers little support for the view that historical state formation can be reduced
to geographic influences. The only significant association between the level of historical
state formation (whether by 1500 or by 1750) and geography is how closely a country was
situated to the tropics. An interesting subsidiary question arises as to whether, as Sachs
(2000) and others have argued, this may be linked to disease burden, or even as Mon-
tesquieu (2011 [1748]) suggested, to climatic conditions and culture. That question cannot
be answered here, but what can be said is this association is supportive of the argument
made by Gerner (2012), that when Europeans went to conquer the tropics after 1450, they
were conquering areas that already had weak states: non-European areas with stronger
states - such as China, Japan, or the Ottoman Empire - were defeated only much later, if at
all. However there is no relationship between early state formation and many of the factors
highlighted in the earlier “hydraulic” or “circumscription” theories of Wittfogel (1957) or
Carneiro (1970), such as terrain ruggedness or the extent of desert. In general, geographic
theories of suitability for state formation are not well-supported by the data.

Table 7.5 also shows that imperial powers were not only more likely to colonize areas
in the tropics, but also areas with lower soil fertility and more desert - though this largely
follows from the fact that most deserts and low quality soil areas are situated in the tropical
zone. Interestingly, though desert and poor soil areas were more likely to become colonies,
European settlement occurred disproportionately in zones with less desert and better soil;
this is consistent with the suggestion by Acemoglu, Johnson and Robinson (2001, 2002)
that the primary purpose of colonialism was extraction, but in those areas not suitable
for exploitation of population or minerals, a settlement strategy was chosen instead. Fi-
<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>State Antiquity, 0-1500 AD</th>
<th>State Antiquity, 0-1750 AD</th>
<th>Percent European Origin</th>
<th>Colonized (by any power)</th>
<th>Duration of Interstate War, 1800-1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruggedness Index</td>
<td>1.321</td>
<td>0.599</td>
<td>-8.69***</td>
<td>0.002</td>
<td>-0.135</td>
</tr>
<tr>
<td></td>
<td>(9.48)</td>
<td>(8.966)</td>
<td>(2.296)</td>
<td>(0.023)</td>
<td>(0.396)</td>
</tr>
<tr>
<td>Land Area</td>
<td>-0.055</td>
<td>-0.043</td>
<td>0.038*</td>
<td>0.000</td>
<td>0.011***</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.052)</td>
<td>(0.015)</td>
<td>(0.000)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Percentage of Fertile Soil (FAO / UNESCO)</td>
<td>-0.195</td>
<td>-0.191</td>
<td>0.569***</td>
<td>-0.003**</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(0.473)</td>
<td>(0.125)</td>
<td>(0.001)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Percentage of Desert Area</td>
<td>0.69</td>
<td>0.079</td>
<td>-0.912**</td>
<td>0.008*</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>(1.327)</td>
<td>(1.255)</td>
<td>(0.287)</td>
<td>(0.003)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Tropical Climate Index</td>
<td>-1.554***</td>
<td>-1.278***</td>
<td>-0.52***</td>
<td>0.007***</td>
<td>-0.051***</td>
</tr>
<tr>
<td></td>
<td>(0.263)</td>
<td>(0.249)</td>
<td>(0.067)</td>
<td>(0.001)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>% Within 100km of Ice-Free Coast</td>
<td>-0.315</td>
<td>-0.164</td>
<td>0.062</td>
<td>0.000</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.371)</td>
<td>(0.351)</td>
<td>(0.094)</td>
<td>(0.001)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Average distance to Coast</td>
<td>-61.247</td>
<td>-36.807</td>
<td>-13.524</td>
<td>-0.16</td>
<td>-0.094</td>
</tr>
<tr>
<td></td>
<td>(34.538)</td>
<td>(32.667)</td>
<td>(9.282)</td>
<td>(0.102)</td>
<td>(1.457)</td>
</tr>
<tr>
<td>Constant</td>
<td>268.284***</td>
<td>267.679***</td>
<td>43.998***</td>
<td>0.563***</td>
<td>4.307**</td>
</tr>
<tr>
<td></td>
<td>(37.894)</td>
<td>(35.841)</td>
<td>(9.501)</td>
<td>(0.097)</td>
<td>(1.606)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>145</th>
<th>145</th>
<th>162</th>
<th>188</th>
<th>137</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj r^2</td>
<td>0.227</td>
<td>0.161</td>
<td>0.388</td>
<td>0.392</td>
<td>0.277</td>
</tr>
</tbody>
</table>

Table 7.5: Hypothetical “First Stage” Regressions of Geographic on Historical Variables. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level.
nally, Table 7.5 shows that geographically larger countries in general have longer interstate wars. This reflects the established fact of military strategy that smaller states, such as Israel or Rwanda, must opt only into lightning conflicts and are vulnerable beyond this point, whereas larger states, such as Russia, China, and the United States, possess the option of engaging intercontinental and multi-theater conflict.

Overall, the coefficients in Table 7.5 suggest that geographic factors do appear to predict the pattern of events after 1450, namely the likelihood of European conquest and settlement; for each of these models the r-squared is 0.39, implying that two-fifths of variation in such outcomes might be reducible to geographic considerations, largely relating to the ease of conquest. However, in contrast to much of the literature on early state formation, geography offers a poor explanation of patterns of state development prior to the modern age. In the absence of further explanatory schema, contingency appears to be the overwhelming factor behind such early legacies. Moreover, contingency also remains the major of variation also in patterns of European conquest and settlement, even if non-stochastic factors account for a greater share of variance here.

**Searching for the Causes of Weak States: Does Colonialism Matter?**

The foregoing discussion has touched upon one of the more longstanding hypotheses for the weakness of postcolonial states, namely that western colonial rule undermined the development of state capacity (Amin 1972; Rodney 1972). This literature has suggested a wide variety of mechanisms, ranging from the deliberate policies of colonial rulers, to the disruption to indigenous institutions, to the accidental attributes of post-colonial states such as ethnic heterogeneity or territorial dispersion. Acemoglu, Johnson and Robinson (2002) maintain that areas ruled but not settled by Europeans were subject to “institutional reversal” as European powers established predatory regimes to exploit labor and natural resources and Engerman and Sokoloff (2002) argue the same with respect to the labor-rich areas of meso-america, such as Mexico, Peru, and Bolivia. This hypothesis is also supported in the findings of Melissa Dell (2010) regarding the mita in Bolivia and Peru. Though the labor and resource-extractive nature of colonial regimes may have entailed es-
establishing a minimal layer of state capacity, it also meant that legal institutions were not extended beyond the colonial elite, and the failure of colonial institutions to establish legitimacy beyond this circle resulted in pervasive non-compliance, resistance, and insurgency outside of the metropole (Herbst 2002). James Mahoney (2010) similarly argues that in the former Aztec and Inca domains of Mexico, Bolivia and Peru, colonial authorities engaged in mercantilist restrictions of trade, ownership and economic participation, resulting in an entrenched patrimonial elite unable to extend rule of law or secure the legitimacy among the wider population.

Another argument against the colonial state comes from Lange (2009), who, following Mamdani (1996) argues that the reliance by colonial regimes upon strategies of indirect rule left local power brokers entrenched, and that this resulted in fragmented postcolonial regimes dependent upon distributing patronage to local stakeholders, unable to enforce regulatory compliance, taxation or administrative control in many areas of their territory. According to Lange, direct rule entailed the creation of bureaucratic states with extensive territorial reach, while indirect rule relied upon local intermediaries and led to patrimonial states. However, other authors have also sought to question the benefits of direct rule, in particular in former French colonies; Kwon (2011) contends that in Cambodia and Laos colonial state institutions had adverse effects on post-colonial state formation, while La Porta et al (1998) argue that French directly-ruled areas of Africa were left with less legitimate and effective institutions than indirectly-ruled British Africa; a theory tested by and Lee and Schultz (2011), who show higher public goods provision on the formerly British than the French side of the border in Cameroon. Likewise in the context of colonial India, where both direct and indirect patterns of governance prevailed, Iyer (2010) argues that areas under direct rule fared significantly worse during the post-colonial era than those in which indigenous institutions were preserved through the princely states, a hypothesis tested by identifying princely states that narrowly fell to direct rule following a lapse in royal succession.

Finally, a wide number of authors suggest that colonisation left unviable states, as colonial borders were drawn around vast territories following arbitrary fiat, without respect to prevailing ethnic boundaries or considerations of geographic governability. As a result, post-colonial states were shared among large numbers of ethnic groups, increasing the cost
of collection action and requiring greater patronage in order to maintain governing coalitions (Alesina et al. 2004). In addition, the territorial span of the post-colonial state, above all in Africa, made the challenge of securing cooperation and administrative control over these groups even more difficult (Herbst 2002). Colonial regimes also deliberately exacerbated ethnic tensions: by relying on strategies of ‘divide and rule’ in order to create ‘favored’ ethnic groups, as in Rwanda (Mamdani 2002), or in India following the uprisings of 1857 (Varshney 2003), rulers made it more difficult to secure collective action from ethnic and sectarian minorities in the post-colonial era, requiring instead the distribution of rents (Keefer and Khemani 2004).

These varying arguments for the weakness of the post-colonial state can be summarized into three main categories, the limitation of scope, the displacement of indigenous state capacity, and the creation of unviable states.

i) The limitation of scope. Colonial regimes sought only to extract resources and tax, rather than provide public goods, fight wars, or implement a wide-ranging platform of societal transformation and behavioral regulation, in the fashion of post-revolutionary regimes. The colonial state was therefore a ‘thin’ state by design; a “gate-keeper” state based on the coast, with interest only in ruling and extracting natural resources rather than institution-building (Cooper 2002). Colonial regimes did not establish, or seek to establish, the state’s presence outside of the administrative capital or those areas which contained valuable resources. Nor did colonial regimes generally invest in building tax infrastructure, such as registration of income or cadastres of land, opting instead for ‘easier’ sources of revenue such as tariffs on trade or the use of commodity marketing boards (Bates 1981). Post-colonial regimes were left without ready sources of revenue, and without any means of accomplishing the ambitious transformative schemes of post-colonial leaders other than by resorting to the same distortionary apparatus and the recourse of international debt markets.

ii) The displacement of indigenous capacity. Colonial regimes relied for their functioning on a cadre of administrative officers provided by the colonial core state. Therefore, on independence many post-colonial regimes found themselves facing a severe shortage of capable bureaucrats. This lack of administrative capacity has been reinforced by the tendency
for post-colonial states to rely on their former colonial rulers and international donors and multilateral agencies for technical support, military defense, finances and soft loans, and in some cases, active suppression of threats to the ruling regime. The ready availability of soft loans and aid has hindered the responsibility to develop extractive fiscal capacity, while the availability of external technical support has inhibited the development of parallel bureaucratic capacities domestically (Easterly 2001, 2006).

iii) The creation of “unviable” states. Whereas European states formed endogenously to the conditions of state viability, i.e. geographic defensibility, ethnic cohesion, and access to resources, colonial regimes were formed arbitrarily by European powers with limited respect for existing ethnic boundaries or natural geographic conditions of governability. As a result, colonial regimes unintentionally left behind unviable states, either because ethnic fractionalization led to weak collective identities and clientelism (Alesina et al. 2004, Keefer and Khemani 2004), or because geographic conditions were unfavorable to territorial consolidation (Herbst 2000, Fearon and Laitin 2003).

In their most extreme form, they assert not only that colonial regimes stagnated or were ill-designed for independent statehood, but polities that were the most economically and politically advanced at the start of the modern era had become the weakest by modern times. States that were most developed in 1492 were the first to be targeted by European colonial powers: polities such as the Inca and Aztec Empires, the Malacca sultanates, and Indian Bengal were attractive to European powers due to their reserves of labour and mineral resources, and subsequently turned into extractive colonies for the purpose of mining and taxation. By contrast, the more sparsely populated territories of Australasia, Northern America and the African Cape offered fewer possibilities for colonial exploitation, and were instead turned into centers of settlement, inheriting European models of political and social organization. In this way the most powerful empires outside of Europe on the eve of the modern age became peripheral zones of the world economy, while once marginal territories in the Americas and Antipodes were granted secure property rights and transformed into sites of productive growth. Acemoglu et al. (2002) use data on city size and population density in 1500 (McEvedy and Jones, 1978) to show that the most densely settled areas and
the largest urban centers were to be found in the indigenous polities of the Americas, and in South and Southeast Asia.

Problems with the “colonial underdevelopment” thesis

There are, however, several problems with the thesis of colonial underdevelopment. First, theories that root the low state capacity of non-western countries in the legacy of colonialism lack a “counterfactual conditional” (King, Keohane, and Verba 1994) – a hypothetical alternative scenario in which the treatment, in this case colonial rule, had not been applied. What kind of state system would have emerged in the absence of the colonial enterprise? This is difficult to answer. Though we can try to imagine a counterfactual scenario in which European powers had agreed in 1885 to not interfere in the African subcontinent, there is no African territory that was left undisturbed by foreign powers, and even more difficult to imagine would have been the consequence of non-intervention in India, or Southeast Asia, or the Americas, given the length of time between initial European contact and their eventual independence. Moreover, while there are limited examples of non-European regimes that avoided colonial rule, notably China, Persia or Japan, it was often their high initial level of state capacity that allowed them to repel attempts at foreign incursion. This endogeneity between strategies of foreign rule and pre-existing indigenous state strength problematises any attempt at causal identification through simple cross-sectional analysis (Gerring et al. 2011). We might compare the state capacity of European nations that were narrowly subject to long-term foreign rule (such as Ireland, Finland, or Poland) to those that narrowly survived as independent powers (such as Belgium, Norway, or Portugal), though this might not provide strong evidence for the hypothesis.

Second, the critique of colonialism thesis “reads history sideways,” by comparing the current state capacity gap between former colonial powers and post-colonial states in order to conclude that the latter were at some point “deinstitutionalized.” Yet as the reversal thesis asserts, at the time of early colonization in the sixteenth and seventeenth centuries, many colonial powers were themselves relatively weak states. They augmented their capacity in the eighteenth and nineteenth centuries for reasons that were largely endogenous to Europe - interstate competition, socioeconomic modernization, and technological inno-
vation (e.g. the spread of printing, or improvements in transport and chronometry). It is not so much that non-European regimes were “deinstitutionalized” by Europe, but rather, that during the period of the colonial endeavor, Europe was engaged in its own endogenous path of political modernization in which non-European states were not equal participants. A parallel can be made to endogenous growth theory, which asserts that the economic development of Europe and the West since the early modern period came not through their relations to other countries, but by changes internal to western capitalism, notably, investment in human capital and better property rights (Romer 1994, North and Thomas 1973). A stronger version of this hypothesis might even assert that European interstate competition had positive spillover effects for the state capacity of non-European regions: by serving as competitors, by providing models of reform, technical knowledge, and resource-intensive weaponry (above all, firearms), European states inadvertently augmented the capacity of their Middle Eastern and East Asian rivals. From the nineteenth century onwards, European states were engines of state capacity, both domestically and in the wider world. Regimes in competition with European powers – Japan, the Ottoman Empire, China – were committed to defensive modernization, and in countries under European colonial tutelage, techniques of state administration and control – the census, conscription, the land cadastre, the civil service exam – became transferred to colonial subjects, in ways that were unlikely had these regions been left isolated from world affairs, as European powers sought to enhance as best as possible their control over territorial possessions. The effect of European innovation in the domain of public administration was comparable to administrative innovation in the private sector, i.e. eventual organizational isomorphism as models of best practice spread to subsidiaries and competitors (DiMaggio and Powell 1983).

Third, we cannot attribute the weakness of the post-colonial state to the colonial past alone, because this colonial state was not simply a product of European design but heavily endogenous to pre-existing precolonial institutions. European powers did not discover the world as a tabula rasa but as a series of small and large states, equally struggling with challenges of resource mobilization, military modernization, and the consolidation of their administrative reach (Lieberman 2010). Therefore, both the strategies of foreign rule and the resultant levels of state capacity were determined by pre-existing patterns of institutional development. In a simple version of this hypothesis, there is “raw” precolonial persis-
tence: patterns of post-colonial state capacity largely reflect the extent of precolonial state formation because colonial administrators had little effect on the societies they nominally governed. Stated in a more nuanced form, there is also the interaction between strategies of colonial rulers and precolonial legacies. Whether a territory fell European rule at all, depended on the capacity of local rulers to resist, as in many areas conquest entailed repeated warfare. And the policies and institutions of colonial rule depended on the organizational resources precolonial polities made available as options for colonial administrators. This was most obviously the case in the option whether to govern by indirect rule, which required the existence of a sufficiently centralized precolonial polity (Gerring et al. 2011), yet it also applied with respect to specific aspects of administrative or fiscal policy. To give one example, the division of India into individual-cultivator (ryotwari) and landlord-based (zamindari) systems of taxation was not simply a consequence of colonial fiat, but also, as intended, a codification and simplification of the precolonial land tenure systems in place (see Bayly 1983, Kulke and Rothermund 2010). However most econometric analyses treat the precolonial era as a “year zero,” taking the differences among post-colonial states as reflecting the colonial past.

Finally, the “treatment” of colonial rule was not constant over time and space, and colonial rule meant very different things in different places. The conquest of mesoamerica by the conquistadores in the 1500s was clearly a different process, in just about every relevant regard (the degree of bureaucratization, relationship to indigenous population, nature and goals of the governing actors) to the administration of, say, Hong Kong under the British empire. Trade with Europeans let to a significantly growth in the slave trade in the seventeenth and eighteenth centuries with clear institutional legacies (Acemoglu et al. 2001, Nunn 2007, 2013), yet, colonial administration were also instrumental in ending the slave trade in the nineteenth and twentieth centuries, including intra-African trade. The literature on legacies of colonial rule acknowledges this diversity and includes many competing arguments as regards the legacies of specific colonial powers: Mamdani (1996) and Young (1994) for example criticizes the colonial state as generating high levels of despotic capacity, that subsequently became the basis for post-colonial, authoritarian rule; while other authors criticize the colonial state precisely for its weakness, serving only as a ‘gatekeeper’ (Cooper 2002) state with no real authority over the interior. Various authors attribute posi-
tive legacies to British colonialism either as a result of spreading common law institutions, preserving indigenous capacity through indirect rule, or promoting policies of trade (La Porta et al. 1999), though other authors have argued that British colonialism largely left a negative institutional legacy (Iyer 2010) or that indirect rule created weak states (Acemoglu et al. 2013, Goldstein and Udry, 2008, Acemoglu, Reed and Robinson, 2013). Moreover, European powers were surely not the polities engaged in colonial acquisition at this time, and the effects of non-European imperialism are insufficiently studied and understood. Most independent powers, including Abyssinia, China, Japan, Persia, the Ottomans, and Burma (until its conquest) were similarly engaged in warfare and the attempt to acquire new territories during the eighteenth and nineteenth centuries, and many polities today are the former territories of these Asiatic powers, notably Taiwan and the two Koreas (from Japan), Mongolia (from China), Eritrea (from Ethiopia), Afghanistan and Azerbaijan (from Persia), and much of the Balkans, Caucasus and the Arab Middle East (from the Ottoman Empire). The legacies of non-western empires are insufficiently studied, yet for now seem at best ambiguous (Kohli 2004, Becker et al. 2013).

Reversal or Persistence of Fortune?

In recent years, an alternative has been what we might call the “persistence of fortune” hypothesis, according to which it is precisely those states that were most established in the early modern era, such as China, Japan, or Ottoman Turkey, were the most likely to resist European encroachment, undergo defensive modernization, and experience “catch-up” growth in the late twentieth century (Putterman 2013, Gerner 2012). Meanwhile the colonial territories in Africa, Oceania, and Central Asia that have the weakest states today, were also largely pre-state societies prior to European arrival. Historical studies such as Fukuyama (2011) and Mann (1986) argue that non-European powers such as China and the Ottoman Empire were the first to implement measures such as military conscription or merit-based examination, while Japan after the Meiji Restoration offered a path to defensive modernization subsequently taken up by other Asiatic powers (Moore 1966). Advocates of the “persistence” view point out that the countries experiencing the fastest rates of eco-
onomic growth in the era following decolonization were almost universally non-European states with long state histories, that such countries were most successful at state-led development, and that none of these countries ever became a European colony. The persistence of fortune argument thus contends that powers which entered the modern age strongest also left it the strongest. In this view, the first have finished first, while the last did not last.

The “persistence” argument appears at first a radical departure from the reversal of fortune view. To the extent that clear causal mechanisms are specified, there are at least three potential mechanisms. i) The persistence of “mental models” of governance. Post-colonial governing practices based on neopatrimonialism or sultanism are a legacy of pre-colonial political orders and the familistic norms that governed prestate or early-state society. Upon independence, postcolonial leaders reverted to these patterns because the brevity of formal colonial rule had failed to establish alternative ‘legal-rational’ modes of legitimation (Weber 1978, Bratton and van de Walle 1994). By contrast, where colonial regimes built upon existing models of bureaucratic authority, as in Vietnam, Taiwan or Korea, post-colonial regimes were able to maintain strong, insulated states with limited opportunities for patron-client relations. Similarly, where colonial regimes preserved longstanding traditional modes of legitimation, as in the kingdom of Morocco or Malay Kedah dynasty, these provided a basis for institutional stability in the post-colonial era. ii) Endogenous colonialism. A colonial “treatment effect” cannot be easily identified, because whether countries were colonized at all was endogenous to prevailing levels of state capacity; much of the variance in state capacity attributed to a colonial legacy or to a mode of colonial administration (e.g. direct versus indirect rule) may simply reflect differences in the ability of the precolonial state to resist foreign incursion (Hariri 2012, Gerring et al. 2011). In other words, the treatment effect of colonial rule may well have been negative, but, it was only administered in those territories that were already occupied by weak states; not in those areas subject to stronger and more centralized precolonial regimes. This is obviously the case for those states such as China or Japan that resisted colonial rule actively and by and large successfully, and more ambiguously so in those areas that were sufficiently resistant to merit a strategy of indirect of governance. iii) The preservation of abolition
of tribal institutions. Along with colonial regimes, precolonial states played an important role in preserving or abolishing local tribal chieftaincies, which if preserved led to negative consequences for post-colonial states due to widespread clientelism and the balkanisation of legitimate authority. This is the reason indirect rule in India is correlated with positive institutional performance (Iyer 2010): because India’s large, centralizing precolonial states such as Mysore or the Maratha Peshwa had worked to abolish village and tribal chief (“poligar”) institutions, laying the foundation for effective state-level government in contemporary Karnataka or Maharashtra, whereas the weaker post-Mughal states of the North (which eventually came under direct British rule) preserved the power of landholders (zamindars) and village (mahalwari) governance. Similarly, in Africa, positive development consequences are associated with indirect rule in the case of the larger and stronger pre-colonial states such as Buganda or the Ashanti Empire, but not where indirect rule meant devolution of government to a fragmented network of local chieftaincies as in Sierra Leone (Acemoglu et al. 2013).

Though much work has yet to be done, a number of recent econometric studies have begun to provide preliminary evidence for the persistence of fortune viewpoint. Bockstette, Chanda and Putterman (2002) for example develop an index of “state antiquity” which measures the stock of accumulated time over which a supra-tribal polity has existed within the present-day boundaries of each of 149 countries, and find that an early and durable history of political organization is a powerful predictor of modern rates of economic growth. In the African context, Gennaioli and Rainer (2007) find that countries with a greater proportion of centralized ethnic groups have more paved roads, a greater percentage of infants immunized for DPT, lower infant mortality, a higher adult literacy rate, and greater schooling attainment, while Michalopoulos and Papaionnou (2013) investigate the role of precolonial institutions in shaping comparative regional development within African countries, using nighttime satellite density maps to measure electricity output around precolonial ethnic boundaries.

**Beyond Reversal and Persistence – A Partial Synthesis**

Perhaps we can transcend the “persistence” versus “reversal” debate by arguing that
Figure 7.2: GDP per Capita Growth (1950-2006), and State Antiquity in 1500. Notes: All non-European, non-settler regime countries.
both sides are correct, subject to a threshold condition. This condition is whether a state had achieved sufficient institutional development prior to European arrival in order to successfully resist colonization, in which case the encounter with colonial powers gave rise to defensive modernization; or not, in which case a territory was conquered and subjected to institutional reversal. The path of defensive modernization was that taken by ancient polities such as China, Persia and Turkey. These countries responded to the European challenge by reforming their fiscal and military infrastructures, securing their territory, and mobilizing against foreign rule. However, states that fell just below this threshold suffered a reversal, as they were the first to be targeted by colonial powers, and especially so if they were conquered in the initial phase when colonialism was led by the commercial (and extractive) impulses of the Spanish conquistadores and the English, French, and Dutch East India companies. This was the path taken by the Aztecs, the Inca, or the Sultanates of Bengal or Malacca, as well as the few early African trading posts, such as Senegambia or the Kingdom of Kongo, that became a base for slave-trading into the African interior. Finally, non-European regions that were not suitable for European settlement, and in which much of the population remained in pre-state societies – such as (much of) sub-Saharan Africa, upland southeast Asia, and Central Asia – were the last to be fought over by imperial powers, largely in the late nineteenth century. Here again, there was institutional persistence: non-state societies were left by colonial powers with the nominal authority of a central postcolonial bureaucracy, yet, the institutions of tribal government still largely in place.

This pattern can be seen in Figure 7.3, which shows the century of conquest by European powers and the average level of state history in 1500. The persistence of fortune argument is correct in observing that the most established states on the eve of the modern era were those most likely to resist European colonialism, maintain indigenous political institutions, and experience catch-up growth in the twentieth century. However, in accordance with the “reversal of fortune” argument, it can be shown that the first territories to be targeted by European colonial powers were indeed established polities offering possibilities for extractive surplus, and that conquest of pre-state societies was left until late in the colonial era.

Why does this non-monotonic relationship exist? The decision of European powers to conquer or not conquer a territory was a function of two variables: the cost of conquest, de-
Figure 7.3: Century of Conquest by European Powers and Average Level of State History in 1500 (Non-European Colonies Only).
terminated by the degree of indigenous state capacity and the ability of its elites to withstand military attack; and the payoff to the imperial ruler of victory, determined by the ability to extract rents through such practices as mining and tax farming. The cost of conquest rose linearly with the degree of indigenous state-building, as a high degree of state organization, comparable to that of European powers, was required to achieve the most effective military defense through practices such as conscription, and a sufficient fiscal extraction to invest in expensive new gunpowder weaponry. However, taxation and corve labour are established even in very early states, and even younger polities had such institutions. These were the polities most attractive to European colonialism, as they offered possibilities for extraction without also possessing the military institutions required to ward off foreign invasion.

Figure 7.4 illustrates this relationship, regarding the date of conquest (or non-conquest) by European colonial powers. In the first or ‘mercantile’ phase of European colonialism that existed approximately from the fifteenth to the nineteenth centuries, colonial ventures were led by private enterprise. This was either the case formally, as with the Dutch, English, Portuguese, Swedish, Danish and French East India Companies, or informally, in the case of the Spanish conquistadores. Under this system, the relevant decision to engage in the colonial enterprise was a logic of cost and benefit; whether the expected return on investment of conquering a territory would exceed the expense of manpower and resources required to engage in such an acquisition. (Among the company directors of the English East India Company, the term “imperialism” was one of abuse: it referred to the policy of engaging in expensive territorial ventures without regard to whether this would be justified by its economic benefits). As such, European colonizers were selective in their targets. Smaller powers such as the Portuguese, Swedes and Danes would rather avoid military expeditions altogether, preferring to establish trading posts along the main routes for sugar, slaves and spices. Yet even the colonial companies of the larger powers preferred to adopt peaceful means where possible, eschewing explicit conquest in favor of establishing strategic trading forts and treaties with local rulers. To the extent that trading companies engaged in war, it was more often with one another; and not with the local potentates whose cooperation they required.

Where early colonizers did engage in territorial acquisition, priority was given to those holdings that allowed for opportunities for profit through mining, tax farming, or slaving.
Figure 7.4: Two Phases of Colonialism
This greatly restricted the range of potential ventures that might be reasonably entertained. The armies of the Ottoman, Mughal, and Chinese Empires were vastly greater in manpower and resources than the largest of Europe’s domestic armies, to say nothing of the resources that could be mobilized at great distance by a motley collection of privateers. Even the largest European powers during this phase of world history were minor players in the global system, economically, diplomatically, and militarily. European colonialists needed to be selective in their choice of targets, opting for acquisitions in areas that offered the weakest resistance relative to the potential benefits available. In practice, this meant engaging in wars against established polities in which institutions of slavery and vassalage had already been established by local elites, but which did not yet have sufficient knowledge of Eurasian military technology or bureaucratic organization in order to mount an effective resistance. These polities were found in the Americas and in South and Southeast Asia; in the Americas because of the absence of familiarity with cavalry and gunpowder weaponry, as well as the devastating impact of European disease; in southeast Asia due to a relative lack of political centralization (Diamond 1999). Thus the first colonial acquisitions were in the territories of the Inca and Aztec empires, and in the weaker sultanates of the South China Seas. Yet despite these acquisitions, European colonial companies inserted themselves into a world economy that was still dominated by Asiatic powers. Fifty per cent of all silver mined from the Americas by the Spanish made its way to China, and to the extent that the Spanish acquired posts in Formosa and the Philippines, it was to be closer to their main client (Tignor et al. 2002). Similarly, much of the trade along the ‘spice route’ operated by the Dutch and Portuguese from Indonesia to Europe via Malabar and the Arab Gulf occurred internally, and only a small amount in fact originated or terminated in Lisbon or Amsterdam. Even the African slave trade remained dominated by Asiatic and internal African demand until the eighteenth century (Unesco 1979). European companies operated in a world of Asiatic powers, and required their cooperation: a Dutch colony on Taiwan was destroyed in 1662 when 50,000 Chinese soldiers turned up under the generalship of Koxinga, and forts operated by the Danish, French and Dutch East India Companies were seized by indigenous rulers when they displeased their local hosts².

²The Dutch lost their entire Malabar possessions to Travancore after losing the Tranancore war in 1741.
During this “mercantile” phase of European colonialism, the established Eurasian poli-
ties such as Persia, China, Vietnam were left undisturbed; likewise untouched were those
areas still populated by pre-state societies, whether in the Amazonian interior, upland south-
est Asia, or much of Africa. The latter were difficult to subjugate and for the sake of little
obvious benefit; no sooner had one tribe been pacified, than another might erupt in rebel-
lion. Thus the simplest policy in dealing with pre-state societies was a logic of displacement
and extermination, as carried out to remarkable effect in the ‘neo-Europes’ of the Americas
and in Australasia (Reynolds 2001). Yet even in these cases of European settlement, the
land had to be comparatively sparsely populated by indigenous peoples; the Americas had
been decimated by the impact of European disease, and the African Cape not yet reached
by the great Bantu migration.

A second phase of European expansion abroad began from the mid-nineteenth century,
which we may refer to as the “imperial” phase. During this era, colonial ventures were
no longer being led by private actors, but rather by governments that sought to maximize
territorial control in their competition with other states. The English East India Company
was brought under parliamentary authority in 1773 and finally nationalized in 1858, the
French East India Company in 1769, and the Dutch in 1798, and decisions over statecraft
transferred to national administrators. As a result, military expeditions were frequently
conducted despite the lack of obvious economic benefits, and in areas of limited economic
value, but of geostrategic importance. This is illustrated by the shift in utility curve shown
in the second chart of Figure 7.4. Overseas colonies became an extension of national state-
craft, and were administered not with the goal of profit but with the aim of enhancing the
state’s competitive position against other states. In the nineteenth century, European pow-
ers began waging wars against strong polities that they had previously left undisturbed; the
British fought a new round of wars in South Asia, primarily against Burma, Afghanistan,
and Nepal; and the French fought new wars in southeast Asia, claiming the ancient states
of Vietnam and Cambodia. These were difficult and costly conflicts, that a fully commer-
cial logic would have avoided. It took three Anglo-Burmese wars to subjugate Myanmar,

The Danish East India Company’s fort at Tranquebar was repeatedly raided by Nayaks, and much of the
French territories in Tamil Nadu were retaken by indigenous rulers in the late eighteenth century, albeit
working in alliance with the British East India Company.
the first of which being the longest and most expensive war in British Indian history, while Afghanistan was never successfully conquered. It took France over thirty years to gain territorial control over Vietnam, from the first expedition by Rigault de Genouilly in 1858 to the conclusion of the Tonkin campaign in 1886. On the other hand, not only ‘strong’ states became targets of imperialism, but also the vast remaining areas inhabited by pre-state societies that had previously remained untouched. For the first time European powers began to acquire territory in Africa for administrative ends, rather than simply engage trading settlements on the coasts for the end of securing slaves; in this way Germany declared its possession of Togoland, the Cameroons and South West Africa in 1884, and King Leopold of Belgium claimed Congo as his personal possession the following year. The conquest of pre-state societies also came at high military cost, due to repeated rebellions such as the Dervish, Herero, and Namaqua uprisings. A similar logic of territorial consolidation dictated that Russia, having completed its administrative hold over the tundras of Siberia, begin subjugating the nomadic peoples of Central Asia by acquiring the territories that were to become Kazakhstan, Kyrgyzstan, and Uzbekistan; both Russia and Britain eyed Afghanistan. The “great game” in Central Asia and the “scramble for Africa” were the defining events of the last phase of imperialism: a competition among colonial powers for territorial control with little regard to economic benefit and often at substantial net economic cost.

The encounter with European colonialism therefore created four kinds of state in the modern world: persistent states, arrested states, new states formed out of pre-state societies; and settler states formed from repopulation by settlers.

i) “Persistent” states are polities which survived the era of European colonialism as independent powers, and indeed were more likely to have been strengthened by their interaction with Europeans as a result of engaging in defensive modernization. Persistent states were already longstanding powers at the start of the modern era (circa. 1500). Examples of persistent states with long state histories were China, Japan, the Ottomans (Turkey), Morocco, Persia (Iran), and Siam (Thailand).

ii) “Arrested” states were already centralized polities at the point of contact with Eu-
Table 7.6: Persistent, Arrested, and New States outside of Europe.

<table>
<thead>
<tr>
<th>Persistent States</th>
<th>Arrested States</th>
<th>New States</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>India</td>
<td>Papua New Guinea</td>
</tr>
<tr>
<td>Japan</td>
<td>Pakistan</td>
<td>Kyrgyzstan</td>
</tr>
<tr>
<td>Persia (Iran)</td>
<td>Bangladesh</td>
<td>Central African Republic</td>
</tr>
<tr>
<td>Ottomans (Turkey)</td>
<td>Indonesia</td>
<td>Congo (DR)</td>
</tr>
<tr>
<td>Siam (Thailand)</td>
<td>Malaysia</td>
<td>Cameroon</td>
</tr>
<tr>
<td>Morocco</td>
<td>Mexico (Aztecs)</td>
<td>Senegal</td>
</tr>
<tr>
<td>Bhutan</td>
<td>Peru (Inca)</td>
<td>Tajikistan</td>
</tr>
<tr>
<td>Nepal</td>
<td>Bolivia (Inca)</td>
<td>Tanzania</td>
</tr>
</tbody>
</table>

European colonialism, but were either insufficiently technologically advanced or bureaucratically organized to resist foreign rule. Arrested states had high state history scores at the outset of the modern era, but not as high as the persistent states, and this is likely to have been covariate with weaker bureaucratic penetration, fiscal capacity, and military readiness. Arrested states were the first to be targeted European colonists in the sixteenth and seventeenth centuries. Such reversal states include India and Pakistan (Mughals), Indonesia and Malaysia (Malacca sultanates), Mexico (the Aztecs), Peru and Bolivia (the Incas), Cambodia (the Khmers), and Myanmar (Burma).

iii) “New” states are found where colonial powers encountered pre-state societies, there was limited indigenous state capacity on which to build, and by consequence political institutions were developed entirely under the aegis of colonial rulers.

These three states are summarized in Table 7.6.

i) Persistent States – the Case of Siam

An example of a “persistent” state would be Thailand, governed by a succession of kingdoms from the 13th to the 18th centuries that are referred to as the kingdom of Siam. While Siam began as a relatively weak and patrimonial state, repeated warfare between
Siam and its neighbours – above all Burma, against which the kingdom fought no less than twenty wars in the years between the mid-sixteenth to the mid-nineteenth centuries – forced a greater degree of centralization. By the nineteenth century, France had succeeded Burma as Siam’s main challenger on the peninsula due to its expanding empire in Indochina, and the state engaged in extensive defensive modernisation in order to prepare militarily for this threat. Under king Rama V (r. 1853 – 1910) in 1887 a Royal Military Academy was founded along western lines, in 1888 the executive branch was restructured into ministerial government (including a new ministry of war), and in 1898 a Royal Naval Academy was founded, and in 1905 a law of universal conscription introduced. The Siamese kingdom also sought to gain better legibility over its population and rationalise its method of taxation; to this end the Thai Survey School was established in 1882, and a first land cadastre conducted several years later, in 1896. In 1932, a “breakthrough” coup by the army established a constitutional monarchy, and during the Second World War Thailand declared war on France in order to regain control over Laos.

Siam’s survival as an independent state was contingent, and it is unlikely the Siamese state would have survived a lengthy anticolonial war as endured by Burma (at war with the British Empire on three occasions from 1824 to 1886) and Vietnam (at war with France, from 1858 to 1886). Having persisted through to the late nineteenth century, however, the kingdom was able to rationalise its administration and modernise its army in order to effectively ward off attempts at colonial rule. In common with most other persistent states – notably Japan, Nepal, Bhutan, Morocco – Thailand has survived as a constitutional monarchy, and persistent states are well represented among the surviving monarchies of the world today. This is largely due to the greater legitimacy accumulated by historical sovereigns; “artificial” monarchs imposed by colonial powers, such as the post-colonial kings of Egypt, Iraq, or Libya, had much shorter duration in office. Persistent states are also responsible for the observation that early states carry an ‘autocratic legacy’ (Hariri 2012); namely as the stock of legitimacy and compliance built up by absolute monarchs allowed for the construction of highly centralized non-competitive political orders in the contemporary era.

**ii) Arrested States – the Case of Burma**
An example of an “arrested” state would be Burma, or contemporary Myanmar. During the early modern era, Burma was one of the paramount powers of Indochina, fighting repeated wars against neighbouring Siam and repulsing four invasions by Qing Dynasty China in late eighteenth century. In the nineteenth century, the polity’s strongest regional rival became the British Empire, due its consolidating control over neighbouring India. After losing a first war against the East India Company in 1826, Burmese elites made an attempt to modernise their polity, including introducing a salary system for the bureaucracy, fixed judicial fees, penal laws, increases in direct taxation and comprehensive military reform. However, this was not sufficient to stave off successive British attempts at wresting control over its territory, and in 1886 Burma was incorporated into the British Indian Empire.

Under British colonial rule, Burma’s indigenous monarchical institutions were not preserved. The king and his family were instead sent into exile in India, and a colonial administration was formed in the new capital of Rangoon, in place of the palace capital at Mandalay. As in most ‘arrested’ states, the colonial era saw the construction of a parallel administrative apparatus that marginalised the existing, precolonial structures, and in many cases marginalised its former administrators and elites. This was the case in Burma, where Anglo-Burmese dominated the civil service. The result was a divided society, a strong anticolonial movement, and a post-independence politics marked by conflict between radical factions demanding revolutionary change and military officers seeking to restore political order. With the existing sources of political order displaced, the post-colonial history of many arrested states has often been unstable. In the case of Myanmar, a military coup in 1962 established a junta that subsequently ruled the country, punctuated by widespread through unsuccessful civic uprisings in 1974, 1988 and 2007.

iii) New States – the Case of Papua New Guinea

Perhaps the clearest example of a “new” state, lacking any precolonial political centralization, is Papua New Guinea. Formed under colonial administration by the merger of Papua and New Guinea in 1945, the country is one of the mostly heterogeneous in the world, with hundreds of ethnic groups and over eight hundred languages, and the country’s entire process of state formation has occurred under colonial tutelage until the country’s late independence in 1975. Everything that defines Papuan statehood, from the capital
(founded in 1873 and named after a British captain), to the language (a form of patois English), to the legal system (imported from English common law), was constructed under the aegis of its colonial rulers, and in this respect it is an ideal-typical ‘new’ colonial state. While Papua may represent an extreme case, however, it typifies the dynamics of many post-colonial states in Sub-Saharan Africa, Oceania, and Central Asia, in which French, British, and Russian administrators attempted to craft a metropolitan state in the image of the colonial power, grafting atop a tribal, precolonial society. Due to the absence of state penetration into local communities, the postcolonial state is typically weak and detached, with limited territorial control outside of the capital, and clientelism towards competing tribal groups and affiliations.

I argue, therefore, that European colonialism and indigenous institutions interacted to create three kinds of state in the modern world: “persistent,” “arrested,” and “new” colonial states. Which of these outcomes arose was due to the level of precolonial state formation that had occurred prior to colonial influence. States at the highest level of political development at the point of contact with European powers, responded to this challenge through defensive modernization: these states ‘persisted’, or remained independent, and continue to exhibit the highest levels of state capacity today, as measured by such conventional metrics as fiscal depth, the rule of law, or the quality of the bureaucracy. On the other hand, a second group of polities were sufficiently developed by the early modern era to offer opportunities for labour-extractive practices such taxation and mining, but not sufficiently developed in their fiscal or military infrastructures to resist foreign rule. These were the first to be targeted by European colonialists in the sixteenth and seventeenth centuries, resulting in an interruption of their indigenous institutional development. Finally, a third group of polities were created entirely under colonial rule, in areas in which there was little or no indigenous state capacity beforehand. These territories were targeted only in the final phase of European imperialism, and combine a post-colonial central state with almost complete lack of state capacity outside of the capital territory.
Conclusion: The Cross-Country Relationship Between Early States and Modern State Capacity

The general conclusion of this analysis is that contingent historical factors play the greatest, and indeed perhaps the only major role in historical state formation. This becomes more intuitive when we consider actual cases: Why did the Siamese state survive the age of European colonialism unscathed in order to emerge as modern Thailand, while the Burmese, Khmer, and Vietnamese polities were colonized? Why did mountainous Bhutan escape British colonialism in South Asia, while mountainous Kashmir and Sikkim did not? Why did the French patrimonial state survive repeated invasion and defeat in order to re-emerge as a centralizing bureaucratic power, while Poland was dissolved only to re-emerge in the age of nationalism? While ad hoc theorizing is always possible, the simplest answer in these cases is chance. There are plausible alternative histories in which Poland might have become a great power in place of France, in which Siam would have fallen to the French while Vietnam could have been independent, and in which Bhutan would have joined the Raj and Kashmir, like Nepal, could have retained full independence.

The results of the regressions in Tables 7.2 to 7.5 are broadly positive for a cluster of theories that we might characterize as ‘historical institutionalism’ - that see advances in state capacity, not as the result of environmental or geographic determinism, but rather as an accumulation of contingent historical circumstances - such as the policies of colonial rulers, state survival, a legacy of warfaring, or settlement by peoples with a historical experience of state-building. An historical institutionalist approach does not argue the improbable view that historical factors are independent of geography, economics, or other explanatory variables, but simply that there is no one single covering ‘meta-explanation’ possible among them. As a result, the greatest utility level of explanatory reduction that we are likely to achieve remains at the level of historical factors themselves: leaving us to debate such factors as colonial legacies, state antiquity, and settlement patterns, rather than, say, relative factor endowments or the cost of extending rule over a particular type of territory.
Chapter 8

State History and State Capacity in Russian Regions

The previous chapter has shown that the association between precolonial political development and contemporary political order found within India, can also be found among countries in the world. But does the subnational pattern of variation found in South Asia also exist in other large territorial states with fragmentary legacies of early state development? Ideally, we should seek a confirmatory case in a country that is ‘most different’ from the case of post-colonial South Asia.

In this chapter I show that the breakup of the Soviet Union also provides a remarkable experiment in testing for the legacy of historical states in raising local state capacity and the means for local public goods provision. With the end of centralized rule from Moscow, regional governments were suddenly faced with responsibility for wide-ranging duties of fiscal administration, policing, and revenue mobilization. This chapter shows that the strongest correlate of success in handling this transition was a region’s historical legacy of state formation, prior to the reform era. Using a new dataset of state history for 83 provinces of the Russian Federation for ten centuries until 1991, cumulative state formation is shown to be the decisive factor explaining success or failure at provision of public goods and the management of public administration. By examining case studies from within the Russian Federation, it is argued that a combination of the legitimacy of local elites, subnationalist mobilization, and regional bargaining strategies can explain the outperformance
of areas with longer state history vis-à-vis those which did not build on existing traditions of governance.

As yet, while there are a growing range of cross-country studies of the effects of historical state formation on contemporary quality of governance, there have yet to be systematic econometric studies showing these effects at the sub-national level. Further, while cross-sectional studies are able to correlate certain outcomes (economic growth, public goods provision) to measures of historical state formation, such studies do not satisfy the chronological condition of experimental method, that a variable be manipulated prior to the onset of the observation. Accordingly, this chapter remedies these deficiencies by examining state capacity in the period following the break-up of the Soviet Union, after a substantial transfer of sovereignty to its constituent units.

**State History and State Capacity in the Former Soviet Union**

The Soviet Union operated upon a highly centralized system of center-periphery relations. Its collapse was concomitant with a centrifugal cascade of power that saw a massive devolution in responsibilities handed to subnational entities inside the new Russian Federation, as the central government was no longer able to provide clear lines of authority, career paths for regional elites, or generous subsidies and transfers (Holmes and Colton 2006). By the early 1990s Yeltsin was inaugurating a sweeping and chaotic decentralization program that gave political autonomy first to the two main cities (Leningrad and Moscow) and six republics. A five year moratorium on elections for regional leaders appointed by Yeltsin was proclaimed at the end of 1991, but during this period 31 regional elections took place with his consent, and a further two without (Zhuravskaya 2010). By mid-1992, 4 of the 5 autonomous oblasts became republics, and Chechnya and Ingushetia were separated, bringing the number of republics to 21. A Federation Treaty was signed on 31 March, specifying types of power allocated to four types of unit – the federal government; the 21 republics; the 10 autonomous okrugs and 1 autonomous oblast; and the 49 oblasts, 6 krais and 2 federal cities (Herrara 2005).

Political decentralization was accompanied by an uneven fiscal decentralization (Figure 8.1). By the late 1990s, 26.5 per cent of Russian budgetary expenditures were being
disbursed at the regional level, and a further 18.5 per cent at the local level: proportions that are similar to federal regimes such as Germany or the United States (IMF 1999). Analysis of the sectoral composition of regional and local expenditure shows that in some areas, budgetary management was almost fully decentralized: by 2001, regional and local government accounted for over three-quarters of spending in the domains of public transportation, culture, healthcare, roads, and education (Andreeva and Golovanova, 2003).

In addition to decentralization of public expenditure, Russia also underwent a wide-ranging decentralization of revenue mobilization. By 1999, 24 per cent of revenues in the general government budget were raised at the regional level, and 12.3 per cent at the local level; for basis of comparison, these are similar to the levels in the United States (25.9 and 15.7 per cent, respectively) and Germany (21.3 and 12.6 per cent). To a significant degree, Russia’s provinces were left to fend for themselves: according to calculations by Andreeva and Golovanova (2003), own-source revenues accounted for as much as 45 per cent of the spending of Russian federal subjects by the end of 1990s. Russia therefore underwent a thorough federalization, both of its spending, and its revenue raising functions during this era.

Finally, as well as decentralization of revenue mobilization and spending, Russian re-
gions also gained the authority to issue debt on the open market, as a means of managing revenue uncertainty from one year to the next. After one decade of decentralization, this left widespread regional discrepancies in public debt management. In small Ust-Orda Buryat oblast, over 10 per cent of budget expenditures were financed by government securities issues, and in Tomsk Oblast over 15 per cent (Andreeva and Golovanova, 2003). By January 2001, six provinces had outstanding liabilities amounting to over 10 per cent of their budgetary expenditures.¹ This, of course, is similar to other federal regimes where states are able responsible for managing their own debt and interest burdens.

As well as participating in this wave of fiscal decentralization, Russia’s 21 constituent republics, each nominally autonomous and home to a specific ethnic minority, went much further. Following Yeltsin’s call for the constituent republics to ‘take as much sovereignty as you can swallow,’ fourteen of the then sixteen Russian ASSRs (Autonomous Soviet Socialist Republics) declared themselves sovereign, subsequently establishing their own constitutions, legislatures, and even foreign ministries (Dunlop 1993). This was formalized by the federation treaty of 1992 which gave republics the right to their own foreign and trade policies (Treisman 2000). Over the course of the 1990s, a number of these constituent republics continually tested the limits of Moscow’s willingness to devolve power, by appropriating as much fiscal, administrative, and legislative autonomy as they could for themselves.

Empirical Strategy

How should we go about testing subnational variation in state ‘capacity’? Recent studies of state capacity have tended to break the term into the state’s various component ‘capacities’, for example, extractive capacity (the raising of revenues), coercive capacity (the ability to defend one’s borders) and administrative capacity (the enforcement of contracts, protection of rights, and the delivery of public services; Hanson and Sigman, 2011). However, most studies concern administrative capacity as the central aspect of state performance, the empirical measurement of which is made easier by the existence of ‘actionable’

¹10.7 per cent in Kursk Oblast, 11.2 per cent in Omsk Oblast, 15.0 per cent in Ulianovsk Oblast, 16.7 per cent in Tambov Oblast, 18.6 per cent in Altai Krai and 18.8 per cent in the Republic of Mari El.
or hard indicators of service delivery - without need to rely on expert-assessment measures that may be subject to halo effects or response biases (Knack 2007). While critics of the use of public goods indicators have alleged that the extent of public goods delivery may reflect policy choice rather than administrative capacity, this is less of an issue in a sub-national context where all provinces are in principle obliged by the central government to deliver certain goods to their citizens.

Accordingly, a set of regressions are estimated using as the dependent variable a range of 'hard' indicators designed to test the extent of public service delivery and rule of law. Included are a series of measures of public goods delivery from across Russian provinces, which include the proportion of slum dwelling and failing housing stock among all housing stock in the region; the provision of public buses (per 10,000 inhabitants); the proportion of roads that are paved; the proportion of paved roads that have improved road surfacing; the number of hospital beds per capita; the number of inhabitants per doctor, the number of museum visits per 1,000, and the number of children per educational place. In addition, as a measure of the rule of law, we include the homicide rate, which reflects upon the efficacy of police services, the prevalence of organized crime, and the efficacy of judicial means of
resolving disputes.

These models are estimated by a series of OLS regressions of the form:

\[ Y_{nrt} = \alpha + Y_{\Sigma nrt-1} + \beta_1 S_r + \beta_2 X'_{rt} + \epsilon_{rt} \]  

(8.1)

Where \( Y_{nrt} \) refers to one of \( n \) metrics of public goods delivery in region \( r \) at the present time \( t \), \( \alpha \) to the intercept, \( Y_{\Sigma nrt-1} \) is an index of public goods delivery in the earlier time period 1994-5, \( S_r \) refers to that region’s accumulated state history in the centuries preceding the transition era, \( X'_{rt} \) refers to a matrix of control vectors for covariant social and economic attributes, such as local social capital or regional income per capita, whose independent association with \( Y \) is estimated in the vector of betas \( \beta_2 \).

As our independent variable of interest, this chapter measures state history following the method outlined in Bockstette, Chanda and Putterman (2002) for each of the 83 subjects of the Russian Federation. The index is constructed by taking each fifty-year period from 1000 to 1950 AD, and allocating points to countries if there was i) a government above the tribal level; ii) if that government was locally based rather than that of a foreign empire; and iii) a fractional point to represent the extent of the country’s modern territory that was under the control of this earlier government. Data was compiled by Russian research assistants using Russian language historical sources and following the coding rubric, without alteration by the author. To arrive at an index score for each region, the data from the fifty periods is combined. In order to represent the fact that a more recent legacy of state formation is likely to have a greater impact than a relatively more distant one, a discount rate is applied in the aggregation of the scores, as in Bockstette, Chanda and Putterman (2002), thereby assigning relatively greater weight to more recent episodes of state history. Areas

\[ 2 \] In the original work by Bockstette, Chanda and Putterman (2002) state history indices are calculated back to 1 AD, but the discount rate is such that legacies of state formation before 1000 AD have minimal effect upon the index scores. Therefore, the subnational state history indices take only the period from 1000 AD onwards.

284
with a longer history of government of any kind receive a greater score, and a yet higher score still if this government was domestic, and was in charge of much of the territory of the contemporary provincial boundaries. For example, the region of Pskov was a self-governing republic from the time between 1000 and 1399, receiving the full score of 50 for each period. After 1399, it was transformed into a viceroyalty of Muscovy, receiving a fractional score of 37.5 to reflect indirect foreign rule: until 1510, when its independence was ended and it became a full vassal of Moscow, receiving 25 points per period after this point, to reflect direct foreign rule. This continues until Muscovy becomes the Russian Empire, after which Pskov is treated as an indigenously ruled region of Russia, and again accumulates state history at 50 per period, resulting in a final score of 0.84 (on a 0-1 scale) in the present day. By contrast, a region such as the Chukotka, in the northeast of Siberia, has a more recent history of state formation: traditionally home to the nomadic Chukchi people, no organized state existed in the area until the arrival of Cossacks in the seventeenth century, who fought the Chukchi in a series of battles from 1701-47. These prompted a higher degree of political integration among the Chukchi, who in their peace treaty of 1778 secured a tax exemption in exchange for indirect rule of a portion of the oblast’s area under suzerainty of the Russian Tsar, resulting in a partial state history score of 11.25. Though the Chukchi remained the predominant inhabitants of the area, Russians, Americans, and Canadians made competing claims to the area until 1923, when Russia expelled all remaining foreign settlers; though as the paramount administrative rulers of the territory, from the late nineteenth century the area is treated as Russian imperial land, with a state history score of 50 per period, resulting in a final state history score of 0.41 (on a 0-1 scale) in the present.

In order to provide something approaching an experimental test design, a control variable is included for the quality of public goods delivery in 1994-5: the point of onset for fiscal and administrative federalism in Russia. Only a reduced subset of public goods variables are available for this earlier time period, but do include the number of doctors per 10,000 inhabitants, the volume of surfaced roads, the proportion of hospital beds per

---

3However ethnic Russians remain a minority in the province until the 1950s; as late as the 1939 census, Chukchis and Chuvans continued to account for 56.2 per cent of the region’s population.
Figure 8.3: State History Across Russia. Index constructed by taking each period from 1 to 1950 AD, and allocating points to countries if there was i) a government above the tribal level; ii) if that government was locally based rather than that of a foreign empire; and iii) a fractional point to represent the extent of the country’s modern territory that was under the control of this earlier government.

1,000 inhabitants, and the number of annual bus passengers (Rosstat 2012). The volume of surfaced road and public bus passengers were normalized to population in 1994, and standardized values for each of these four indicators aggregated into an index of public goods provision by province in 1994-5. Observation of the descriptive statistics show that in the mid-1990s, the pattern of public goods provision across Russia continued to reflect the strategic priorities of the Soviet state, with the best performing regions being the capital, Moscow, plus the Far Eastern provinces to which official Soviet policy was to encourage the flow of migration (Zaslavskaya and Kalmyk, 1981). By contrast, the worst performing regions at this time were the North Caucasus, as they remained two decades later, as well as the Central and Volga regions.

As well as including a measure for state history as an independent variable and a control for public goods provision at the start of the federal reform era, several other indicators are also included. First, where funding for public goods provision is calculated based on locally raised revenues, the provision of public goods is likely to be positive affected by a higher
level of regional GDP. Accordingly, the models comprise a variable for regional income per capita in 1994.

In addition, a significant body of recent literature has alleged that ethnic fractionalization is negative for the provision of public goods (Alesina et al. 2003, Alesina, Baqir, & Easterly, 1999; Miguel & Gugerty, 2005). As the Russian Federation contains a wide diversity of ethnic groups, it should provide a good test case for the ethnic heterogeneity and public goods hypothesis. Accordingly, a variable for ethnic fractionalization, by subregion, is calculated based on Russian census data, and utilized.

Finally, a longstanding literature has argued that social institutions matter for local governance, in particular participation of citizens in local governance through civic activities such as protest, petition, attending neighbourhood associations, writing to local newspapers and officials, can deliver better quality of governance (Putnam et al. 1993, Norris 1999, Tsai 2007). A variable is therefore included which registers the proportion of citizens, by subregion, who have been involved in one or more of the following activities: actions of political protest; peaceful demonstrations; organizing a group of people to solve their own or someone else’s problem; cleaning up public space; participating in meetings of people living in the same house; communicating in writing with state institutions about personal issues; communicating in writing with state institutions about collective issues; participating in public hearings; organizing to help to people who are in the difficult situation; publicly expressing their opinions on the internet; expressing their opinion at meetings and state institutions; and participating in the activities of NGOs (Georatings, 2008).

Results are shown in Table 8.1. As a check against multicollinearity, a pairwise correlation matrix of the independent variables is included in the appendix; in all cases the correlation is extremely low (-0.16 < r < 0.28).

The estimated coefficients show that state history is very strongly associated with the rule of law and with quality of urban infrastructure: areas with a longer legacy of state formation have significantly lower homicide rates, a higher proportion of improved surface roads, lower proportions of slum housing, and greater use of museums. State history is more weakly associated with welfare provision, such as greater access to health, education or public transportation, which vary instead with income per capita. This is consistent with the observation that spending in these areas is disproportionately financed by local-raised
<table>
<thead>
<tr>
<th></th>
<th>Homicide rate</th>
<th>Paved roads (%)</th>
<th>Surfaced roads (%)</th>
<th>% Slum Housing</th>
<th>Persons per doctor</th>
<th>Hospital beds / 000</th>
<th>Children / place</th>
<th>buses per 10,000</th>
<th>Museum visits (log)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public goods provision, 1994-5</td>
<td>-0.282</td>
<td>-2.735</td>
<td>-3.342</td>
<td>-1.48</td>
<td>-47.045***</td>
<td>16.646***</td>
<td>-4.559*</td>
<td>9.35</td>
<td>-0.342*</td>
</tr>
<tr>
<td></td>
<td>(1.638)</td>
<td>(4.069)</td>
<td>(5.419)</td>
<td>(1.013)</td>
<td>(10.655)</td>
<td>(3.916)</td>
<td>(2.042)</td>
<td>(5.6)</td>
<td>(0.163)</td>
</tr>
<tr>
<td></td>
<td>(3.802)</td>
<td>(9.5)</td>
<td>(12.652)</td>
<td>(2.351)</td>
<td>(24.733)</td>
<td>(9.089)</td>
<td>(4.74)</td>
<td>(13.346)</td>
<td>(0.378)</td>
</tr>
<tr>
<td>Log GDP per capita, 1995</td>
<td>0.537</td>
<td>-2.185</td>
<td>-6.933</td>
<td>-0.731</td>
<td>-30.137**</td>
<td>7.516*</td>
<td>-0.894</td>
<td>17.912**</td>
<td>0.373*</td>
</tr>
<tr>
<td></td>
<td>(1.561)</td>
<td>(3.87)</td>
<td>(5.153)</td>
<td>(0.966)</td>
<td>(10.158)</td>
<td>(3.733)</td>
<td>(1.947)</td>
<td>(5.352)</td>
<td>(0.155)</td>
</tr>
<tr>
<td>Ethnic Fractionalization</td>
<td>-0.2</td>
<td>-10.274</td>
<td>-17.396</td>
<td>2.274</td>
<td>-19.02</td>
<td>-6.216</td>
<td>4.253</td>
<td>5.985</td>
<td>-1.411***</td>
</tr>
<tr>
<td>Civic Activity</td>
<td>0.461</td>
<td>0.108</td>
<td>32.334</td>
<td>-4.123</td>
<td>-51.015</td>
<td>3.668</td>
<td>-6.444</td>
<td>-13.532</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>(6.405)</td>
<td>(15.865)</td>
<td>(21.129)</td>
<td>(3.962)</td>
<td>(41.669)</td>
<td>(15.313)</td>
<td>(7.985)</td>
<td>(21.918)</td>
<td>(0.638)</td>
</tr>
<tr>
<td>Moscow dummy (0/1)</td>
<td>-4.397</td>
<td>(-)</td>
<td>(-)</td>
<td>-0.09</td>
<td>-22.943</td>
<td>-26.268</td>
<td>7.087</td>
<td>-4.913</td>
<td>0.693</td>
</tr>
<tr>
<td>Constant</td>
<td>18.177</td>
<td>101.157**</td>
<td>113.204*</td>
<td>13.632</td>
<td>444.651***</td>
<td>64.912</td>
<td>112.916***</td>
<td>-144.293**</td>
<td>2.161</td>
</tr>
</tbody>
</table>

Table 8.1: Determinants of Public Goods Provision across Russian Federal Subjects. * significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level.
revenues, rather than federally provided grants: thus richer regions have more resources available to spend in these domains (Andreeva and Golovanova 2003).

By contrast, ethnic fractionalization is not associated with any of the measure of public goods provision or usage, except for the number of museum visits, with which it is significantly negatively related. The number of museum visits also varies significantly with both state history and income per capita. Though the sign for ethnic fractionalization is in the expected direction for paved roads, surfaced roads, slum housing, hospital beds per 10,000, children per educational place, and public buses per 10,000, none of these effects are significant or robust to the inclusion of controls for state history and income per capita.

What do these differences mean in practice? The mean score on the state history index of 0.71 with a standard deviation of 0.19 entails that a one-standard deviation increase in state history is associated with a 2.71 per 100,000 reduction in the homicide rate; a difference that is similar to the gap between continental Europe and the United States. Meanwhile a one standard deviation increase in state history is also associated with a 1.49 per cent reduction in the proportion of slum housing (the Russian provincial average is 4.49 per cent) and a 10.1 per cent increase in the proportion of paved roads with improved road surfacing, where the Russian provincial average is 66 per cent. These differences, therefore, are not only of statistical significance, but would be visible to the naive observer. All of the provinces in the top quartile of state history have less than 10 per cent slum housing, over half of roads with quality surfacing, and a homicide rate below 15 per 100,000: relative to other provinces, in these areas the streets are noticeably safer, surfaced, and lined by apartment buildings, rather than makeshift wooden structures. By contrast, in many of the areas with low state history – including a number, which due to resource rents, are now comparatively wealthy – criminality, absence of urban infrastructure, and the presence of shanty housing are apparent.

Decomposing the State History Index

As the state history index is composed of three elements, namely the presence of state, the degree to which that state was indigenous rather than foreign, and the amount of the contemporary territory governed by that state, a natural question that arises concerns which of these three aspects of state formation is most significant for explaining variation in the
current distribution of public goods. Table 8.2 therefore shows the results of a series of regressions in which, in place of the combined state history index used in Table 8.1, each of three separate subindices has been used. These subindices are aggregated for each respective measure (stateness, locality, and contiguity) and rescaled 0-1, whereby 0 represents the minimum possible score and 1, the maximum. As in the regressions reported in Table 8.1, controls are included for the public goods index in 1994-5, the level of ethnic fractionalization, GDP per capita in 1995, a Moscow dummy variable and the social capital index, though, these coefficients are not reported.

The estimated coefficients show that each of the state history subindices, independently used, explains a similar variation in the public goods distribution; and that no single subindex is responsible for the results. The estimated coefficients are in general somewhat larger for the locality index, i.e. the accumulated time over which a region has been governed within a polity indigenous to that area or its titular majority, yet these estimated effects only surpass a higher significance threshold in a few cases. As both locality and contiguity imply the existence of a state the degree of collinearity between the three subindices is high (0.86 < r < 0.93), though the implication is that it is the extent of historical government, and not the form of this government, which explains the accumulation of local governing capacity.

Causal Mechanisms

This strong and robust link between early state formation and modern state capacity raises a series of questions. First, what links historical polities to the functioning of contemporary regions, if social institutions and ethnic homogenization have already been ruled out? By what intermediary chain of events do extant polities pattern the contemporary distribution of state capacity? This raises a recurrent dilemma of historical explanations in the social sciences, where both dependent and the independent variables are temporally distantiated: that of identifying a causal pathway. Stinchcombe (1968), reflecting on this issue, has usefully distinguished ‘constant’ from ‘historical’ causes: whereas a constant cause operates year after year, a historical cause reproduces itself without the occurrence of the original cause, and a satisfactory explanation must in one or both ways account for
<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Stateness</th>
<th>Locality</th>
<th>Contiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homicide Rate per 100,000</td>
<td>-9.904*</td>
<td>-14.035***</td>
<td>-11.814**</td>
</tr>
<tr>
<td></td>
<td>(3.991)</td>
<td>(4.053)</td>
<td>(3.726)</td>
</tr>
<tr>
<td>Paved Roads (%)</td>
<td>6.356</td>
<td>9.529</td>
<td>3.512</td>
</tr>
<tr>
<td></td>
<td>(9.453)</td>
<td>(9.909)</td>
<td>(9.117)</td>
</tr>
<tr>
<td>Roads with Improved Surface (%)</td>
<td>53.377***</td>
<td>56.438***</td>
<td>52.03***</td>
</tr>
<tr>
<td></td>
<td>(12.58)</td>
<td>(13.214)</td>
<td>(12.064)</td>
</tr>
<tr>
<td>Slum Housing (%)</td>
<td>-5.643*</td>
<td>-6.988**</td>
<td>-6.982**</td>
</tr>
<tr>
<td></td>
<td>(2.438)</td>
<td>(2.531)</td>
<td>(2.273)</td>
</tr>
<tr>
<td>Inhabitants per Doctor</td>
<td>33.305</td>
<td>31.42</td>
<td>28.228</td>
</tr>
<tr>
<td></td>
<td>(24.621)</td>
<td>(26.005)</td>
<td>(23.633)</td>
</tr>
<tr>
<td>Hospital Beds per 1,000</td>
<td>-30.343**</td>
<td>-35.92***</td>
<td>-35.92***</td>
</tr>
<tr>
<td></td>
<td>(9.524)</td>
<td>(9.841)</td>
<td>(8.76)</td>
</tr>
<tr>
<td>Children per School Place</td>
<td>-5.55</td>
<td>-7.123</td>
<td>-3.269</td>
</tr>
<tr>
<td></td>
<td>(4.723)</td>
<td>(4.951)</td>
<td>(4.548)</td>
</tr>
<tr>
<td>Buses per 10,000 Residents</td>
<td>11.212</td>
<td>14.819</td>
<td>14.786</td>
</tr>
<tr>
<td></td>
<td>(13.636)</td>
<td>(14.317)</td>
<td>(12.845)</td>
</tr>
<tr>
<td>Log Museum Visits per Capita</td>
<td>1.289***</td>
<td>1.242**</td>
<td>1.223**</td>
</tr>
<tr>
<td></td>
<td>(0.375)</td>
<td>(0.4)</td>
<td>(0.359)</td>
</tr>
</tbody>
</table>

Table 8.2: Decomposing Effects of the Three Components of the State History Index. *** significant at the 0.001 level; ** significant at the 0.01 level; * significant at the 0.05 level. Standard errors in parentheses. All coefficients report partial correlation controlling for public goods index in 1994-5, Log GDP per capita in 1995, Ethnic fractionalization, Civic Activism, Moscow dummy variable.
how some constant cause (e.g. social capital, unequal institutions, the protestant ethic) is initially produced by some originating event (e.g. the renaissance, colonial rule, or the reformation, respectively), or how a historical cause leads to the present outcome via process tracing (Collier 2011). If we begin by considering the proximate cause of the relationship between public goods provision and regions with greater state history, four explanations suggest themselves from the recent literature, which are regional bargaining strategies, norms of vertical accountability, the legitimacy of local elites, and regional subnationalism. These will allow us to trace the historical cause of contemporary state capacity in Russian regions.

**Regional Bargaining Strategies**

The theory that regional activism in Russia has been primarily motivated by regional bargaining was first advanced by Solnick (1995), who argued that stronger regions would be able to extract greater fiscal concessions from the federal centre in exchange for remaining within the Federation. This suggestion has been advanced by several subsequent scholars of Russian federalism. Treisman (2000) for example has suggested that following the breakup of the Soviet state, stronger regions of Russia were able to engage in ‘regional fiscal protection’: whereby regional governments collude with local businesses in order to assist in the evasion of federal taxes and regulation. Stoner-Weiss (2006) confirms similar results through fieldwork interviews, documenting the extensive use of local state capacity in order to undermine the efforts of the central government. A potential mechanism by which state history may affect fiscal resources available for investment in public goods, therefore, would be the role that more entrenched regional governments play in yielding zero-sum gains from the federal center on behalf of local elites.

**Norms of Vertical Accountability and Compliance**

A long sociological tradition has argued that state formation has tangible effects on social norms which may make meritocratic and bureaucratic institutions more functional (Weber 1921). Elias for example (1982) has argued that state formation creates tangible changes in social norms, notably in the form of hierarchy, etiquette, and adherence to for-
mal rules, while the works of Foucault (1962, 1977, 1981) are largely studies in how state institutions, such as schools, hospitals, prisons, and workhouses, socialize individuals into ‘policing themselves,’ thereby removing the need for explicit acts of discipline. More recently the theme has been reprised by Scott (2004, 2010), who examines the ways in which states ‘make legible’ their subjects and internalize norms of governance. Consistent with these arguments, studies have found tentative correlations between empirical measures of behavior and the rise of the state: historical studies of homicide rates, for example, have shown a remarkable diachronic relationship between state formation and declines in murder rates (Gurr 1981; Eisner 2001) while contemporary anthropological studies also suggest that areas populated by non-states peoples have high, sometimes exceptional rates of homicide (Keeley 1996). These suggest support for the basic contention that the existence of formal institutions for dispute resolution lead to measurable patterns of cultural change (Greif 1994). As a result of a tradition of state formation, it may be that social norms in long-governed areas are more conducive to state capacity in a range of areas including fiscal compliance, adhere to the rule of law, or the propensity to corruption or graft.

Evidence of differential patterns of compliance between three subsets of the Russian sample are shown in Table 8.3, which shows variation in perceptions of public order and attitudes towards fare-dodging, tax evasion, and benefit fraud. The table shows three subsamples: first, the all-Russia sample; second, results of a special survey of 1,000 respondents from within Tatarstan, a region of Russia with especially high state antiquity; and third, the subsample from Siberia, a region with relatively low state history scores. In Tatarstan, perception of public order and norms of compliance are more widespread than the Russian average; whereas, by contrast, the opposite is found in the case of Siberia.

**Legitimacy of Local Elites**

An additional mechanism linking state history to contemporary state capacity is the relationship between elites and the governed, specifically the likelihood that elites themselves are recruited locally rather than placed by the federal center. Where areas had a longer state history, local elites were more prominent within the Soviet hierarchy, including local administration. Republics such as Tatarstan and Bashkortostan, for example, were
Tatarstan Russia Siberia

<table>
<thead>
<tr>
<th></th>
<th>Tatarstan</th>
<th>Russia</th>
<th>Siberia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feel “very secure” in Neighbourhood</td>
<td>24%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>Robberies “very frequent” in Area</td>
<td>2%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Alcohol in the streets “very frequent”</td>
<td>26%</td>
<td>31%</td>
<td>43%</td>
</tr>
<tr>
<td>Police harassment “very frequent”</td>
<td>1%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Drug sales in street “very frequent”</td>
<td>2%</td>
<td>5%</td>
<td>7%</td>
</tr>
</tbody>
</table>

Confidence in Courts, % Respondents 56% 32% 39%
Confidence in the Police, % Respondents 51% 32% 42%

Never justifiable: Avoiding Fare on Public Transport 40% 29% 37%
Never justifiable: Cheating on Benefits 47% 39% 55%
Never justifiable: Cheating on Taxes 52% 42% 50%

Table 8.3: Public Order and Institutional Perception, Tatarstan, Russia, and Siberia Compared. All items from the World Values Surveys, Wave 6. Tatarstan items from a special subsample survey conducted within Tatarstan, 1,000 respondents.

leaders in the Soviet policy of korenizatsiia, under which members of the titular minority were promoted through the Soviet army, bureaucracy, and the communist party (Martin 2001; Hirsch 2005). It is noticeable that in both cases members of the titular minority governed the region already during the communist era, and later went on to win popular election as president of the republic in the post-perestroika era, whereas in other republics, such as Chechnya or Buryatia, this role fell to an ethnic Russian (Soulemainov 2007). This meant that areas with a greater domestic state history were endowed with political elites better placed to bargain with the federal center, and also arguably more committed to the advancement of the economic and social interests of their local constituents.

Each of these factors – cultures of compliance, the legitimacy of local elites, and regional bargaining strategies – is likely to explain the outperformance of regions with greater endowments of state history. Moreover, these explanations are not mutually exclusive, but if anything, part of a complementary process. An early state history led to the formation
of a local intelligentsia from which to recruit party officials during the Soviet era, and the empowerment of stronger indigenous elites may have enhanced the capacity of post-Soviet regional autonomy movements, which led to more successful bargaining vis-à-vis the federal centre, and subsequently, more resources to distribute to the regional subnationalist support base through provision of public goods. This confluence of vertical accountability, elite legitimacy, subnationalist mobilization and regional bargaining can be seen from comparing the contrasting cases of Tatarstan and Buryatia, which represent respectively a high and a low state history republic from within the Russian federation.

**Process Tracing in Tatarstan and Buryatia**

The Republic of Tatarstan is one of constituent republics of the Russian Federation, located on the eastern edge of the European continent, that is often considered a relative “success story”. The World Bank’s 2009 subnational *Doing Business Indicators* report for Russia, for example, awarded the Tatar capital, Kazan, top place among cities in which to
do business in Russia (World Bank 2009). The city has also been ranked as that within Russia with the ‘highest quality of living’ (Mercer 2012). Ratings by private banks have consistently awarded the Tatar government the highest investment grade among Russian regions (e.g. Deutsche Bank, 2009). Opinion polls similarly demonstrate that the Tatar government enjoys a high degree of legitimacy among its constituents. In one survey 76 percent of citizens surveyed said they trusted Tatarstan’s leadership, and 65 percent felt that the economic situation was ‘good and calm,’ while the corresponding numbers regarding the Russian federal government were both at less than 35 percent. Finally, in interviews conducted by the author with policymakers and civil society activists in the region, the high quality of public goods provision was frequently cited as one of the factors explaining the region’s political stability. Unlike other regions, Tatarstan experienced few delays to wages or pension payments, proactive intervention by the authorities to prevent food price spikes, and extensive provision of social housing during the 1990s. This is not to say that corruption, clientelism and graft do not exist in the Tatar Republic. Yet these appear moderately less prevalent in Tatarstan than in other subjects of the Russian Federation. Some evidence of these differences is reported in Table 8.3, which shows the difference between Tatarstan and Russia on survey items such as the acceptability of tax avoidance and fare evasion, as well as perceptions of crime and institutional performance.

However beyond its relatively high purchasing-power parity per capita income (of $12,793, against a regional average of $8,198), Tatarstan does not have the social conditions typically associated with strong and successful political institutions. First, it is highly ethnically polarized: according to the October 2002 census of the Russian Federation, of Tatarstan’s 3,780,000 residents 51.3 per cent were titular Tatars, while ethnic Russians account for most (41 per cent) of the remainder. As well as being ethnically divided, the region is also religiously split, both between Muslims and Orthodox, and among Islamic denominations. Finally, the region also exhibits low levels of civic engagement. In fact Tatarstan ranks last of all Russia’s 83 regions on the proportion of respondents who had

---

4In the follow-up report of 2012, it was ranked fifth among an expanded sample of 30 cities, citing in particular the ease of “starting a business” and “registering property” (World Bank 2012).

recently engaged in some form of civic activism – with just 16 per cent reporting having done so. The region is also fourth last on engagement in voluntary associations, suggesting somewhat weak networks of local civil society (GeoRatings 2008).

If we are to understand Tatarstan’s record of political stability and efficiency, explanations other than ethnic structure or “social capital” are required. The region’s unusual legacy of historical state formation, tracing to the period from 1445 to 1552 when the Khanate of Kazan rivalled Muscovy for domination of the Volga river delta, must be considered a strong candidate. Indeed each of the four aspects of the link from state history, namely the legitimacy of local elites, norms of vertical accountability, subnationalist mobilization, and regional bargaining strategies, can be derived from Tatarstan’s early state formation. To begin, the existence of an indigenous political elite has its origin in the Khanate of Kazan, whose army and bureaucracy created the first sedentary, urban bourgeoisie. After the Russian victory, members of this Tatar aristocracy were allowed to assimilate into the Russian imperial administration and commercial trades, becoming known as “service Tatars” (Graney 2009). Throughout the Russian imperial period Tatars became a market-dominant minority, and in 1812 owning 90 per cent of Kazan industrial enterprises (Zenkovsky 1960). Among the merchants and artisans of the city civic and educational institutions continued to flourish. The Kazan State University, founded in 1804, formed only the second university of the Russian Empire – later to enrol a young Vladimir Lenin – and Kazan’s mosques and tea-houses remained the center for the intellectual life of the Russian Empire’s Muslim population. Due to the existence of an indigenous intellectual and civic realm, by the early twentieth century Tatars were leaders of the intellectual life of Islamic Russia, and were the basis for a movement known as Jadidism, which sought to reform and modernize Islam. Later, Tatars were the central actors in the ethnic nationalist movement among Volga Muslims within the Russian Empire. Though the Tatar intelligentsia was particularly hard hit by the purges of the 1920s and 30s, Tatars were thus among the key constituencies mobilizing to support Soviet indigenization policies of the 1920s (Zenkovsky 1960).

Due to this legacy of strong, educated and mobilized local elites, under the Soviet era Tatarstan was a leader in the policy of *korenizatsiia*, dedicated to the advancement of “underprivileged” (non-Russian) ethnic groups, and by consequence Tatars entered into senior
positions in the Communist Party, including First Secretary of the Tatar ASSR. This had profound consequences for Tatarstan’s transition in the post-perestroika period. The Tatar elites who dominated the transition years were a essentially continuation of the Tatar political leadership from the Soviet era, already accustomed to acting as mediators between their local base and the capital, as well as representing their domestic Tatar constituents. Thus the first elected President of Tatarstan, Mintimer Shaimiev, was formerly a Soviet apparatchik as well as a Tatar nationalist, who was able to switch allegiance from the Soviet Union to his native Tatar cause. Though the Tatar nationalist movement had a grassroots basis, Shaimiev and the existing generation of Soviet Tatar elites were successfully able to co-opt this movement by appropriating many of its core demands, drawing support away from the more radical Ittifak party, while at the same time attempting not to alienate either the Russian population or Moscow. Upon election, Shaimiev implemented a comprehensive plan for reappropriating the symbolism of Tatar sovereignty, including its own flag, national anthem, airline, foreign ministry, as well as extensive renaming of city streets and institutions and construction of statues to Tatar heroes and poets (Graney 2009). Linguistic reforms and changes to educational syllabi were also implemented. Shaimiev cemented his nationalist credentials by calling for citizens of the republic to boycott Yeltsin’s 1993 constitutional referendum and the subsequent parliamentary elections, enhancing his image as defender of Tatar interests against the federal government. By the time Tatarstan and the Russian Federation reconciled their differences by signing a bilateral treaty in 1994, Shaimiev had won generous concessions from the federal center, which ensured substantial control over Tatarstan’s resource revenues and a steady flow of funds for the purpose of investment in welfare and public infrastructure. Having won concessions from Moscow, investment in local infrastructure became a central means of solidifying his support base among local nationalists and business interests.

In fieldwork interviews in Tatarstan, the quality of public goods delivery was frequently cited by interviewees as a factor behind the legitimacy of both political elites and institutions. This investment cuts across a wide range of domains, from housing, to welfare, 6

---

6This marks a contrast with regions such as Chechnya, which lacked strong indigenous political elites during the Soviet era, and whose rulers were Russian appointees from Moscow with no legitimacy to represent local interests.
to transport infrastructure and museums and cultural institutes. Since the mid-1990s, for example, the Shaimiev administration as spent over $685m on a comprehensive slum clearance policy designed to give 30,000 former inhabitants modern-built apartments on the city suburbs (Graney 2009). Other key investments in public goods since the collapse of the Soviet Union include the construction of a metro system for Kazan, reconstruction of the city center, and construction of a panoply of sports facilities and concert halls that have earned Kazan the moniker as the ‘sports capital of Russia’. In addition to public goods, policymakers and scholars interviewed in Tatarstan highlighted the role of such social programs in maintaining the government’s stability and legitimacy. Social programs implemented in Tatarstan since the mid-1990s include a range of measures to soften the impact of economic transition in the early years of the Shaimiev administration, and additional welfare transfers at later times. The so-called ‘soft entry’ strategy of the 1990s sought to ease the pain of economic transition for citizens of Tatarstan by avoiding shock therapy, pursuing a slower rate of reform and maintaining adequate social protections throughout the process of transition. In practice, this meant maintaining price controls on basic foodstuffs and fuel as well as subsidies to consumers for housing and transport, and giving priority to the agro-industrial sector of the Tatarstani economy by supporting it with extensive government subsidies and credit (Graney 2009). One of Muratov’s first acts was to raise the minimum wage in Tatarstan, and to present a “radical, maybe even revolutionary” program of social reform in Tatarstan based on maintaining the incomes of citizens at a level 1.5 times the subsistence level, as determined by the Tatarstani government (Izvestiya Tatarstana January 24, 1995, 35; Respublika Tatarstan January 25, 1995, 1). Similarly, the Tatarstani government used the Putin administration’s controversial decision to monetize in-kind social benefits, introduced in January 2005, as an opportunity to reinforce the positive benefits that sovereignty has brought to its citizens. While Tatarstan agreed to introduce the new reforms, and was one of only a handful of regions to do so fully and on-schedule, thus allowing the republic to play the role of good federal subject, Kazan also supplemented the new cash payments with additional outlays drawn from the republican budget (Graney 2009). In October 2005 the government announced that the World Bank would be giving the republic over $1.3
Figure 8.5: Location of the Buryat Republic in the Russian Federation.

million to help families living below the poverty line in Kazan.

Why did Tatar elites use these resources to invest in public goods, rather than exclusively expropriate them through corruption or distribute them to supporters in the form of consumption goods? When asked this question, a local newspaper editor simply remarked that it is because ‘the degree of social irresponsibility of our elite is much lower than the degree of social irresponsibility of the Moscow elite’ Yet if this is to be more than a circular understanding, it is necessary to also understand the role of Tatar nationalism, political stability, and the historical existence of a Tatar ‘service class’. Tatarstan is not free of corruption, graft, or the use of public office for private gain. Yet the mentality of the Tatar elites with respect to their region is less that of a rapacious kleptocrat, suddenly given a window to loot and steal - as many post-Soviet leaders arguably have been - and more akin to Olson’s ‘stationary bandit’: rooted in their cultural and historical homeland, and proud of their shared history, its leaders see little trade-off between their private interests, and the public need to invest in the region’s economy, schools, and urban infrastructure.

Whereas in Tatarstan we find the confluence of a long state history, indigenously formed

8Interview with Rashid Akhmetov, 19/04/2012.
political elites, regional subnationalism and a high degree of contemporary state capacity, a very contrasting picture is to be found in the Buryat Republic, located in the southeast of Siberia. Like Tatarstan, Buryatia possesses natural wealth: the region is blessed with extensive deposits of gold, coal, various non-ferrous metals, and other materials. By consequence, Buryatia’s GDP per capita, of $7,071, is only marginally below the Russian regional median of $8,198. Yet despite this economic potential, the republic has an abysmal record on most measures of administrative state capacity. The homicide rate of 26.1 per 100,000 is the third highest in the entire Russian Federation, and slum housing accounts for 7.5 per cent of the total stock, against an average of 4.5 per cent across Russian regions. Incredibly, for a region in which many lack car ownership, there are only 8 public buses per 10,000 inhabitants. For comparison, in Tatarstan there are 53, in Moscow 64, and in St Petersburg, 106. The region is also affected by a wide range of social problems, which the authorities have done little to address. The rate of drug-related crimes committed in Buryatia is reported to be twice as high as Russia’s average. In the late 1990s, more than 1,200 people were registered as drug addicts, with two-thirds of them being young people under 30. Buryatia in the 1990s was also the site of a major public health epidemic, as the number of tuberculosis cases more than trebled from 1991-1996, reaching 123.9 patients per 100,000 of the population, more than double the average figure for Russia of 57.8 per 100,000; this figure reached 2.9 times by the end of the decade. Buryat public officials have consistently been behind in responding to the scale of the mounting crisis.

In addition, in contrast to Tatarstan’s post-Soviet record of public goods provision and relatively good governance, Buryatia’s recent history has also been marked by extensive reports of corruption, graft, and bureaucratic complacency. During the period of late wage and social security payments in the 1990s, Buryatia was listed among the regions with the most severe delays in government payments, with waits between three and seven months. This despite the fact that the Russian Finance Ministry stated in 1998 that Buryatia had already received 539.1 million roubles, of which only 338.2 million it could account for - a

---

9British Broadcasting Corporation, 12/6/96.
10TASS, 2/24/99.
37 per cent shortfall.¹² In 1995, the mayor of Ulan-Ude, Valery Shapovalov, was suspended on grounds of corruption, and subsequently found guilty of tax evasion and document forgeries; a financial inspection of the privately-owned Shapovalov and Company had revealed the concealment of more than 29m roubles in tax payments, as well as a forged payment order.¹³ In 1997, the vice-mayor, Andrei Firsov, was also declared a suspect in at least two major embezzlement schemes, including one in which he allegedly misappropriated 9 billion roubles for ‘the city’s needs’. Subsequently, he was believed to have fled Russia.¹⁴

Not only media reports, but also comparative ratings of institutional quality by investment advisory bodies also rate the region poorly. The official website of the Republic of Buryatia, for example, trumpets the fact that the republic has risen ‘from 56th to 48th place’ on a ranking of investment potential by the Expert RA group (Republic of Buryatia 2012). Yet this omits the fact that most of the performance is due to high ratings for natural resources and ‘tourism potential’, while on ‘infrastructure’ and ‘institutions’ – the two measures which track the quality of public services and governance - the region receives 71st and 62nd place, respectively, from among Russia’s 83 regions (Expert RA, 2012).

Why has Buryatia’s post-perestroika trajectory been so different from that of Tatarstan? Overall, at the collapse of Soviet rule Buryatia was poorly prepared to assume administrative responsibilities, with absent indigenous bureaucracy, weak regional or subnational identity, and little protection against unscrupulous elites willing to take their share of the region’s great natural resources. Whereas the Tatars trace their state history to the Khanate of Kazan, the Buryats are historically cattle-breeding nomads living in felt yurts that they would relocate seasonally between pastures, and did not develop an indigenous urban elite or intelligentsia until the Soviet era. Indeed, the region’s capital, Ulan-Ude, was originally founded by Russian Cossacks, and until 1934 known by its Russian name, Verkhneudinsk.

Whereas Tatarstan was a leader in the movement for autonomy within the Soviet Union, Buryatia was a by-product of Stalin’s nationalities policy. Following the creation of autonomous republics for Tatarstan and Bashkortostan, the Buryat-Mongolian Autonomous

---

¹² British Broadcasting Corporation, 12/12/98.
¹³ British Broadcasting Corporation, 12/20/96.
¹⁴ What the Papers Say, 7/1/97.
Region (BMAR) was created in 1921, but without a clearly identifiable ethnic group. In 1937 the BMAR was arbitrarily detached and merged as the Aginsk Buryat National Area as a part of the Chita Region; it was later reconstituted in July 1958, at the stroke of a bureaucrat’s pen, as the Buryat ASSR. The Buryats themselves can lay only a weak claim to possession of a distinct ethnic identity before the Soviet era, having previously been considered a branch of Mongols, who until 1931 had used the Old Mongolian written language. On the state history index Buryatia has a score of 0.72, half a standard deviation below the median score of 0.78, and more than one standard deviation lower than the score of Tatarstan.

During the Soviet era, Buryatia did not benefit greatly from the program of korenizitsiya, and, facing a lack of qualified candidates of the titular group, and was instead governed by ethnic Russians, as during the Russian imperial era. In the post-Soviet phase, this has continued. Unlike Tatarstan, where an educated and empowered indigenous elite was able to form a platform for defending Tatar interests, Buryatia never saw an effective political mobilization for defending Buryat concerns. In the early 1990s Buryat nationalist parties (the Buryat-Mongol People’s Party and the Negerdel National Unity Movement) were founded to promote the regionalist cause, but disunited and disorganized, failed to gather many votes even among ethnic Buryats. The 1994 presidential elections were won overwhelmingly by Leonid Potapov, an ethnic Russian who in Soviet times was chairman of the Buryat Supreme Council. Though born in Buryatia, Potapov assumed his role after being nominal vice-president of Turkmenistan. Potapov was further re-elected to office in 1998 and 2002, and in 2007, replaced by Vyacheslav Nagovitsyn, another ethnic Russian. Nagovitsyn’s main qualification was that he was former Deputy Governor of Tomsk Oblast, situated 2,000 kilometers to the west (Heaney 2012).

From the start, Buryatia has lacked of domestic elites capable of advancing its sovereign interests, and the inadequate provision of public services, wage payments, and the absence of public accountability have, accordingly, been defining features of the contemporary region. Perhaps particularly characteristic of this maladministration is one episode from the 1990s, during which civic protests and strikes against non-payment of wages were widespread. Rebuffed by the Federal government for having already paid out funds earmarked for the payments, and unable to account for the shortfall, the Buryat Ministry for
Social Welfare and Labor:

... announced plans to allow its pensioners to receive part of their pensions in foodstuffs. The republic has around 21 per cent of the 26 million rubles ($1 million) necessary to pay pensions in November, and local officials were prepared to try to cover some of the payments due with high-quality and fodder grain\textsuperscript{15}.

**Conclusion**

The breakup of the Soviet Union provides a remarkable experiment in testing local state capacity and public goods provision. With the end of centralized rule from Moscow, regional governments were suddenly faced with responsibility for wide-ranging duties of fiscal administration, policing, and revenue mobilization. This chapter has shown that the strongest correlate of success in handling this transition was a region’s historical legacy of state formation, prior to the reform era. Using a new dataset of state history for 83 provinces of the Russian Federation for ten centuries until 1991, cumulative state formation is shown to be the decisive factor explaining success or failure at provision of public goods and the management of public administration.

Chapter 9

Conclusion: Ancient Polities, Modern States

This research has sought to address a number of core questions for political science. Why do patterns of state capacity vary so widely not only between, but within countries? What are the origins of state capacity? And what explains why such patterns can prove persistent over many generations, and through changes in formal administration and governance? By examining district-level variation in state capacity in the Indian subcontinent, I have chosen to study a persistent legacy of states from the precolonial era. Contrary to the widespread perception that India before the onset of colonialism was a “village” or pre-state society, I argue that there existed wide variation in precolonial state centralization. The eighteenth century was a time of rapid defensive modernization across the subcontinent, driven by the requirements of gunpowder weaponry, and the increasing frequency and intensity of warfare among South Asian regimes and against colonial powers. These changes included a range of centralizing fiscal and administrative reforms which restructured state-society relations at a local level, and persisted under the aegis of the colonial regime. Moreover, I argue that this pattern is not unique to the South Asian context, but exists at a cross-country level globally, and can also be found subnationally in another case study, namely across districts of the former Soviet Union. These finding has a number of important implications, both theoretical, methodological, and empirical.
Theoretical Contributions of this Project

First, this research has several important theoretical contributions. A longstanding behavioralist tradition in comparative political science, rooted in the works of Almond and Verba (1963) and more recently, Putnam (1993), has argued that the effect of hierarchical political institutions is to undermine the voluntaristic norms upon which political order depends. The “Tocquevillian” tradition sees the basis for civic compliance in horizontal norms of reciprocity and trust, which for theorists such as Alexis de Tocqueville (1835), were generated in the civic norms of early American settler communities, and for Putnam (1993), in the communal republicanism of renaissance Italy. Such a paradigm remains the preferred perspective for understanding why collective action leads to payment of taxes and the provision of public goods (De Cremer 1999, Ostrom 1999, Herreros 2004, Parks et al. 2013, Glaeser et al. 2000, Anderson et al. 2004, Hermann and Thoeni 2011).

By contrast, this dissertation takes a contrasting view, which may be labeled as “Hobbesian.” I argue that the vertical norms generated by historical political authority also generate a distinct set of norms, a culture of compliance, that is conducive to “making states work” even in the absence of strong affective or solidaristic ties, similar to what Levi et al. (1998) refer to as “cooperation without trust”. The stock of norms of compliance, traditional legitimacy, and authority vested in a polity explains the capacity of states to enforce cooperation among their subjects, raise levies, and enlist manpower, and such “political” capital, if one wills, is a resource, valuable yet distinct from the “social” capital that theorists such as Putnam (1993, 2000) see in horizontal bonds among citizens that are generated through civic association, shared identity and personal ties. The concept of a culture of compliance fills a significant gap in the literature between voluntaristic theories of compliance, based on the microfoundations of reciprocity and trust, and the implication of recent historical studies such as Ertman (1997), Gerner (2012) or Fukuyama (2011), which propose a relationship between historically centralized institutions and contemporary political order, yet without the microfoundations underpinning social capital theories. Whereas social capital emphasizes voluntary association, the “political” capital of compliance is a matter of explicit or implicit coercion; whereas social capital emphasizes equality, political capital entails hierarchy and subordination. The essence of social capital is the capacity of individuals to solve problems of collective action through simple free association; the essence of compliance,
is the resolution of these problems through the “Hobbesian” subordination of one man, or group of men, to another man, or group of men. Such a “culture of compliance” offers a very distinct mechanism for solving the problem of order from the mechanism proposed by theorists of social capital. High levels of societal compliance make possible taxation and investment in public goods; they also make possible many of the negative aspects of state capacity, such as authoritarianism and even genocide. The capacity to achieve one end is the same capacity that achieves others: today’s police state, may become tomorrow’s welfare state, and today’s absolutist regime, raising great revenues for the grandiose vanity projects, may become tomorrow’s developmental state, investing in infrastructure, health and education.

Second, this dissertation shows the relative importance of the precolonial over the colonial past. Discussions of institutional legacies in political science and economics have been dominated by studies which evaluate the impact of the policies and institutions of colonial powers, whereas less weight has been given to the political structures and institutions that colonial authorities inherited (Dell 2010, Herbst 2002). Yet state-building is, in Weber’s phrase, a “slow boring of hard boards,” and such institutional inheritance cannot simply be ignored. The evidence of this dissertation is that in the South Asian context at least, precolonial legacies trump the legacy of efforts at state formation by British colonial authorities. Why should this be so? I consider several, related, hypotheses, including the persistence of precolonial administrative structures, and the ability of precolonial states to war down local-level actors and institutions resistant to the state’s authority. Yet underlying each of these explanations is a more general theory, regarding cultures of compliance and the incentive structures imposed by historical states, following the model of compliance

---

1History furnishes us with many examples of countries which have made the transition from one to the other. Within the context of this project, Travancore in southern Kerala for example is the clearest example of a theocratic regime which extracted vast resources for expenditure on military and religious works, yet which redirected such an apparatus in the nineteenth century towards the goals of public health and education. The Soviet Union could not have achieved its high level of state-directed economic growth from the 1930s to the 1970s were it not for the legacy of Tsarist absolutism, and developmental states in East Asia have drawn upon a long legacy of bureaucratic centralism (Lee 1998). Similarly, the “warfare states” of nineteenth and early twentieth century Europe became the “welfare states” of the latter half of the century, precisely because warfare encouraged fiscal deepening and the sacrifice for the sake of the common good (Skocpol 1992, Tilly 1985). A failing state, however, will never be one nor the other: examples of transition in this regard are more scarce.
outlined in chapter 2. Seen through the lenses of compliance, states lack capacity because public servants and citizens see no reason to comply with its rules, regulations, and edicts. As Pritchett (2009) observes of contemporary India, in many areas “the everyday actions of the field level agents of the state – policemen, engineers, teachers, health workers - are increasingly beyond the control of the administration at the national or state level”. We can then think of the compliance-generating strategies of precolonial polities, such as Mysore or Travancore, and British rule, as a principal-agent dilemma offering two different means of raising compliance with authority. The compliance problem was solved by precolonial ruler such as Haider Ali of Mysore or Marthanda Varma of Travancore through a ruthless system of surveillance and coercion. As long as the center held, people therefore enlisted when they were demanded, paid the taxes that were due, and officials lived in terror of an accusation of embezzlement or fraud.

By contrast, colonial authorities attempted to raise compliance through attempts to increase the flow of information between the districts and the authorities – via the use of land cadastres, court records, and reports such as those compiled in the district gazetteers – thereby, making it more difficult to conceal acts of non-compliance, such as tax evasion or land seizures. While the colonial state did also have means to prosecute cases through the courts – and was certainly described as “despotic” by its critics – it cannot be considered as such in comparison to the genuine despotism of precolonial rulers. The consequence was a state that was “all fingers and no thumbs”, or the phenomenon that Pritchett (2009) refers to as India’s “flailing state,” namely, a central bureaucracy that was very good at analyzing and diagnosing local problems, but limited in its capacity to correct them. In contrast, by acting despotically, precolonial regimes had not only increased the level of surveillance, but also very severely raised the cost of non-compliance. If patterns of political order in contemporary India continue to follow precolonial state boundaries, it is because this is ultimately a more effective strategy: sanction without information is arbitrary, yet information without sanction is toothless.

Finally, an additional theoretical contribution of this work is that it provides external validation to a number of longstanding hypotheses on the causes of state formation that have been developed in the European context, and in particular, to the ‘bellicist’ arguments of Charles Tilly and Otto Hintze, regarding the role of warfare in causing an increase in
state infrastructural power in early modern Europe (Hintze 1906, 1975; Tilly 1975, 1990; Downing 1992). According to this view, “states made war, and war made states,” in that warfare required growing numbers of soldiers, and new technologies of war entailed growing fiscal expenses (Tilly 1992; Roberts 1956; Parker 1976 1996). Until now, such bellicist theories have rarely been successfully applied outside of the European context (cf. Centeno 1997). By examining the validity of bellicist arguments in the context of eighteenth century India, I show that it is possible to replicate the expectations on the argument in a context where warfare and the introduction of gunpowder weaponry had occurred relatively exogenously, and in which causal inference is less muddied by the covariant processes that were occurring in Europe at this time.

Methodological Contributions of this Project

In addition to these theoretical contributions, this dissertation also makes several methodological innovations. First, this dissertation combines econometric analysis of contemporary data with data from archival research, and combines these quantitative data sources with qualitative assessments of institutional variation in precolonial India. Second, and more broadly, this project is the first to take advantage of district-level variation in order to examine persistent institutional legacies from precolonial South Asia, based on the 640 districts of the most 2011 census. India’s administrative districts constitute a logical unit for testing the persistence of precolonial institutions into the colonial era, for, as Brass (1994) notes, “the entire system of British district administration... concentrated authority at the district level” and following decolonization “was retained virtually intact, particularly the central importance of the District Magistrate and the district courts and the police”. In contrast to the growing literature that discusses the effect of institutional legacies at the State and Union Territory level within India (e.g. Kohli 1988, Singh 2010, Verghese 2016), use of district level data allows analysis along the boundaries of precolonial polities, and the examination of persistent variation within India’s post-1956 states. While the number of districts has expanded steadily over time - from 356 in 1971 to 640 by 2011 - this has occurred almost entirely though the splitting of existing districts into two or more sub-entities, with fewer cases of re-apportionment between multiple districts.

Third, in examining precolonial state formation, this dissertation has not only relied
upon existing sources but also generated new data in an attempt to understand the historical origin of patterns of compliance and state capacity across contemporary India. For example, in order to assess levels of bureaucratic compliance, in this dissertation I have presented results of a field experiment within Karnataka, designed to test the degree to which local state agents comply with the legal and administrative demand to respond to requests for information, and also presented an extensive review of primary and secondary literature on eighteenth-century Indian states in order to assess their levels of fiscal capacity, the organization of their executives, and the judicial administration.

Empirical Significance of the Findings

There are also a number of important empirical implications of this research. Cross-country studies of the institutional legacies of colonialism treat India as a single case in a dataset of many cases, this neglects to recognize the fact that, during the “high phase” of imperialism between 1870 and 1945, a majority of the world’s colonial subjects fell under just one single administration – British-ruled India (Figure 9.1). While European empires may have extended their reach from Africa, to the Americas, to the Middle East and Central Asia, India constituted, therefore, a “special case”, a collection of cases nested under a single polity, and by this fact, hidden from comparative analysis.

Underneath the monolithic imperial regime however, there was extreme diversity in the range and development of precolonial political institutions, from pre-state hunter-gatherer societies in the interior northeast, to the centralized, bureaucratic states of the southern Deccan. A number of recent studies of institutional variation have taken advantage of this variation in order to make empirical contributions to the study of the effects of colonial policies, yet to date there have been no district-level analyses of the effect of precolonial regimes in India (Mukherjee 2012, Iyer 2010, Banerjee and Iyer 2005). Such an absence is especially notable given that while political scientists have neglected the importance of precolonial past in explaining contemporary outcomes in India, in recent decades historians have made considerable progress in analyzing this relationship (Bayly 1983, Pati and Ernst 2010, Ramusack 2004, Stein 2010, Roy 2013b, Robb 2002). Whereas both “imperial” and “nationalist” histories of India tended to focus disproportionately upon the positive and negative consequences, respectively, of colonial policies and institutions, in recent
decades revisionist histories have brought out the significance of India’s precolonial era. It is now generally accepted among historians, for example, that colonial era patterns of land administration were largely inherited from Indian precolonial regimes, and that far from being an era of universal stagnation a number of regimes in eighteenth-century India made significant moves towards centralization and defensive modernization (Stein 1984, Roy 2013b, Gordon 1993). Yet these advances in historical understanding of the eighteenth century in Indian history have occurred in relative isolation, and there remains a significant gap in econometric studies of long-run institutional legacies in India.

Conclusion

In this dissertation I have traced the boundaries of precolonial eighteenth-century South Asian polities, and shown that districts of India that lie narrowly within the boundary lines
of historically centralized states perform significantly better today on a wide variety of
district-level indicators of state effectiveness than those narrowly outside these boundaries,
despite the fact that these borders largely ceased to exist in the early nineteenth century.
These estimated effects are robust to a wide variety of controls, placebo tests for border
displacement, the exclusion of individual polities, and controls for the boundaries of In-
dia’s contemporary federal states. In order to explain these persistent patterns of variation,
I propose a new theory, based on the effect of historical states in generating persistent cul-
tures of compliance, and that such persistent cultures are the basis for contemporary state
capacity in the world today. Such patterns of were generated in India during the period of
intense warfare that racked the subcontinent in the eighteenth century, and has persisted
through to the present day.

I verify this persistent legacy of precolonial states using a combination of archival re-
search, district-level colonial data on taxation and public goods from 1853 and 1901, and
a field test of bureaucratic responsiveness conducted in the state of Karnataka. Using ex-
tensive archival research on the fiscal and bureaucratic structure of Indian states in the
eighteenth century, I argue that following the decline of the Mughal Empire, warfare be-
tween “challenger states” prompted an accumulation of bureaucratic and fiscal capacity at
the local level, and that this capacity has persisted through the colonial era to the present
day. In contrast to “bottom-up” theories of state capacity which root institutional strength
in societal characteristics such as ethnic homogeneity, shared identity, or social capital, it
is argued that government effectiveness is cumulatively built through long-term historical
investments in state capacity, and that, in India, an important phase of investment occurred
during the warring states period of the eighteenth century.

As discussed in this chapter, the findings of this research have a number of important
implications for political science, and for policy makers more generally who are concerned
with the existence of persistent gaps in state capacity and the ability of developing country
states to raise taxes, provide goods, and deliver the rule of law. At first glance, the finding
that the best predictor of public goods delivery at the district level in India today is the level
of one century ago, would not appear very optimistic for policy makers seeking to effect
change within the rather shorter duration of their political mandates. Yet policy makers con-
sistently overestimate the impact of their policies in the short run, while under estimating
their impact over the longer horizon. In modern India, central governments have consistently failed to challenge the structure of incentives at the local level, instead focusing on short-term, cosmetic reforms while abdicating responsibility for policy implementation downwards. Under the colonial regime, for example, efforts to reform governance were stalled after 1858, with the uprising of 1857 vitiating policymakers’ enthusiasm for risking reform to land administration or the displacement of local elites by centrally appointed bureaucrats. In post-independence India, policymakers have continued in this vein, devolving governance through federalism and panchayat reforms, while announcing ambitious policies to tackle shortages of food, shelter, or sanitation, that are implemented only at the discretion of statewide politicians, who in turn devolve implementation to districts and panchayats. As Gunnar Myrdal observed in the 1960s, in India many laws have been “intentionally permissive” with the consequence that while “social legislation [has] pointed the direction in which society should travel” they have left “the pace indeterminate” (Myrdal 1968). The consequence of the de facto and de jure decentralization of administrative responsibility is that local institutional legacies have largely persisted, and that lack of local state capacity has been the problem preventing more effective service delivery in India. Indeed, to the extent that public funds are provided from the center in order to meet policy demands, this in many ways has exacerbated the institutional cleavage, as well-governed districts are able to use these funds for investing in public goods and services, while in low capacity districts the funds are misappropriated and used to further entrench the authority of existing elites, who have often sought to undermine state capacity (Wilkinson 2004, Suryanarayan 2015).

However, simply because current patterns of state capacity date to the precolonial era, does not entail that persistent non-compliance equilibria must endure indefinitely, or that these cannot be shifted. It only entails that the last major efforts to reform local institutions date from the centralizing efforts of warring states from 1707-1857. The evidence of the eighteenth century suggests that change occurred relatively rapidly in districts that were subject to the mobilizing initiatives of precolonial rulers, and that, such rapid change could occur again if institutional incentives were imposed which ensured stricter penalties for non-compliance and more effective infrastructure for bringing about this outcome. The combination of democracy and decentralization reforms, of course, are one attempt
to provide reform incentives ‘from below’ via the ballot box, yet have proven limited in the context of a society with low levels of literacy, low political expectations, and riven by clientelistic ties based on personal association, religion and caste. Where such reforms fail, then the only possible source for institutional change must be “from above,” in the creation of clearer lines of vertical accountability, linking local authorities to the more responsive agencies of Indian public administration. And here there is a more optimistic implication of the current research. India is not uniformly deficient in state capacity, and just as there are pockets of state weakness, so too are there pockets of excellence – both spatially, and also hierarchically, i.e. at differing levels of competence and within different administrative units of the state (Pritchett 2009, Oldenburg 1987). Leveraging such persistent pockets of capacity is key to addressing persistent pockets of state weakness in India today, and in other polities suffering from extreme spatial inequalities in governance.
Bibliography


Johannesburg, South Africa.


Khan, I. (2004). *Gunpowder and Empire: The Indian Case*. Indian History Congress Section.


Lindsay, J. H. (1964). “Local Self-Government”, in *The Cambridge History of India*. Lon-


London.


Publications.


Institute.


