# Indirect rule and the unequal burden of persistence: Evidence from Hyderabad *

Varun K

Harvard University

October 27, 2026

**Abstract**

We study the distributional consequences of colonial institutions in India by analyzing the effect of direct and indirect rule on the outcomes of historically disadvantaged and advantaged caste groups using a border discontinuity design. We find that disadvantaged caste groups in indirectly ruled areas have 1.4 fewer years of schooling and significantly less wealth than similar groups in direct rule areas. A similar analysis for the Advantaged caste in indirect rule areas suggests that they have small and mostly insignificant differences in education and wealth. Cross-caste wealth inequality and land inequality are higher in indirectly ruled areas. Differences in extent of agricultural labor repression, as measured by agricultural wages and bonded labor, led to the emergence of these differences during the colonial era. These differences persist partly due to differential growth in credit institutions after independence. Indirectly ruled areas have fewer banks per grid cell post independence, and this negatively affected productivity. As a result, farmers in these areas have larger non-institutional debt than directly ruled areas, and disadvantaged caste farmers drive this effect.

Keywords: caste, indirect rule, India, colonialism.

JEL Classification: O15, F54, N45, D72, O43, Z13, D6, Q15

*I am grateful to Arun Chandrasekhar, Rachel Glennerster, Lakshmi Iyer, Chinmaya Kumar, Michael Kremer, M.R. Sharan, Raul Sanchez de la Sierra and especially James Robinson and Eduardo Montero for valuable comments, advice and guidance. I benefited greatly from the comments of seminar participants at the University of Chicago. Dhruvika Dhamija provided excellent research assistance. This work would not have been possible without the data entry work of Ajeet Singh, Aman Raj, Trilok Kumar Jha, and Gufran Ahmad.

# 1 Introduction

Previous work in political economy has argued that historical institutions cast a long shadow by impacting the evolution of later institutions, and this causes a divergence in outcomes (Acemoglu et al. (2001), Nunn (2008), Banerjee and Iyer (2005), Dell (2010), Iyer (2010), Lowes and Montero (2020), Mamdani (1996)). While the fact that differences in institutions can lead to a persistent difference in outcomes is well acknowledged, which social group within the native society bears the burden of this persistence is not as well studied. We try to answer this question by looking at the long-term effects of colonial institutions on different caste groups across the boundary of princely state of Hyderabad in modern day India.

In India, colonial institutions were categorized into two types: direct rule, where the colonial power governed territories directly, and indirect rule, where they governed through a local ruler loyal to the colonial administration. In principle, these two types of rule can differ not only in their policies but also in their implementation. Indirect rule may involve working with traditional elites, while direct rule could disrupt the existing social structure. Indirect rule could empower the advantaged groups, while direct rule could have a more egalitarian impact. It can also be the case that indirect rulers may have been more aware of social inequities and better equipped to address them, while direct rulers might have been cautious about disturbing existing hierarchies. As a result, the distributional consequences of direct and indirect rule are uncertain. Understanding these consequences becomes especially important when these institutions are imposed on top of the existing hierarchy in society, which was the case with the caste system in India.

The caste system that orders social groups into a hierarchy originated around 1000-1200 BC after the arrival of Aryans in India (Sharma (1958)). The social groups were created from the population through the practice of caste endogamy (Ambedkar (1916)). These groups specialized in different occupations, and were organized in a total order on the basis of the purity of their occupation. The groups at the bottom of this hierarchy were subjected to extreme forms of discrimination, including untouchability, by the groups above them in the hierarchy. In the modern era, the outcomes of these social groups vary significantly. Advantaged castes in India have 35% more years of education compared to disadvantaged castes. There is significant geographic heterogeneity in these educational disparities — 0.4 years in the state of Maharashtra, 1.7 years in

Andhra Pradesh and 2.6 years in Bihar. This study will examine the role that colonial institutions played in widening/narrowing the gaps between advantaged and disadvantaged groups.

In this paper, we attempt to answer two questions - *Firstly*, when there are persistent differences in outcomes due to colonial institutions (direct and indirect rule in our case), which group in the social hierarchy bears the burden of these differences? *Secondly*, how do colonial institutions affect present-day inequality? These two questions together answer the question of the distributional consequences of colonial institutions.

The British colonial rule in India exercised direct rule in some areas called presidencies, while they implemented indirect rule in princely states. In these princely states, native kingdoms were governed with British protection. We focus on the princely state of Hyderabad (henceforth *Hyderabad*), which a native ruler, the *Nizam* of Hyderabad, ruled before independence, and compare its outcomes to nearby areas under the British using a geographic regression discontinuity design. Consistent with the idiosyncratic manner with which the historical boundaries of the state of Hyderabad were determined, we demonstrate that areas inside *Hyderabad* are geographically similar to the British areas just outside of the boundary.[1] We use individual-level response data from the Demographic and Health Survey of 2015-16 for our primary analysis.

In answering the first question, we look at the heterogeneity in outcomes for different caste groups across the boundary of the princely state. While the number of caste groups in an area can vary, there are four large categorizations into which all castes fall.[2] The top-most in this hierarchy is the group of advantaged castes (also known as upper castes), and the bottom three (SC, ST and OBCs) are historically disadvantaged castes. Hence, we classify the scheduled castes, scheduled tribes and the other backward caste individuals as belonging to disadvantaged castes. We classify the upper castes as advantaged castes.[3]

[1] Figure 1 shows the division of India into princely states (in yellow) and presidencies (in pink) in 1909. The princely state of Hyderabad was surrounded by British presidencies on all sides, with the Bombay Presidency towards the west, Central Provinces and Berar to the North and Madras presidency to the East and South.

[2] Here we do not refer to the four caste groups in the *chaturvarna* system. Instead we refer to the categorization used by the India government to implement affirmative action policies.

[3] While it is the Hindus who have been practising the caste system for a long time, other religions have caste too. Followers of Sikhism, Christianity, Islam and Buddhism in India practice caste as well. However, the intensity of this practice is lower for religions other than Hinduism. We do not include Muslim upper castes in the sample of advantaged castes because there was an out-migration of this group from Hyderabad after the *Nizam*'s rule ended in 1948 (Sherman (2011)). While there was out-migration of all Muslims at the time of independence, this was higher in Hyderabad because the state had a Muslim ruler with a significant Muslim bureaucracy which had roots in North India. This bureaucracy largely belonged to advantaged caste.

We find that the disadvantaged castes in the indirectly ruled areas (the princely state) have large differences in educational and wealth outcomes compared to disadvantaged caste groups in geographically close direct rule (British) areas. These castes in the indirect rule areas have 1.4 fewer years of education, 15% less literacy rate and 0.2 $\sigma$ less wealth than similar caste groups on the other side. These results are robust to a number of analyses, such as alternative RD specifications, bandwidths, kernels, analysis along different parts of the princely state, standard error adjustments to account for spatial auto-correlation (Conley (1999)) and spatial placebo simulations (Kelly (2021)). Importantly, the differences in outcomes for advantaged caste groups across the boundary are small for education. The estimates for literacy rates are close to zero and those for number of years of education are negative, but small and insignificant at 10% level. The differences for wealth are negative but insignificant for all specifications except one. These differences are small compared to what we observe for disadvantaged caste groups and are not as robust to alternate specifications. In particular, when we use survey weights in our analysis, the above documented estimates for disadvantaged groups remain negative and significant, but the estimates for advantaged castes flip sign and become positive. The negative estimates for advantaged castes are not robust to placebo noise simulations (Kelly (2021)). This lends support to the theory that the social groups towards the bottom (the disadvantaged castes) bear the burden of differences in outcomes caused by indirect rule.

To answer how colonial institutions affect present day inequality, we calculate caste wealth inequality using the DHS data at a cluster level and compare it across the boundary of the princely state. Caste wealth inequality measures the difference in median wealth of advantaged and disadvantaged caste groups at the cluster level. We find that areas in the princely state are more unequal on this measure. As land is an important factor in our setting, we compare land inequality across the boundary of the state. We find that areas in the former princely state have higher land inequality as measured by inter-quartile ratio, and median absolute deviation from the median. While caste land inequality is higher in indirect rule areas it is not as robust to different specifications.

Having established that the disadvantaged caste population in *Hyderabad* bears the unequal burden of persistence in contemporary times, we look at intermediate outcomes. Differences in literacy and agricultural yield (a proxy for wealth) had arisen by independence in 1947. Literacy

rates for both disadvantaged and advantaged groups were lower at the time of independence compared to similar groups on the other side of the boundary.[4]

Taken together, our findings suggest that there is an overall negative impact of being in *Hyderabad*, both in the colonial era and in contemporary times. Advantaged groups fare slightly worse than comparable groups on the other side, while disadvantaged groups bear the brunt of this negative impact. The overall negative effect could be because of the differences in the agrarian relations between advantaged and disadvantaged groups across the boundary, or other factors, such as lack of public goods provision. However, the disproportionate negative impact on disadvantaged castes implies that the aforementioned relation may have an important role to play. To understand the driver of these differences, we turn to mechanisms. The mechanism behind these differences should be able to explain the following facts — the emergence of differences in education (for all caste groups) and yield during the colonial rule; the persistence of differences in educational and wealth outcomes across the boundary of *Hyderabad* for disadvantaged castes from independence till contemporary times, but the lack of persistence in educational outcomes and relatively weak persistence in wealth outcomes for advantaged castes. We propose three mechanisms — labor repressive institutions during the colonial era, differences in public good provision during the colonial era, and credit institutions that continued to the independence era — which together explain these facts.

Historical work (Thirumali (2003), Pavier (1981), Iyengar (1931))) has described the exploitative labor practice of *vetti* that was pervasive in *Hyderabad* under colonial rule, where individuals, mostly from disadvantaged castes, performed free agricultural labor for the local landlords.[5] Thirumali (2003) and Mukherjee (2021) argue that areas which were given as grants to military chieftains were more exploitative and repressive. Therefore, we decompose the effects of indirect

[4]We use the data from the census of 1931. Unfortunately, the census does not differentiate between OBCs and non-Brahmins, and hence we restrict our analysis to SC/ST (disadvantaged castes) and Brahmins+Jains (advantaged castes).

[5]While the practice of forced/bonded labor (such as the *bhagela* system) was imposed upon a subset of disadvantaged groups, mostly untouchables, *vetti* was applicable to a broader group of poor individuals in the village (Iyengar (1931) page 125). While caste groups in OBCs were higher in the caste hierarchy than SC/STs, they were socially and economically backward compared to the advantaged castes, which contributed to their indebtedness to the large landlords, who almost always belonged to advantaged castes. Some castes among the OBCs were landed, but the DHS data which we use for our analysis does not have information about the subcaste of individuals which prevents us from excluding economically dominant OBCs from our analysis. As long as these dominant groups were doing better on average than other subcaste within disadvantaged caste, pooling them together will provide an underestimate of the effects of labor repression. It must be noted that there is no dataset which contains both subcaste and geographically granular data on outcomes for India.

rule by the intensity of labor repression. Our analysis shows that contemporary effects of indirect rule are worse for disadvantaged castes in areas where the intensity of labor repression was higher. To directly test for differences in labor repression, we look at the presence of bonded labor in contemporary data, and find that households in indirectly ruled areas are twice more likely to have a member who was released from bonded labour.[6] Agricultural wages in indirectly ruled areas were 30% lower after independence after controlling for price and geographic factors. Taken together, these three pieces of evidence suggest that indirectly ruled areas were more labor repressive than directly ruled areas. As agricultural labor was performed mostly by disadvantaged castes, this difference in repression had a persistent negative impact on the outcomes of these groups.

Differences in labor repression do not explain why advantaged caste groups had lower literacy rates in 1931. To address this question, we examine the provision of education infrastructure and find that direct rule areas had more primary and high schools than indirect rule areas at the time of independence. However, contemporary data does not show a difference in education infrastructure, indicating that indirectly ruled areas closed the gap. This suggests that for the disparities in disadvantaged castes to persist, either repressive labor institutions continued after independence or another factor caused the persistence. The effects of labor repressive institutions can be exacerbated by imperfections in financial institutions, as these can push individuals into debt bondage (the fact that indirect rule areas have a higher fraction of bonded laborers supports this hypothesis). Additionally, it can lead to an increase in the concentration of land, and thereby inequality, when a debtor is unable to pay back the loan to the moneylender (who was very often the landlord) and is forced to give up the land. Given this intrinsic link between forced labor and debt in agrarian economies, we analyze the difference in post-independence credit institutions.

We find that indirect rule areas had 65% fewer banks per grid cell in 1980, and 40% fewer banks per grid cell in 2005, than direct rule areas. Less access to formal credit meant that farmers in indirect areas were more reliant on informal credit through moneylenders. We show that this was indeed the case using farmer debt data from NSS 59.[7] We find that the differences in non-

[6] Bonded labour, a practice in which a debtor enters into an agreement to render labour to the creditor either by himself or through any member of his family, for a specified or unspecified period, either without wages or for nominal wages, was prohibited in India in 1976.

[7] The National Sample Survey or the NSS is an annual survey done on a representative sample of the population.

institutional debt across the boundary are higher for disadvantaged castes than for advantaged castes. These differences in credit institutions can drive the inequality that we observe through channels highlighted in theoretical work by Galor and Zeira (1993) and Banerjee and Newman (1993).

**Related literature and contributions** Extractive institutions in hierarchical societies can create conditions for those at the top to exploit the groups at the bottom. We provide evidence that social groups at the lower end of the hierarchy do worse due to colonial institutions and this was because the elite groups were able to exploit them due to the weakness of the indirectly ruled state. This is the most important contribution of this paper.

Several studies have highlighted the disparities in economic outcomes of different caste groups in India (Deshpande (2001), Thorat and Neuman (2012)). This paper shows that some of the geographic variation in the gap in socio-economic outcomes between disadvantaged and advantaged caste groups is due to differences in colonial institutions. Lee (2017) is another paper which has analysed colonial rule through the caste lens. They compare the contemporary economic fortunes of landed caste groups under British areas which were annexed at different times (wartime vs non wartime) with non-landed caste groups and find that landed caste groups in areas annexed during European wars faced a reversal in outcomes compared to non-landed caste groups. While they are concerned with the effects of timing of annexation on outcomes of landed caste across the country, we are concerned with the effects of direct and indirect rule on outcomes of disadvantaged and advantaged caste groups.[8]

This study also contributes to the literature on long-term effects of labor-repressive agriculture. Political scientists and economists have tried to understand the role that labor-repressive agriculture plays in the emergence of different types of social order (Moore (1984)). The landlords in *Hyderabad* dominated the local economy and enforced extremely oppressive labor practices (Thirumali (2003), Mukherjee (2021)). The social order that emerged as a result of these repressive institutions increased the subordination of disadvantaged caste groups to advantaged caste

[8]We do not look at landed and non-landed caste groups but instead at historically advantaged and disadvantaged groups. Our mechanism focuses on labor repression by large landlords and even small landed peasants were expected to perform *vetti* (Iyengar (1931), Pavier (1981)). There were caste groups within the backward castes (OBC) who were landed. But they were not part of the landed aristocracy who were the drivers of forced labor practices.

groups. While the social hierarchy was similar outside the state of Hyderabad, the extent of dominance of the advantaged caste groups and repression of disadvantaged caste groups was higher within the state, causing a higher divergence in outcomes of disadvantaged caste groups.

Previous work has highlighted that banks play an important role in reducing poverty (Burgess and Pande (2005)) and improving health (Cramer (2021)) but their long-term impacts have not been explored. We show that former British areas had much better access to banks after independence compared to *Hyderabad*. This mechanism provides empirical support to theoretical work by Galor and Zeira (1993) which shows that capital market imperfections can lead to differences in steady state output (wealth in our setting) and investment in human capital across regions which differ in initial wealth distribution (which was the case at the time of independence across the boundary of the princely state).

Previous work by Jain (2017) has emphasized the importance of language instruction for educational outcomes of the population. The paper shows that parts of India in which the language of instruction did not match the vernacular language during colonial era had lower literacy rates after independence. *Hyderabad*'s official language under the *Nizam* was Urdu, while the local population spoke one of Telugu, Marathi or Kannada, and hence the mechanism highlighted in Jain (2017) may have been relevant in our setting as well. However, differences in agricultural productivity and subsequently in wealth cannot be explained by this channel. The differential catch up rates across caste groups that we have shown is hard to explain with this mechanism without the presence of differences in other variables.

Lastly, this paper follows many studies which have looked at the long-term effects of colonial institutions in India (Banerjee and Iyer (2005), Iyer (2010), Colleoni (2023), Ratnoo (2024), Chaudhary et al. (2020), Chaudhary and Garg (2015), Chaudhary (2010)) and the effects of historical institutions across the world (Becker and Pascali (2019), Mendez and Van Patten (2022), Grasse (2023), Dell and Olken (2020)). Our outcomes utilizes granular data from DHS IV while most of the previous literature uses aggregate data at the district level. In a recent paper Colleoni (2023), using granular data, shows that British ruled areas in Central India do better on average than nearby indirectly ruled areas. While our overall findings align, the primary focus of our paper is to underscore the variation in this persistence across historically advantaged/disadvantaged caste groups in Indian society.

Banerjee and Iyer (2005) and Iyer (2010) use data from the whole country and employ an instrumental variables strategy to estimate their effects, whereas we focus on just one princely state and use a geographic RDD to compare areas which are close to each other and are therefore, arguably more comparable. In Online Appendix I we show that the IV approach to study colonial institutions in India would give correct results in the finite sample of (modern day) Indian states only if the sample is restricted to states where the overlap condition holds. This means that only those (modern day) Indian states can be used for analysis where there exist treated districts (created through the instrument) and untreated districts. We show that the results for Iyer (2010) change significantly (see Appendix Table I-1) when we impose this condition.[9]




Figure 1: Political map of India in 1909 . Boundary of the princely state of Hyderabad is in yellow

This paper is organized as follows. In section 2, we provide a brief history of the Princely state of Hyderabad which will help establish the plausible exogeneity of our empirical framework. We will describe the land and revenue administration of the state in detail and explain how it led to an increase in intensity of labor-repressive agricultural practices. Section 3 will lay out

[9] In the geographic RD our condition translates to requiring that the treated and control sample do not lie along the boundary of a modern day state (or institution which may lead to divergence in outcomes for the two groups).

the empirical framework we adopt and estimate the long-term adverse effects of being part of *Hyderabad* on different caste groups. Section 4 shows intermediate outcomes and channels that drive these results and section 5 concludes.

# 2 Setting

After seizing Bengal from the Nawab in the Battle of Plassey in 1757, the British gradually brought other native rulers under their control over the course of the next 100 years. Some of these rulers became tributaries to the British and paid them a yearly sum. In other areas called the presidencies, the British had complete administrative control. On the other hand, the princely states were administratively ruled by native rulers, although their foreign policy was controlled by the British. Hyderabad was one such princely state that occupied a large land area in the south-central part of the subcontinent.

## 2.1 A short history of the Hyderabad state

The boundary of the state of Hyderabad that we use for our analysis was settled in 1853. The area that came to be the state has a long and rich history. It was originally part of the Bahamani Sultanate (founded in 1347). Aurangzeb's Deccan charge led to the fall of the dynasty and large parts of the Deccan plateau became part of the Mughal Empire. Aurangzeb appointed Mir Qamaruddin as the Viceroy to the Deccan. After Aurangzeb passed away at Ahmadnagar in 1707, the Viceroy declared himself ruler of the Kingdom of Hyderabad in 1724. Both the British and *Hyderabad* area that we use in our analysis were part of the Kingdom of Aurangzeb in 1688.

The Asaf Jahi dynasty, whose ruler was known as the *Nizam*, ruled Hyderabad till independent India's army forced them to surrender in September of 1948. During their rule of around 225 years, the Nizams were tributary to different rulers, first the Marathas and then the British. Even though the internal affairs of the state were under the *Nizam* the British had an influence on the administration of the state. The British appointed Residents to the princely state whose job was to ensure that the interests of the East India Company and later the Crown were not being undermined.[10] The Prime Minister, who was the most important administrator in the state, was

[10] British India came under the direct rule of the crown after the first war of Independence in 1857.

appointed by the *Nizam* after conferring with the British Resident. Throughout its existence, the state of *Hyderabad* reeled under severe financial constraints due to a constant state of war. These financial constraints led to the Prime Minister, Salar Jung I, formulating a new land revenue policy in 1853.

**Administration in Hyderabad** After the land reforms implemented by Salar Jung I in 1853, around 50% of land came under the *Dewani*, where the revenue was collected by the state machinery. The majority of land in these areas was under *ryotwari* system which involved direct taxation of cultivators. However, the system became perverted over time as landlords (*deshmukhs*) were able to acquire lands from individual cultivators rendering them tenants. 10% of the land was *sarf-e-khas* or meant to collect taxes for Nizam's private purse. This revenue was not used for the administration of the state. The remaining 40% of the land came under the *Jagirdari* system which was a feudatory setup within the princely state where each *Jagir* operated as a separate state under a *jagirdar* appointed by the *Nizam* (Bhanumathi (1992)). The *Jagirs* were land grants made to individuals who had provided military services to the state and served the purpose of establishing the authority of the ruler by placating the military chiefs (Nayakwadi (2006)). The *Jagirdars* had absolute power in these areas and they implemented policies of revenue extraction from the peasants without spending on providing administrative infrastructure to help them. According to Bhanumathi (1992) the cultivators in *Jagir*areas were subject to *oppressive and arbitrary rule unknown to the modern world*.[11]

**Labor repressive institutions** After Salar Jung's reforms, the *deshmukhs* (landlords) were removed from their position of revenue collectors and were instead given ownership of land in proportion to the area they previously collected revenue from (Pavier (1981)). During the land survey that followed the reforms, the *deshmukhs* were able to acquire deserted and forest lands. Developing and cultivating this land required clearing, irrigation, and fertilizers, which only larger landlords could afford. These landlords served as moneylenders and this role became more important towards the end of the nineteenth century when the state required revenue to be paid in cash (Pavier (1981)). When cultivators were unable to repay their debts, the landlords seized their

[11] This statement is made about the conditions of the cultivators in the 1920s.

land and forced them into tenancy or agricultural labor. According to an enquiry conducted by S.K. Iyengar in 1931, 25% of wet land in Warangal district had been transferred in the 25 years preceding the inquiry, largely to *deshmukhs* (Iyengar (1931)).[12] Around 50% of the families surveyed were in some kind of non-land mortgage debt, and 70% of these loans were taken on personal security or personal services (Iyengar (1931) page 32).[13]

Thirumali (2003) (page 27) states that *landlords forcibly transformed the large mass of rural poor into servile labor*. Most of those forced into labor belonged to historically disadvantaged caste groups. Due to the high tax rate imposed by the state, only large landlords were able to generate profits. This was primarily because they had access to free labor of *vetti*, which was not available to small landlords. While some groups within backward castes may have owned land, most of the large landlords belonged to advantaged castes. Disadvantaged castes had limited occupational opportunities, such as small-scale agriculture, agricultural labor, or traditional caste-based occupations. This made them highly vulnerable to the demands of local landlords. On the other hand, advantaged caste individuals, who had higher levels of education and social capital, and often belonged to the same subcaste as the local landlord, were protected from exploitation by local landlords.

An important historical event serves as evidence of the extent of repression of small peasants and agricultural laborers in the state of Hyderabad. In July 1946, peasants in Hyderabad rose up in rebellion against the *Nizam* and his policies. The rebellion was sparked by an incident in Jangaon taluka, where Visnur Ramachandra Reddy, the *deshmukh*, forcibly tried to take possession of land belonging to a washer-woman, Ailamma (Gupta (1984)).[14] The local peasants prevented that from happening. In the fight that ensued, Doddi Komarayya, the local peasant leader, died from a bullet fired by the *deshmukh's* men. The peasants, who were fed up with the exploitation they had suffered at the hands of the landlords, took up arms with the help of the Andhra Mahasabha and the Communist Party of India. Some attacked the houses of landlords, and others refused to perform forced labor. The rebellion lasted for four years, during which the peasants occupied and governed large tracts of land (Sundarayya (1972), Pavier (1981))).

[12] The corresponding figure for dry land for Aurangabad was 29%.

[13] The report likely underestimated the extent of repression by landlords and moneylenders. Iyengar was concerned about the influence of these landlords and their ability to discredit his inquiry if he were to report the true state of affairs (Pavier (1981) page 7).

[14] *Reddy* is one of the dominant and advantaged caste group in *Hyderabad*, where as caste groups whose traditional occupation includes washing clothes, are historically disadvantaged castes.

**Land Administration outside the state of *Hyderabad*** There were three directly ruled presidencies which bordered *Hyderabad* — Bombay, Madras and Central Provinces & Berar. In Madras, the British largely collected tax directly from the cultivators through the *ryotwari* system, though some districts in the Northern and Central parts had the *Zamindari* system as well (Ratnoo (2024)). The tax was high, and cultivators frequently defaulted, often falling prey to moneylenders (Raju (1974)). Similarly in Bombay presidency, the British implemented direct taxation. In the 1870s cultivators rioted in the district of Poona and Ahmadnagar, as their lands were being taken over by usurious moneylenders. The British administration passed the Deccan Agriculturalists Relief Act in 1879 which put a check on the moneylender's power (Chaudhary and Swamy (2017)). This was in sharp contrast to Hyderabad where landlords had become the sole providers of credit and justice in *Dewani* areas in the absence of the state (Thirumali (2003)). The British on the recommendation of the first Law commission brought the Indian Penal Code (IPC) in 1862 which was only applicable in direct rule areas. Section 374 of the IPC banned forced labor. *Hyderabad* passed a similar act prohibiting forced labor, but only in 1936 (Pavier (1981)). On the whole, the landlords in the direct rule areas surrounding *Hyderabad* were small, and agricultural practices were less oppressive.[15]

Thus, the administrative differences between the direct ruled British areas and indirect ruled *Nizam* areas could have led to differences in the agricultural institutions. The state of *Hyderabad* was weak, creaking under financial burden imposed by constant wars and taxes paid to the British which allowed *Jagirdars* and *deshmukhs* to exploit the rural peasants and laborers, most of whom belonged to the disadvantaged castes. The dominance of landlords was limited not only to agriculture but also to credit markets, which allowed them to lend at usurious rates, forcing peasants to become tenants and forced laborers for the landlords. While the labor institutions were undermined (to some extent) after independence, differences in credit institutions between the former direct and indirect rule areas could have caused a persistent difference in outcomes.

[15]We focus on the relative and not on the absolute level of oppression. There is no doubt that agriculture was labor repressive outside *Hyderabad*.

## 2.2 Borders of the Hyderabad state

The borders of the Hyderabad state were drawn as a result of wars after the death of Aurangzeb. Towards the west, the Marathas gained huge chunks of territory in wars. These territories were won from them by the British in the Anglo-Maratha Wars.[16] The western and northern boundary was settled after the Third Anglo-Maratha war in 1819 (except for Berar). Towards the east, the East India Company expanded its territories in the second half of the $18^{th}$ century. In 1765, the area north of Guntur to Vizag was obtained from the *Nizam* by Robert Clive using a treaty of alliance. The area south of Guntur which was won by the *Nizam* in the Anglo-Mysore wars was ceded to the British in 1796 in exchange for a military alliance. The state of Berar, northwest of Hyderabad, was ruled by the Nizams till 1853 after which it was taken over by the British as payment for military protection. The border of *Hyderabad* that we use is after the annexation of Berar.

The area under *Hyderabad* as well as that surrounding it was part of the Mughal territory of Aurangzeb, and the state's borders of 1853 were carved out after several wars and cessation. To the extent that this happened over territory and not the kind of land, there should not be any differences between treatment and control areas at the beginning of 1853. Evidence from the Governor General of India's correspondence suggests that he wanted to bring Hyderabad under their rule, indicating that *Hyderabad* was not disadvantaged compared to surrounding areas.[17] The British however did annex Berar in 1853 from *Hyderabad* in lieu of non-payment of military debt. That raises concerns about selection. In the Online Appendix, we show results for areas excluding the districts that were under Berar and find that the results do not change. The northeast border of the princely state overlaps with the border between the state of Telangana and Chattisgarh, and Telangana and Maharashtra in modern day India. We remove data along this border for analysis.

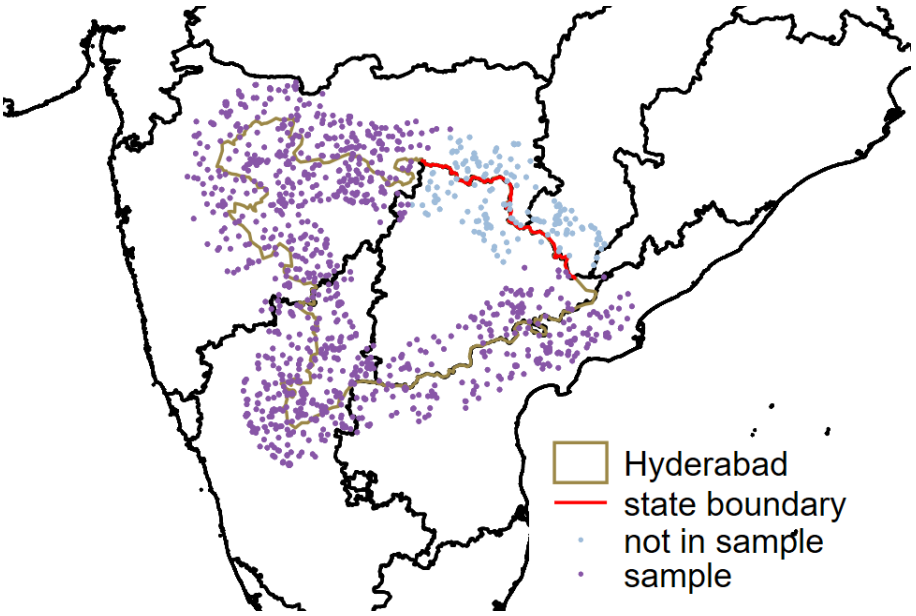
Borders formed from wars tend to overlap with natural geographic boundaries such as plateaus and rivers. The eastern border of the state of Hyderabad is separated from the British territories by the Krishna and Tungabhadra rivers flowing from southwest to northeast. The river provides water to both treatment and control districts and there is no difference in altitude on the two sides of the border. To make sure that all our effects are not being driven by this segment of the state we

[16] The Marathas ruled over a large part of the subcontinent up to the Northern plains in 1760.

[17] "I have got two other kingdoms on hand to dispose of — Oudh and Hyderabad. Both are on the high road to be taken under our management." Lord Dalhousie's correspondence (Rahim (1963))

conduct robustness checks by not including the segment along the river in our analysis.

Figure 2: Location of clusters




Notes: The figure shows location of all clusters within 75 km of the boundary of *Hyderabad*. The state boundary between Maharashtra, Chattisgarh and Telangana is in red. We do not include clusters near this boundary in our sample.

# 3 Results

## 3.1 Data

We use the political divisions from the Imperial Gazetteer of India (Hunter (1907), see 1) to identify the boundaries of the princely state of Hyderabad. This is then superimposed on a map of India from 2003 using GIS to identify the areas that will be part of the sample. Data on education and wealth is taken from Demographic and Health Survey (henceforth DHS) of 2015. Data on geographic characteristics of various clusters is taken from the DHS covariates manual (Mayala and Dontamsetti (2018)). The data on banks is taken from SHRUG (Asher et al. (2021)). Soil nutrient and crop suitability data for balance checks has been accessed from FAO Global Agro-Ecological Zones dataset.[18] We use the Census of 1911, 1931 and 1961 for data on literacy rate, public infrastructure and banks. We use the Agricultural Wages of India dataset from the 1960s to estimate difference in wages across the boundary of *Hyderabad*. More details are provided in

[18]The data was accessed from https://gaez.fao.org/

Online Appendix A.

### 3.2 Estimation Framework

We employ a geographic discontinuity framework to estimate the long-term distributional effects of being part of *Hyderabad*. Our estimating equation is:

$$Y_{id} = \alpha + \beta Hyderabad_d + \delta X_{id} + f(geography)_d + \eta_d + \epsilon_{id} \tag{3.1}$$

where $Y_{id}$ is the outcome for individual $i$ in cluster $d$, $Hyderabad_d$ indicates if the cluster was part of the princely state of Hyderabad, $X_i$ are control variables such as age and gender. $\eta_d$ captures fixed effects of the boundary segment. We divide the area in our sample into seven segments. $f(geography)_d$ is an RD polynomial of the latitude and longitude of the cluster or district (depending on availability of data). We show the results for two types of polynomial for our main set of outcomes: i) a local linear polynomial in distance to the boundary estimated separately on each side of the boundary. We use a triangular kernel and calculate the optimal bandwidth using the MSE-minimizing procedure suggested by Cattaneo et al. (2020). We also show results for a fixed bandwidth of 75 km. This follows previous work in the literature by Lowes and Montero (2020), Bandiera et al. (2022); ii) Following the seminal work by Dell (2010), we also show results for a linear polynomial in latitude and longitude for a bandwidth of 75 km and 50 km.

We show two sets of results — by caste groups and overall. For the former, we divide the population into two groups: disadvantaged castes, which includes the scheduled castes/tribes and backward castes, and advantaged castes, which includes non-Muslim upper castes. We compare the outcomes of each of these groups under indirectly ruled areas with directly ruled areas. The overall results include the entire sample. We use cluster level geographic coordinates (1 for $\sim$ 30 observations) for our main outcomes. For variables where we do not have granular observations (this includes agricultural yield and farmer debt), we use the centroid of districts as a proxy for geographic location. Our results do not change substantially for outcomes where we have both cluster level and district level coordinates. That should support our argument that outcomes for which we have only district level coordinates would still hold if we had access to cluster level

coordinates.

The DHS dataset that is the source of outcome variables randomly displaces geographic coordinates by 2km for all urban clusters, by 5 km for 99% of rural clusters and by 10km for 1% of rural clusters. The observations which are close to the boundary might be displaced to the other side. This might lead to errors in estimation when we use data close to the boundary. To overcome this, we use a donut hole approach and estimate our specification by removing data within 10 km of the boundary (Barreca et al. (2011), Lowes and Montero (2020)) and show the results in Online Appendix Table B-1, C-2.

In a potential outcome framework, the identifying assumption for an RD design is that the potential outcome under treatment varies continuously at the boundary. With treatment varying discontinuously at the boundary, any change in the outcome at the boundary can then be attributed to the treatment. For our identifying assumption to hold, the potential outcome must be a continuous function at the boundary. This requires that outcomes of individuals around the time of British annexation, when the treatment of being in *Hyderabad* started, should be similar as well as variables that could drive these outcomes should be similar. Therefore, variables like altitude and soil fertility should not be very different. Our empirical analysis will compare areas which are close to the boundary of *Hyderabad* and therefore the balance on important geographic and agricultural variables should hold in areas close to the boundary. As there does not exist a granular historical dataset of geographic variables, we proxy for them with contemporary values from the DHS covariates dataset (Mayala and Dontamsetti (2018)).[19] To test for differences in agricultural productivity prior to 1853, we use agro-climatic potential yield data from GAEZ (GAEZ (2021)). This data has been used in other works (for instance Adamopoulos and Restuccia (2022)) in economics to test the role of geography in agricultural productivity. The data is generated by combining cell-specific land quality attributes with agronomic models, for a given level of water supply and cultivation inputs. Potential yields indicate how geographical attributes translate into productivity by crop.[20] This data is imperfect, but a good proxy of potential yield of an area.

In Table 1 we show the difference in geographic, soil and agricultural variables between *Hyder-*

[19] Heston (1973) uses data on yield for different crops from 1897 onwards. However, this data is at the district level and not granular enough for an RD analysis.

[20] Adamopoulos and Restuccia (2022) states that *We stress that the agronomic model parameters are not based off a regression analysis of observed choices on outputs and inputs across countries, regions, or farms, an analysis that would be subject to serious endogeneity issues.*

Table 1: Balance on Geographic and Agricultural variables

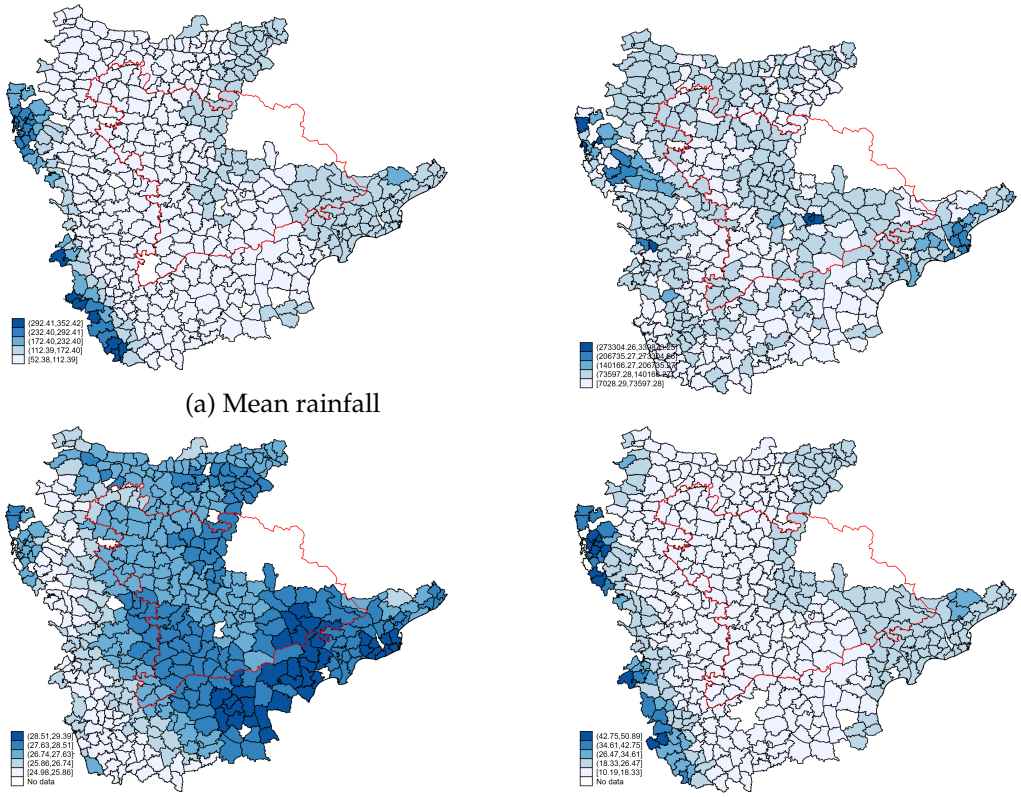
| | Coefficient | SE | Mean | Observations |
|---|---|---|---|---|
| Geographic Characteristics: | | | | |
| *Rainfall* | 4.65 | (2.95) | 111.55 | 573 |
| *Temperature* | 0.09 | (0.12) | 27.10 | 573 |
| *Maximum temperature* | 0.11 | (0.14) | 33.03 | 573 |
| *Aridity* | 0.35 | (0.32) | 18.59 | 573 |
| *Slope* | -0.02 | (0.13) | 0.72 | 573 |
| *Elevation* | -4.20 | (25.68) | 393.91 | 573 |
| *vegetation* | -78.75 | (120.38) | 2298.37 | 573 |
| *Episodes of drought* | -0.04 | (0.20) | 6.69 | 573 |
| Soil Characteristics | | | | |
| *Soil workability* | 0.11 | (0.23) | 6.18 | 680 |
| *Dominant soil type* | -0.60 | (1.05) | 29.68 | 291 |
| *Oxygen availability to roots* | 0.01 | (0.24) | 7.97 | 425 |
| Agro-climatic Potential Yield | | | | |
| *cotton Potential yield (1961-90)* | 0.79 | (1.55) | 188.77 | 2024 |
| *rice Potential yield (1961-90)* | 5.08 | (14.68) | 965.51 | 2018 |
| *wheat Potential yield (1961-90)* | 8.21 | (47.51) | 528.39 | 2023 |
| *Sugar cane Potential yield (1961-90)* | 13.47 | (42.65) | 2679.67 | 1838 |
| *sorghum Potential yield (1961-90)* | 5.83 | (18.10) | 1882.91 | 1921 |
| *Groundnut Potential yield (1961-90)* | 2.51 | (8.46) | 877.77 | 1964 |
| *Pigeonpea Potential yield (1961-90)* | 8.80 | (9.91) | 1448.75 | 1772 |
| *Pearl millet Potential yield (1961-90)* | -0.12 | (2.10) | 865.38 | 1772 |
| *Total growing period* | 1.16 | (1.53) | 145.40 | 788 |

The unit of observation is the NFHS cluster for geographic characteristics. Soil and Agro-climatic potential yields are at the 9 km* 9 km grid cell level. The Table shows the estimates of equation (1) using procedure suggested by Cattaneo et al. (2020). Optimal bandwidth was chosen using this procedure, and estimates were weighted with a triangular kernel. For geographic variables we use the average of optimal bandwidth for all variables. The optimal bandwidth is calculated separately for the remaining variables. We control for boundary segment fixed effects. Variable sources and definitions are in Online Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

*abad* and non-*Hyderabad* areas.[21] We find balance on all geographic, soil and agricultural variables suggesting that areas on either side of the *Hyderabad* boundary were similar. The geographic variables include — temperature, rainfall, aridity (ratio of precipitation to evaporation), elevation, slope, vegetation and episodes of drought in the last 30 years. The soil characteristics include soil workability, the class of the dominant soil type in a grid and availability of oxygen to the roots. These are important determinants of agricultural productivity, which can impact development outcomes in an area. We also show balance on the potential yield of the major crops grown in our sample after accounting for agro-climatic factors. Figure 3 maps the rainfall, temperature, aridity

[21] The geographic characteristics are at the DHS cluster level. A cluster in DHS corresponds to the smallest sampling unit. On average a cluster has 30 individual observations. Soil and agricultural variables are at a 9km * 9 km grid level and are accessed from GAEZ v4.

Figure 3: Balance on geographic features




(a) Mean rainfall

(b) Mean temperature

(c) Land aridity

Notes: The figure presents RD maps for the main geographic variables. The mean value of each outcome variable is calculated at the block level and plotted on a map of block boundaries.

and population at the subdistrict level. It is clear from these figures that none of rain, temperature or aridity jump at the boundary. However, there is a variation in these variables across the area we analyse. That geographical variation is accounted for in our analysis as we compare areas that are close to the boundary where there is little difference. Online Appendix Figure F-1 shows the local linear plots of these balance variables with distance to the boundary.

The second important assumption for the RD approach is that there is no selective sorting across the RD threshold when the state borders were established. There is a possibility that due to oppressive labor institutions in *Hyderabad*, there might have been selective migration out of the state. In Online Appendix H we show that there was little migration during colonial times and the rate of migration would have to be really high to explain our results.

### 3.3 The unequal burden of persistence

If all social groups were affected equally by institutions in the princely state, then all caste groups would lag behind comparable caste groups on the other side of the boundary. The burden of persistence is borne equally in such a scenario. However, if certain caste groups do worse compared to similar caste groups on the British side and some don't, it implies that the institution impacted one group more than the other.

In Table 2 we look at the educational and wealth outcomes for caste groups separately — disadvantaged caste, which includes the scheduled castes (SC), scheduled tribes (ST), backward castes (OBC), and advantaged castes.[22] The first two columns show the estimates for a specification that controls for a polynomial of geographic coordinates for two different bandwidths, 50 km and 75 km. This follows work by Dell (2010).[23] The next two columns show estimates using a local linear polynomial in distance to the boundary for a fixed bandwidth of 75 km and an optimal bandwidth (Cattaneo et al. (2020)). Our primary specification is the one with optimal bandwidth, and we use this throughout the paper unless otherwise mentioned. Our results indicate disadvantaged caste groups have worse education and wealth outcomes than similar groups on the other side of the boundary. The outcomes for advantaged castes are different. Advantaged castes in *Hyderabad* have similar literacy levels to comparable groups on the other side of the boundary. The estimate for years of education is negative but we cannot reject the null for any specification at the 10% level. Robustness checks in Online Appendix show that the estimate is small and close to zero for most of the other specifications (Online Appendix: C-4). Therefore, we conclude that differences in educational outcomes for advantaged caste groups are close to zero. Advantaged caste groups have lower wealth levels than similar groups on the other side, though the coefficient is statistically insignificant at 10% level for all specifications except one. The negative results are not as robust for areas close to the boundary (Online Appendix : C-4). When survey weights are used in the analysis (Online Appendix C-1), the results for advantaged castes flip sign for some specifications but the results for disadvantaged castes remain strongly negative. To understand how the estimates for different castes measure against each other, we plot the standardized effects

[22] The SCs are more commonly referred to as *dalits* and the STs as *adivasis*. In our analysis we restrict advantaged castes to non-Muslim upper castes.

[23] Gelman and Imbens (2019) suggest that higher order polynomials should not be used in geographic regression discontinuity analysis and hence we show the results for a linear polynomial of latitude-longitude.

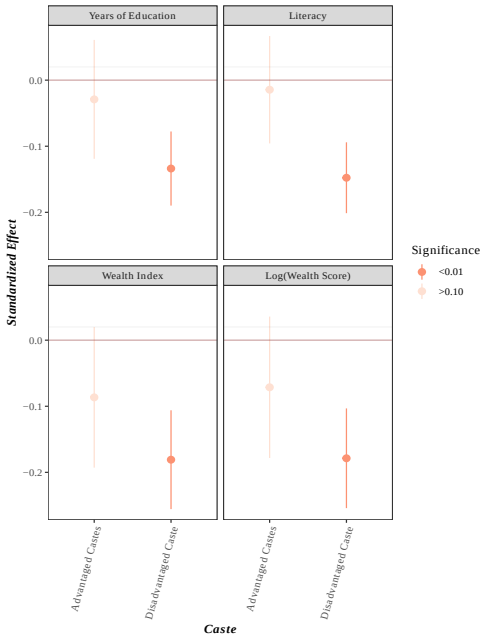
in Figure 4 for different caste groups. Figure 5 plots the linear trend of the education and wealth outcomes for different caste groups with distance to the boundary. These plots reinforce the estimates we showed in Table 2. In Online Appendix Figure H-4 we decompose the educational outcomes for both caste groups by gender. For advantaged castes, these are close to zero for both men and women. For disadvantaged castes, the effects are being driven by both men and women suggesting that selective migration is unlikely as that tends to be male-dominated.

Table 2: Estimates by caste groups

| | *Disadvantaged Castes* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.974*** | -0.883*** | -1.376*** | -1.400*** | -0.082*** | -0.086*** | -0.141*** | -0.148*** |
| | (0.155) | (0.174) | (0.294) | (0.428) | (0.013) | (0.015) | (0.026) | (0.033) |
| Observations | 21,245 | 15,042 | 21,245 | 11,053 | 20,920 | 14,783 | 20,920 | 13,909 |
| Clusters | 1,113 | 781 | 1,113 | 573 | 1,113 | 781 | 1,113 | 734 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 36.35 | 75.00 | 50.00 | 75.00 | 46.77 |
| Mean Dep. Var. | 6.334 | 6.108 | 6.334 | 6.005 | 0.665 | 0.651 | 0.665 | 0.648 |
| SD Dep. Var. | 5.088 | 5.066 | 5.088 | 5.048 | 0.472 | 0.477 | 0.472 | 0.478 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.241*** | -0.231*** | -0.427*** | -0.413*** | -0.076*** | -0.076*** | -0.139*** | -0.145*** |
| | (0.051) | (0.056) | (0.090) | (0.134) | (0.017) | (0.018) | (0.030) | (0.044) |
| Observations | 13,220 | 9,337 | 13,223 | 6,733 | 13,220 | 9,337 | 13,223 | 6,970 |
| Clusters | 1,113 | 781 | 1,113 | 564 | 1,113 | 781 | 1,113 | 583 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.40 | 75.00 | 50.00 | 75.00 | 37.05 |
| Mean Dep. Var. | 2.980 | 2.921 | 2.980 | 2.897 | 12.259 | 12.240 | 12.259 | 12.233 |
| SD Dep. Var. | 1.168 | 1.152 | 1.168 | 1.140 | 0.387 | 0.386 | 0.387 | 0.381 |
| | *Advantaged Castes* | | | | | | | |
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.099 | -0.003 | -0.283 | -0.735 | 0.007 | 0.010 | -0.012 | -0.028 |
| | (0.249) | (0.292) | (0.446) | (0.595) | (0.017) | (0.020) | (0.034) | (0.045) |
| Observations | 6,052 | 4,215 | 6,052 | 3,151 | 6,000 | 4,188 | 6,000 | 3,955 |
| Clusters | 767 | 530 | 767 | 396 | 764 | 530 | 764 | 502 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 36.35 | 75.00 | 50.00 | 75.00 | 46.77 |
| Mean Dep. Var. | 7.765 | 7.650 | 7.765 | 7.741 | 0.781 | 0.771 | 0.781 | 0.770 |
| SD Dep. Var. | 4.853 | 4.871 | 4.853 | 4.932 | 0.413 | 0.420 | 0.413 | 0.421 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.030 | -0.040 | -0.223* | -0.277 | 0.002 | 0.000 | -0.057 | -0.083 |
| | (0.079) | (0.090) | (0.132) | (0.169) | (0.026) | (0.029) | (0.042) | (0.053) |
| Observations | 3,808 | 2,669 | 3,811 | 1,970 | 3,808 | 2,669 | 3,811 | 2,035 |
| Clusters | 767 | 530 | 767 | 390 | 767 | 530 | 767 | 402 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.40 | 75.00 | 50.00 | 75.00 | 37.05 |
| Mean Dep. Var. | 3.336 | 3.276 | 3.336 | 3.300 | 12.364 | 12.347 | 12.364 | 12.355 |
| SD Dep. Var. | 1.219 | 1.216 | 1.219 | 1.219 | 0.390 | 0.396 | 0.390 | 0.392 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p < 0.10, ** p <0.05, *** p<0.01.

Figure 4: Standardized estimates for different castes




Notes: The figure presents standardized RD coefficients for our main outcomes for different caste groups. Regressions control for age, age squared, household size, gender, and boundary segment fixed effects. We use a fixed bandwidth of 50 km and use a triangular kernel.

**Robustness checks** We estimate our main specification for different bandwidths (Online Appendix: C-4), for different specifications of the RD parameters (Online Appendix: C-5), with Conley SE (Online Appendix: C-7, C-6), after excluding Berar (Online Appendix: C.4), with a donut hole (Online Appendix: C-2), after excluding different segments (Online Appendix: C-3). The pattern that we have documented above holds for most specifications. In particular, while advantaged caste groups do have small negative effects for wealth, these are not robust for distances close to the boundary and to the spatial noise simulations of Kelly (2021) (see C-5). The results for disadvantaged caste are robust to spatial noise simulation exercise (see C-5). This leads us to the conclusion that the burden of persistence of colonial institutions is borne by disadvantaged caste groups.

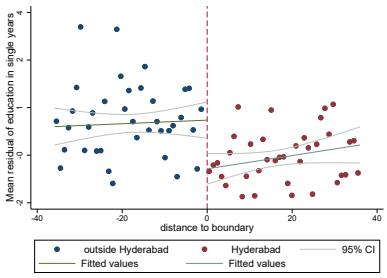
The castes that make up the SC, ST, and OBC may vary across the boundary of the princely state and cultural differences in these caste groups may drive the evolution of their long-term outcomes. However, there is unlikely to be a sharp discontinuity in the type of castes at the boundary because the areas being compared are only 15-20 km apart from each other. To address

remaining concerns, we list the largest caste groups within advantaged and disadvantaged castes for each segment of the boundary in Online Appendix E-2 using data from the census of 1931. The table shows that for all segments, the major caste groups on each side of the boundary are the same. We address concerns regarding selective migration of caste groups in Online Appendix H. Historical data on migration suggests that individuals were migrating both into Hyderabad from nearby provinces as well as out of Hyderabad (Online Appendix Figure H.1). In Online Appendix Figure G-1 we repeat our analysis for the princely state of Mysore. The results indicate that disadvantaged castes in Mysore do slightly worse than similar castes in neighboring direct rule areas and advantaged castes do significantly better.

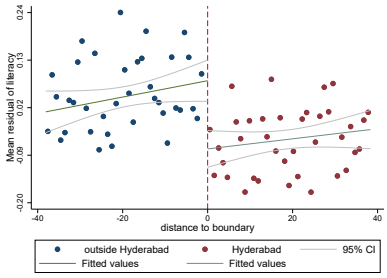
Whether these results hold for princely states other than Mysore will depend on the peculiarities of colonial institutions in those states, which may or may not be similar to *Hyderabad*. However, given that *Hyderabad* was the largest princely state in the colony, with its own postal and railway system, its effect on the weighted average of a similar analysis for all princely states is going to be significant.

**Inequality** We analyse the effects of indirect rule on inequality in Table 3 and Online Appendix D. We show results for four measures of inequality — caste inequality, inter-quartile range, 90%ile - 10%ile ratio and median absolute deviation from median — at the cluster level. Caste inequality is the difference between the median outcome of advantaged and disadvantaged castes at the cluster level. Table 3 shows that there is higher levels of land and wealth inequality in indirectly ruled areas. In Online Appendix D we show results for other measures of land and wealth inequality and perform robustness checks. Our results indicate that land inequality is higher for measures of inter-quartile range and median absolute deviation from median. The results are not as robust for caste land inequality. Similar analyses for wealth inequality indicate that indirect rule areas are highly unequal.

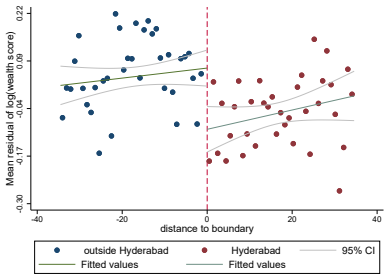
Figure 5: RD plots of outcomes for difference caste groups



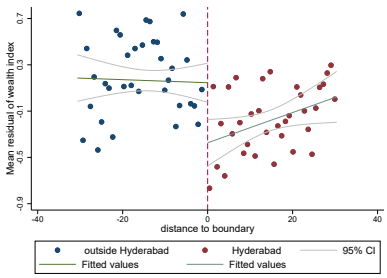

(a) Years of education: disadvantaged



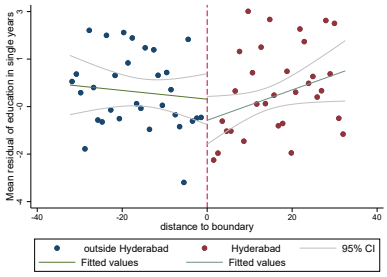

(b) Literacy: disadvantaged



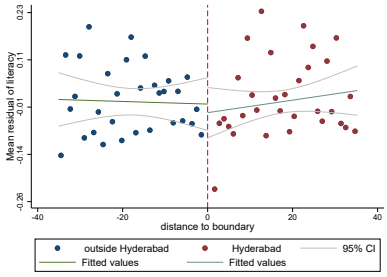

(c) log(wealth): disadvantaged



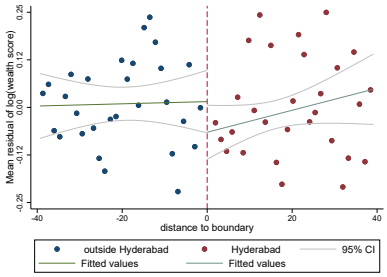

(d) Wealth Index: disadvantaged



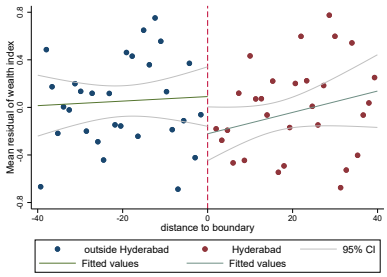

(e) Years of education: advantaged




(f) Literacy: advantaged




(g) log(wealth): advantaged




(h) Wealth Index: advantaged

Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 1 km bin along the running variable with a local linear trend estimated separately on each side of the discontinuity. Figures a-d show the results for disadvantaged caste groups and e-h show the results for advantaged groups. Each regression is estimated using the optimal bandwidth chosen through the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

Table 3: Inequality

| | *Caste Wealth Inequality* | | | | *Median abs deviation from median - Land Inequality* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | 0.088*** | 0.103*** | 0.145*** | 0.155*** | 0.099*** | 0.105** | 0.237*** | 0.282*** |
| | (0.020) | (0.023) | (0.038) | (0.045) | (0.038) | (0.042) | (0.067) | (0.087) |
| Observations | 751 | 522 | 751 | 571 | 1,125 | 786 | 1,125 | 623 |
| Clusters | 206 | 151 | 206 | 163 | 234 | 172 | 234 | 144 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 54.82 | 75.00 | 50.00 | 75.00 | 39.69 |
| Mean Dep. Var. | 0.094 | 0.093 | 0.094 | 0.096 | 0.349 | 0.368 | 0.349 | 0.377 |
| SD Dep. Var. | 0.246 | 0.248 | 0.246 | 0.245 | 0.437 | 0.433 | 0.437 | 0.436 |

The unit of observation is the 20km*20km cluster. Standard errors are clustered at the block level. Linear polynomial regressions include a linear polynomial of grid latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications control for boundary segment fixed effects. Caste wealth inequality measures the difference in median wealth of advantaged caste and disadvantaged caste at the cluster level. Median absolute deviation from median is the measure used for land inequality. * p < 0.10, ** p <0.05, *** p<0.01.

### 3.4 Overall outcomes

Table 4 shows the estimates of equation 3.1 for educational and wealth outcomes for the entire population (including both advantaged and disadvantaged castes). The results in the first panel indicate that areas inside the princely state have significantly lower levels of education across all specifications and bandwidths. Individuals just inside the former princely state are estimated to have approximately 1.3 fewer years of education than individuals just outside. The second panel of Table 4 shows that individuals inside the former princely state have approximately 15% less wealth than individuals outside *Hyderabad*. Online Appendix Figure B.8 shows the standard RD plots for our main outcomes of interest, with distance to the border as the running variable and a local linear trend to each side of the discontinuity.

**Robustness Checks** In the Online Appendix we show robustness to selective annexation of Berar by the British (Online Appendix B.4), to Conley standard errors (Conley (1999)) (Online Appendix B.2), to excluding districts on the boundary of river Godavari and Krishna (Online Appendix B.5), to alternative RD specifications, different RD parameters, alternate bandwidths, and excluding segments, for a quadratic polynomial (Online Appendix B.10), and with a donut hole to

Table 4: Overall outcomes

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.860*** | -0.758*** | -1.228*** | -1.316*** | -0.064*** | -0.065*** | -0.114*** | -0.131*** |
| | (0.142) | (0.161) | (0.265) | (0.405) | (0.011) | (0.013) | (0.022) | (0.033) |
| Observations | 30,095 | 21,144 | 30,095 | 14,097 | 29,639 | 20,808 | 29,639 | 15,154 |
| Clusters | 1,125 | 786 | 1,125 | 523 | 1,125 | 786 | 1,125 | 571 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.24 | 75.00 | 50.00 | 75.00 | 35.67 |
| Mean Dep. Var. | 6.668 | 6.477 | 6.668 | 6.493 | 0.696 | 0.683 | 0.696 | 0.681 |
| SD Dep. Var. | 5.027 | 5.028 | 5.027 | 5.037 | 0.460 | 0.465 | 0.460 | 0.466 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.207*** | -0.207*** | -0.413*** | -0.360*** | -0.060*** | -0.063*** | -0.128*** | -0.134*** |
| | (0.048) | (0.053) | (0.082) | (0.128) | (0.015) | (0.017) | (0.027) | (0.041) |
| Observations | 18,645 | 13,109 | 18,648 | 8,165 | 18,645 | 13,109 | 18,648 | 8,898 |
| Clusters | 1,125 | 786 | 1,125 | 488 | 1,125 | 786 | 1,125 | 533 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 30.07 | 75.00 | 50.00 | 75.00 | 33.33 |
| Mean Dep. Var. | 3.073 | 3.012 | 3.073 | 3.003 | 12.287 | 12.268 | 12.287 | 12.264 |
| SD Dep. Var. | 1.188 | 1.174 | 1.188 | 1.164 | 0.389 | 0.389 | 0.389 | 0.384 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

account for random displacement of clusters by DHS (Online Appendix B.1). The results hold and are significant for most specifications. In Online Appendix Table B.7 we show that our wealth results hold for the Facebook wealth (Chi et al. (2022)) and nightlights data (Henderson et al. (2011)) taken from Asher et al. (2021).

### 3.5 Intermediate Outcomes

We use the census of 1931 to compare educational outcomes across the boundary of *Hyderabad* at the block level. The census has separate categorizations for SC, ST, Brahmins, Non-Brahmin Hindu and other religions. Unfortunately, the census does not separately categorize backward castes within non-Brahmin Hindus. Therefore, we show results for SC/ST and Brahmins+ Jains,

assuming that they reflect the trend for disadvantaged and advantaged castes respectively. Table 5 shows that literacy levels for both advantaged and disadvantaged groups were lower in *Hyderabad* in 1931. This is in contrast to our contemporary results where advantaged castes had no difference in literacy rates. We will show in section 4 that these differences are driven by differences in provision of primary and high schools. Next, we look at intermediate wealth outcomes. We use agricultural yield as a proxy for wealth. Given that agriculture was the largest component of income, this is a reasonable assumption. Table E-3 in the Online Appendix shows the yield for different crops between 1956-61, 8 years after Hyderabad joined the Indian union. We use district level measure of productivity for different crops and restrict our analysis to border districts. Results indicate that yield of almost all crops was lower in *Hyderabad*. This suggests that differences in wealth had emerged by independence.

Table 5: Literacy rates in 1931

| | *ST/SC Literacy rate- 1931* | | | | *Brahmans+Jain Literacy rate- 1931* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.013*** | -0.011*** | -0.009*** | -0.012*** | -0.099*** | -0.107*** | -0.120*** | -0.124*** |
| | (0.002) | (0.003) | (0.003) | (0.004) | (0.015) | (0.014) | (0.023) | (0.027) |
| Observations | 194 | 137 | 194 | 133 | 194 | 137 | 194 | 133 |
| Clusters | 35 | 28 | 35 | 27 | 35 | 28 | 35 | 27 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 47.44 | 75.00 | 50.00 | 75.00 | 47.44 |
| Mean Dep. Var. | 0.013 | 0.012 | 0.013 | 0.013 | 0.384 | 0.371 | 0.384 | 0.371 |
| SD Dep. Var. | 0.014 | 0.015 | 0.014 | 0.016 | 0.093 | 0.091 | 0.093 | 0.090 |

Literacy rates are calculated at the block level for different groups using the census of 1931. Standard errors clustered at the district level. Linear polynomial regressions include a linear polynomial of grid latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications control for boundary segment fixed effects. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

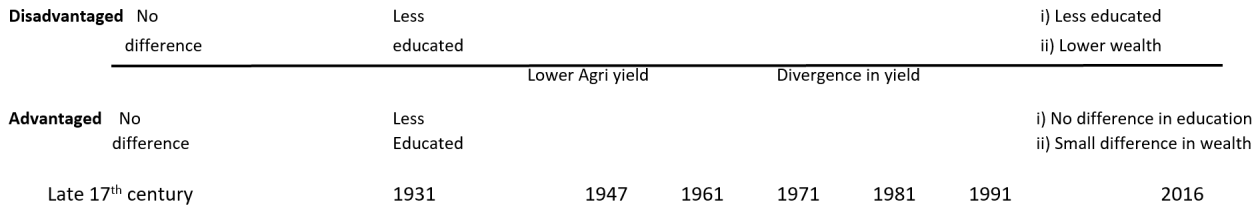
# 4 Mechanisms

## 4.1 Conceptual Framework

The indirect rule under the *Nizam* of *Hyderabad* was characterized by financial turmoil and a weak administration which was not able to keep the extraction of *Jagirdars* and *deshmukhs* in check (Dhanagare (1974), Thirumali (2003)). The mechanisms driving the documented differ-

ences should explain the following facts — i) development outcomes in the indirect rule area of *Hyderabad* are worse and these areas have higher inequality ii) these differences are driven by disadvantaged caste groups iii) Advantaged caste groups have little to no differences in education and small differences in wealth. The timeline of these results is summarized in Figure 6 below. The fact that the boundary of *Hyderabad* was formed arbitrarily after the death of Aurangzeb in late $17^{th}$ century, and there is no difference in agricultural or geographic characteristics across the boundary (as shown in Table 1) make it highly probable that there was no difference in education and wealth of all caste groups across the boundary before *Hyderabad* was formed.

Figure 6: Timeline of documented results




Notes: The figure shows the timeline of results we have documented. All comparisons are between indirectly ruled *Hyderabad* and directly ruled British areas.

Drawing on theoretical research on inequality and growth (Galor and Zeira (1993), Banerjee and Newman (1993), Loury (1981)), we propose three mechanisms that collectively explain our findings. These are labor-repression, credit institutions and public goods.

Previous research has suggested that differences in the provision of public goods under direct and indirect rule can have significant implications (Iyer (2010), Dell (2010), Dell and Olken (2020), Chaudhary and Garg (2015))). These public goods may include education infrastructure, roads, railways, and other factors that can influence growth. Our first hypothesis is that the observed differences cannot be solely explained by the presence of public goods, unless certain extreme conditions are met. The second hypothesis posits that differences in public goods are necessary to explain our findings. We will address the latter hypothesis first.

Advantaged castes in *Hyderabad* had lower literacy levels than similar groups in direct ruled areas in 1931. In a world where agents invest in human capital or seek work this can be true if either the former was not available (i.e. schools were not present) or agents were choosing to work over investing in human capital. Forgoing investment in human capital would be rational (in the short-term) if they were poorer on average. But before the onset of British rule these groups on

either side of the boundary were similar in wealth. Hence, it is likely that differences in school infrastructure across the boundary lowered the literacy rates for advantaged castes.

But can differences in schools explain the cross-caste differences in persistence of education from 1931 to 2016? Differences in provision of schools increased the literacy rates on the British side of the boundary and this difference can persist if the difference in public goods continues to persist. But then this persistence should be observed for advantaged castes too unless the distribution of schools was skewed to the extent that it only catered to areas where advantaged castes lived. The geographic targeting of schools towards advantaged castes is unlikely as Indian villages generally have multiple castes. Hence, factors other than school provision must vary across the boundary to explain the persistence in educational differences for disadvantaged castes and the lack of persistence for advantaged castes. This may include higher levels of discrimination against disadvantaged caste groups on the indirect rule side.

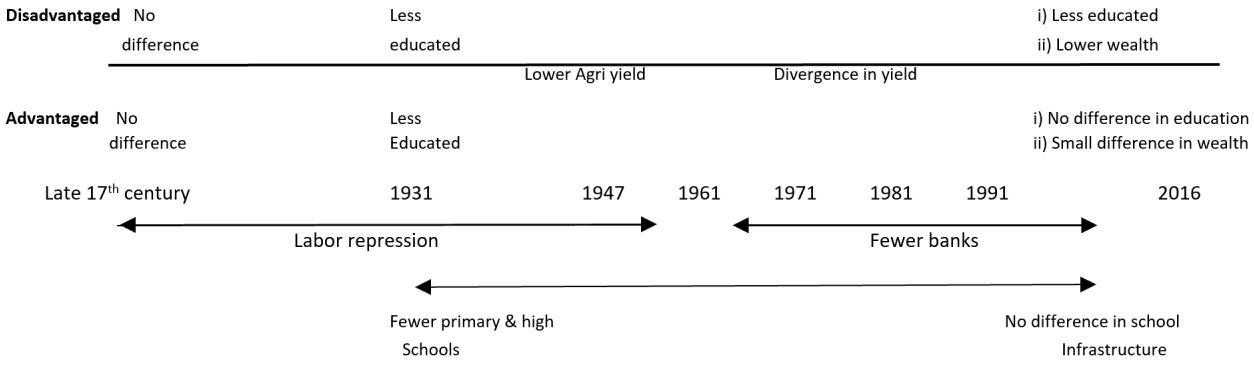
**Labor-repressive institutions** The administration in *Hyderabad* was financially weak and this meant that law and order was not maintained (Bhanumathi (1992)). The state was particularly weak outside the capital city of Hyderabad, allowing large landlords to become de facto providers of work, credit, and justice. In the absence of state administration, advantaged caste landlords became dominant in areas away from the capital, to the detriment of disadvantaged caste peasants and laborers. Most of the large landlords, who also doubled up as moneylenders, were from the upper castes. Advantaged castes moneylenders were creditors in 93% of the total loans given on land security in *Hyderabad* (Iyengar (1931)).[24] They forced a large portion of the rural class into agricultural labor. Most of these individuals belonged to the disadvantaged castes.

While similarly oppressive labor systems existed in areas under the British, the comparatively lower dominance of landlords in those areas due to (mostly) direct tax collection by the administration, meant that repression was low compared to Hyderabad.

[24]This figure is based on the author's calculations for the districts of Aurangabad, Warangal, Nanded and Raichur, using data in Iyengar (1931). The 7% moneylenders from disadvantaged caste belong to one caste group, the Lingayats. Lingayats have different sub groups within them some of whom are advantaged and others are disadvantaged, but Iyengar (1931) doesn't differentiate between the two. Hence, we assume that a fixed percent of Lingayat moneylenders were disadvantaged caste based on their fraction in modern day Lingayats. Our estimate is very likely an upper bound as it may have been the case that all the Lingayat moneylender were advantaged castes.

**Credit institutions** Theoretical work by Galor and Zeira (1993) has shown that credit market imperfections (combined with non-convex returns to investment in human capital) can lead to persistence in wealth inequality because of an inability of agents to access credit. Credit markets are intertwined with labor-repression as their absence may lead to debt bondage. Moneylenders (who are mostly landlords) lend money to cultivators for farming. When they are unable to pay back the money they take over the land and force them to work on the land for gratis till the debt is paid. This causes higher land inequality (which we find) and lower wealth for the disadvantaged caste. The process can persist even after abolition of *Zamindari* as individuals with generational wealth (largely the advantaged castes) can continue to lend to the poor. As the disadvantaged caste groups in *Hyderabad* were poorer than similar groups on the other side at the time of independence they would be more reliant on landlords for loans. Access to credit can break this cycle by reducing reliance on moneylenders. If there is difference in access to credit across the boundary then the initial difference in wealth can persist. As poorer individuals have to work to consume, they have to forego investment in human capital. Hence, there will be persistence in educational outcomes for the group which are more reliant on loans, i.e. the disadvantaged caste. Thus, credit institutions can play an important role in the persistence of outcomes.

Figure 7: Timeline of outcomes and underlying mechanisms




Notes: The figure shows the timeline of results and the mechanisms. All comparisons are between indirectly ruled *Hyderabad* and directly ruled British areas.

### 4.2 Empirical evidence on mechanisms

**Labor institutions** To test for differences in labor-repressive agriculture, we use data on presence of bonded labor in villages from the Socio Economic and Caste Census of 2011.[25] Surveys on bonded labor (other than SECC) suggest that around 89% of bonded laborers worked in agriculture and 93.2% of them belong to disadvantaged groups (Srivastava (2005), Gandhi Peace (1979)).[26] It must be noted that the SECC data captures the number of individuals released from bonded labor between 1976 (when bonded labor was prohibited) and 2011, and not the actual number of bonded laborers. For any difference in this measure to be indicative of differences in labor-repression during colonial times, the number of bonded laborers should be directly proportional to the number of bonded laborers released. Table 6 shows that the fraction of individuals in a village who were released from bonded labor is twice as high in indirect rule areas compared to direct rule areas. As our estimates compare areas close to the boundary which lie within the same modern day state, differences in state capacity cannot explain our results. These results from SECC data are supported by a survey carried out by Gandhi Peace foundation and National Labor Institute in 1979 (Gandhi Peace (1979)) (see Online Appendix J-1).

The second strategy that we adopt is to look at variation in labor-repression within the state of *Hyderabad*. As we have discussed above labor-repression was highest in *Jagirdari* areas and areas where the *deshmukhs* were dominant. Online Appendix Figure J-2 shows the location of non-dewani areas in the state of *Hyderabad*. This includes *sarf-i-khas* (land used for Nizam's private purse), *paigah*, *jagir* and *samastan*. The latter three have similar administration and were highly exploitative. We categorize the latter three as labor-repressive areas and the remaining areas (dewani and *sarf-i-khas*) as comparatively less labor-repressive. We estimate the effect of indirect rule using our main specification separately for these two areas for both the disadvantaged and advantaged caste groups. If our hypothesis of labor-repression were true, then the estimate for disadvantaged castes in labor-repressive areas would be larger than the estimate for this caste group in the re-

[25] The SECC asked every household whether they had a member who had been legally released from bonded labor. The convention on the Abolition of Slavery in 1956 defined bonded labour as: *Debt bondage, that is to say, the status or condition arising from a pledge by a debtor of his personal services or of those of a person under his control as security for a debt if the value of those services as reasonably assessed is not applied towards the liquidation of the debt or the length and nature of those services are not respectively limited and defined*

[26] The National Commission on Rural Labour (Labour (1991)) found that bondage of agricultural labourers was the main form of bondage in India.

Table 6: Fraction of villagers released from bonded labour

| | *Fraction released from bonded labour* | | | |
|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| *Hyderabad* | 0.0031** | 0.0036** | 0.0021* | 0.0016 |
| | (0.0012) | (0.0016) | (0.0011) | (0.0016) |
| Observations | 26,794 | 19,177 | 26,794 | 12,496 |
| Clusters | 606 | 441 | 606 | 294 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 29.97 |
| Mean Dep. Var. | 0.003 | 0.003 | 0.003 | 0.003 |
| SD Dep. Var. | 0.032 | 0.037 | 0.032 | 0.036 |

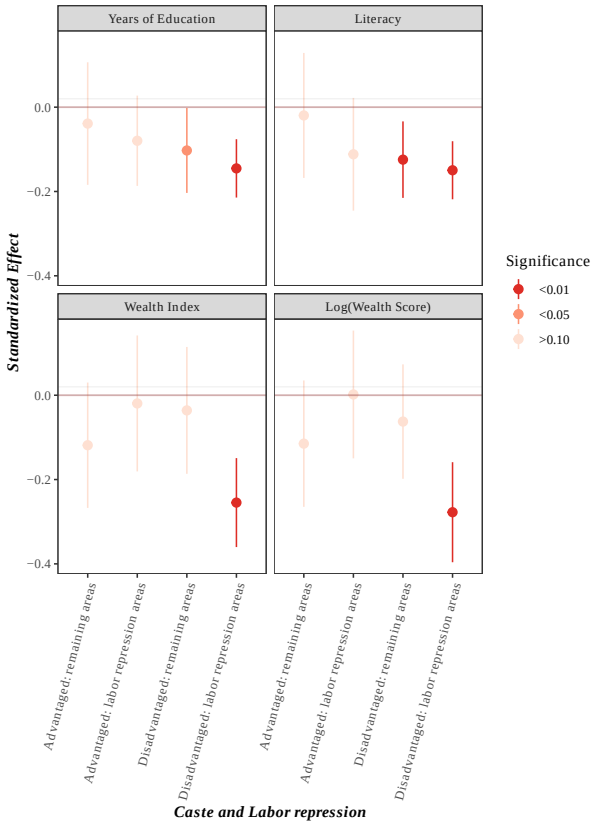
The data is taken from Socio Economic and Caste census of India 2011 (Ministry of Rural Development (2011)) (source Asher et al. (2021)). Standard errors are clustered at the subdistrict level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. RD regressions include a local linear specification estimated separately on each side of the boundary and uses a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications include fixed effects for boundary segment and control for total population. Robust standard errors are in parenthesis. * p < 0.10, ** p <0.05, *** p<0.01.

maining areas. In other words, disadvantaged caste groups in labor-repressive areas would be doing worse than similar caste groups across the boundary by a larger margin than disadvantaged caste groups in remaining areas (which had less labor-repression) when they are compared to similar groups across the boundary. Figure 8 shows that this is the case. Disadvantaged groups in both labor-repressive and remaining areas have significantly less education than similar groups across the boundary, but the effect is stronger for labor-repressive areas. For wealth, disadvantaged castes in remaining areas have an estimate close to zero but these groups in labor-repressive areas have statistically significant and strong negative estimate.

Additionally, in Appendix Table 9, we compare the areas where labor-repression was high to the remaining areas within *Hyderabad* using our local linear RD specification. We find that while the overall estimate of being in a labor-repression area is negative but not significant, it is strongly negative and significant for disadvantaged caste groups, while it flips signs for advantaged castes for education. This provides further support for the theory that labor-repression in these areas was directed against disadvantaged castes.

Now we turn to evidence on agricultural wages. In a standard economic model, (real) wages

Figure 8: Educational and wealth outcomes across labor-repressive areas

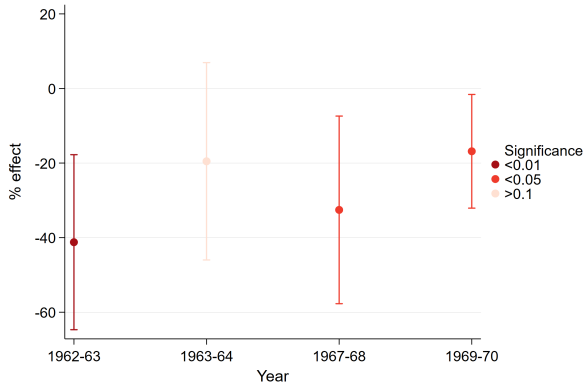



Notes: The figure presents standardized RD coefficients for educational and wealth outcomes for different areas by caste groups. We divide the sample inside the boundary into two groups, labor-repression areas — which include *jagirs*, *paigah* and *samastans* — and remaining areas which makes up rest of the sample. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes. We use boundary segment fixed effects for all specifications. Optimal bandwidth is calculated separately on each side of the boundary using the procedure suggested by Cattaneo et al. (2020). Figure shows point estimates and 95% confidence intervals.

in an area capture the marginal product of labor. Therefore, low wages can be indicative of low productivity in a region. They may also indicate the presence of coercive institutions on average, as these tend to suppress wages (Acemoglu and Wolitzky (2011)). Figure 9 shows the estimates of wages of agricultural labor across the boundary of *Hyderabad* in the 1960s using our main specification.[27] Agricultural wages in *Hyderabad* were around 30% lower than in nearby areas that were under the British. As the majority of agricultural laborers belonged to disadvantaged castes, this points to an important mechanism in explaining the previously documented differences in outcomes.

[27]We use data from the Agricultural Wages of India. The dataset has wages for two to three centers within each district in our sample. However, starting from the 1970s, the dataset only includes one center per district, which limits our ability to extend our analysis to that time period.

Figure 9: Estimates of wage across the boundary of Hyderabad




Notes: The figure shows RD coefficients of a regression of agricultural wages across the boundary of Hyderabad. The data on agricultural wages is taken from Agricultural Wages of India. We take mean of wages for men and women. We restrict the sample to monsoon months of June, July and August to ensure that we are not capturing differences driven by agricultural season. Regressions control for price index, total rainfall in the month and mean temperature in the district. We use month and boundary segment fixed effects for all specifications. Optimal bandwidth is calculated using the procedure suggested by Cattaneo et al. (2020). Figure shows point estimates in terms of percentage and 95% confidence intervals.

**Differences in credit institutions** To provide support for the channel on credit institutions we estimate the difference in banks after independence across the boundary using our main specification. The results in Table 7 suggest that there is a persistent difference in the presence of banks starting from 1960 till 2005 and this difference is narrowing over time. The estimate for 1961 is at the block level and rest of the estimates are at the grid cell level.[28] In Appendix Table 8 we look at the difference in presence of commercial and cooperative banks at the village level in 2001 and find that while there is no difference in presence of commercial banks, villages in former indirect rule areas are 30% less likely to have a cooperative bank.[29]

Having shown the difference in banks across the boundary, we provide two pieces of evidence to support the hypothesis that this policy of the government exacerbated differences in outcomes. Firstly, we look at how the crop yields changes between *Hyderabad* and non-*Hyderabad* areas after the bank policy was implemented in 1970. Areas which had more banks would provide more

[28]For 1961 we use data from the district census. For rest of the years we use Asher et al. (2021). Data from 1961 includes both scheduled and co-operative banks while rest of the data only includes Scheduled banks.

[29]This raises the question about the reasons for the differences across the boundary. The policy of expanding rural credit and encouraging banks to open branches in rural areas without banking services began in 1969. The rules for opening the bank indicated that the likelihood of getting a bank increased in unbanked population (Burgess and Pande (2005)). But the choice of location from the list of unbanked locations that satisfied this criterion was left to the bank. Therefore conditioning on existing banks per capita, the probability of being chosen as a location for a bank might have increased in the financial viability and business potential of the location. In rural areas that loosely translates into agricultural productivity.

credit to farmers which would allow them to invest in productive methods and improve their yields. Online Appendix table E-4 provides suggestive evidence in favor of this hypothesis. Difference in difference estimates for crop yield are negative in indirect areas for many crops after the bank reform. Our estimates are imprecise and that is probably due to small sample size. This piece of evidence ties up well with Burgess and Pande (2005) which showed that the rural bank expansion program led to decrease in rural poverty in India. Secondly, we look at indebtedness of farmers to non-institutional sources in 2003. In the absence of banks, farmers are more reliant on moneylenders and local traders. Appendix table 10 shows that non-institutional debt, the average debt from moneylenders and traders, is much higher in *Hyderabad* and this is driven by disadvantaged caste groups. These groups rely more on moneylenders in the post-colonial era whereas advantaged caste groups do not (their coefficients are small and not significant) due to generational wealth. Thus, access to cheaper credit from banks can partially explain the better outcomes for disadvantaged castes in British areas.[30]

**Public infrastructure** In Online Appendix Table E-6 we compare the number of education facilities. British areas had higher number of primary and high schools just after independence in 1961. However, these differences are not observed in contemporary data suggesting that areas in *Hyderabad* have caught up. Online Appendix Table E-7 looks at differences in presence of other public goods in contemporary data. Results indicate that villages in indirectly ruled areas are less likely to have a paved road but more likely to have hospitals.

## 5 Conclusion

This paper looks at the long-term distributional effects of indirect vs direct rule across the boundary of the largest indirectly ruled state in India. We find that disadvantaged caste groups in indirectly ruled areas have lower wealth and educational outcomes compared to similar groups in geographically close directly ruled areas. The historically advantaged upper caste groups under

[30]Data limitations on geographic granularity of observations means that we have to use the district centroids of observations for this analysis. This is far from ideal, but we hope that these results will be seen in light of what we have shown in the previous section. To illustrate the difference between using granular and non-granular geographic observations, we run the same specification for an outcome variable(years of education) for which we have granular and non-granular geographic data. This is shown in Online Appendix E.7. The estimates are very close in size and significant for both levels of geographic granularity.

Table 7: Difference in banks

| | *Banks in 1961* | | | | *Banks in 1980* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -2.214*** | -2.481*** | -2.974*** | -4.468*** | -1.285*** | -1.216*** | -1.202*** | -1.444*** |
| | (0.322) | (0.315) | (0.621) | (1.299) | (0.265) | (0.303) | (0.361) | (0.549) |
| Observations | 157 | 113 | 157 | 84 | 821 | 597 | 821 | 432 |
| Clusters | 26 | 21 | 26 | 21 | 247 | 185 | 247 | 139 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.65 | 75.00 | 50.00 | 75.00 | 32.56 |
| Mean Dep. Var. | 2.796 | 2.796 | 2.796 | 2.571 | 1.845 | 1.642 | 1.845 | 1.509 |
| SD Dep. Var. | 3.038 | 3.132 | 3.038 | 3.007 | 4.243 | 4.355 | 4.243 | 4.298 |
| Scheduled Bank | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| cooperatives | Yes | Yes | Yes | Yes | No | No | No | No |
| Sample | No Karnataka | No Karnataka | No Karnataka | No Karnataka | Complete | Complete | Complete | Complete |

| | *Banks in 1990* | | | | *Banks in 2005* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -1.869*** | -1.782*** | -1.724*** | -1.974** | -1.927*** | -1.598** | -1.914** | -3.079** |
| | (0.364) | (0.417) | (0.540) | (0.840) | (0.663) | (0.765) | (0.877) | (1.324) |
| Observations | 821 | 597 | 821 | 443 | 821 | 597 | 821 | 416 |
| Clusters | 247 | 185 | 247 | 140 | 247 | 185 | 247 | 137 |
| Bandwidth | 75.00 | 75.00 | 75.00 | 33.66 | 75.00 | 75.00 | 75.00 | 31.11 |
| Mean Dep. Var. | 2.697 | 2.430 | 2.697 | 2.257 | 5.820 | 5.275 | 5.820 | 4.603 |
| SD Dep. Var. | 5.747 | 5.940 | 5.747 | 5.787 | 9.920 | 10.300 | 9.920 | 6.600 |
| Scheduled Bank | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Cooperatives | No | No | No | No | No | No | No | No |
| Sample | Complete | Complete | Complete | Complete | Complete | Complete | Complete | Complete |

Observations are at the block level for 1961, and at the grid level for the remaining years. Standard errors clustered at the district level for the former and block level for the latter. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for boundary segment fixed effects. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

indirect rule have comparable educational outcomes to similar groups in directly ruled areas. The effects on wealth for these groups is negative but not robust to other specifications. The size of effect is also smaller than for disadvantaged groups. Thus, the burden of persistent differences is borne by historically disadvantaged caste groups. Indirectly ruled areas also have higher wealth and land inequality.

Historical work suggests that the difference in intensity of labor-repressive agriculture which mostly affected the disadvantaged caste groups maybe driving these results. We find evidence which supports this hypothesis. Indirect rule areas are more likely to have bonded labor than directly ruled areas. The differences in education and productivity that we find in contemporary data had emerged by the end of the colonial rule itself. Differences in provision of schools across the boundary support these results, but these infrastructure differences vanish over time.

To account for the persistence of these differences in the post-colonial era we look at credit institutions. We show that former direct rule areas have more banks compared to indirect rule areas and this led to a divergence in agricultural productivity during the independence era. The difference in credit institutions also caused disadvantaged caste farmers to take up more informal debt in indirect rule areas. This difference in credit institutions explains relatively less investment in education and less wealth acquisition by disadvantaged caste groups in indirect rule areas. Our results provide support to theoretical work by Galor and Zeira (1993) by showing that initial differences in wealth distribution (which was the case at the time of independence) under imperfect capital markets can lead to persistence in inequality (which we find in contemporary data).

These results should be understood from a deontological perspective. Both the British and the *Nizam* of Hyderabad had exploitative administrative systems of land revenue aimed at optimizing their objective function. The British were aided by revenue from other parts of the country, giving them leeway in policy formulation. This income was not available to the *Nizam*. To pay off debts the *Nizam* had to lease lands to military chieftains which led to the establishment of repressive labor institutions in *Jagirdari* areas. An attempt to increase land revenue led to an increase in concentration of lands and thereby power of the landlords who used that to force swathes of rural masses into labor in *Dewani* areas. It must be kept in mind that a comparison between direct and indirect rule is not a comparison between colonial rule and native rule but rather a comparison between colonial rule and the interaction of colonial & native rule. To get at the former comparison, one must create a counterfactual where there was no colonial rule.

The results in this paper have important policy implications. When countries become independent from colonial rule, they are likely to focus on growth and development and not on the heterogeneity in outcomes that run across the social hierarchy. If the social group at the bottom of the hierarchy is worse off compared to others in areas which had repressive institutions, then post-independence institutions that are built on top of colonial institutions can exacerbate the differences across locations for disadvantaged groups as groups at the top of the hierarchy are better placed to utilize location-agnostic government policies everywhere while groups at the bottom might be more lacking in formerly repressive areas. The inequality that arises, as a result, can have negative consequences of its own (Banerjee and Newman (1993), Galor and Zeira (1993), Mookherjee and Ray (2010)) and can sustain an equilibrium in which the advantaged groups con-

tinue to extract rent (through debt bondage or usurious lending) from the disadvantaged groups. These results can help in accounting for the contemporary differences in outcomes between various caste groups that arise at different points in time. This has important consequences for the caste-fragmented Indian society.

# Appendix

## A.1 Village level credit

Table 8: Village level banks in 2001

| | *Commercial bank in the village* | | | | *Cooperative bank in the village* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.002 | 0.002 | 0.001 | -0.001 | -0.019*** | -0.018*** | -0.015*** | -0.017** |
| | (0.004) | (0.004) | (0.005) | (0.008) | (0.005) | (0.006) | (0.006) | (0.007) |
| Observations | 27,470 | 19,754 | 27,470 | 9,310 | 27,470 | 19,754 | 27,470 | 12,585 |
| Clusters | 612 | 447 | 612 | 240 | 612 | 447 | 612 | 294 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 20.17 | 75.00 | 50.00 | 75.00 | 28.83 |
| Mean Dep. Var. | 0.087 | 0.083 | 0.087 | 0.084 | 0.073 | 0.070 | 0.073 | 0.070 |
| SD Dep. Var. | 0.282 | 0.276 | 0.282 | 0.278 | 0.260 | 0.256 | 0.260 | 0.255 |

The data is taken from Asher et al. (2021). We restrict the sample to villages. Standard errors are clustered at the subdistrict level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. RD regressions include a local linear specification estimated separately on each side of the boundary and uses a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications include fixed effects for boundary segment and control for total population. Robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## A.2 Comparison of labor-repressive areas with remaining areas

Table 9: Comparison of labor-repressive areas with rest of the areas

| | *Years of Education* | | | *Literacy* | | |
|---|---|---|---|---|---|---|
| | Overall | Disadvantaged | Advantaged | Overall | Disadvantaged | Advantaged |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Repression* | -0.536 | -1.188** | 1.195 | -0.048 | -0.080** | -0.000 |
| | (0.603) | (0.534) | (0.956) | (0.037) | (0.037) | (0.067) |
| Observations | 5,421 | 4,455 | 1,046 | 5,478 | 4,301 | 1,183 |
| Clusters | 43 | 53 | 19 | 46 | 53 | 27 |
| Bandwidth | 6.04 | 7.83 | 5.00 | 6.26 | 7.79 | 5.84 |
| Mean Dep. Var. | 6.428 | 6.143 | 7.502 | 0.694 | 0.660 | 0.768 |
| SD Dep. Var. | 4.914 | 5.078 | 4.690 | 0.461 | 0.474 | 0.423 |

| | *Wealth Index* | | | *Log(wealth score)* | | |
|---|---|---|---|---|---|---|
| | Overall | Disadvantaged | Advantaged | Overall | Disadvantaged | Advantaged |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Repression* | -0.216 | -0.258 | -0.180 | -0.086 | -0.121* | -0.078 |
| | (0.176) | (0.217) | (0.263) | (0.056) | (0.066) | (0.091) |
| Observations | 2,679 | 1,766 | 600 | 2,729 | 1,879 | 628 |
| Clusters | 27 | 26 | 16 | 27 | 27 | 17 |
| Bandwidth | 4.53 | 4.42 | 4.41 | 4.65 | 4.68 | 4.67 |
| Mean Dep. Var. | 2.793 | 2.639 | 3.142 | 12.198 | 12.155 | 12.300 |
| SD Dep. Var. | 1.160 | 1.119 | 1.214 | 0.392 | 0.386 | 0.397 |

We compare labor-repressive areas with the remaining areas in *Hyderabad*. Standard errors clustered at the NFHS cluster level. Regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. We show results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We use fixed effects for each *jagir*. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### A.3 Effect of Banks

Table 10: Agricultural debt

| | *Non institutional debt* | | |
|---|---|---|---|
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 5795.219*** | 7763.970*** | 3028.326 |
| | (2061.382) | (2329.964) | (4895.768) |
| Observations | 1,767 | 1,767 | 1,767 |
| $R^2$ | 0.05 | 0.06 | 0.02 |
| clusters | 347 | 347 | 347 |
| mean | 16784.154 | 16784.154 | 16784.154 |
| | *Non institutional debt - Disadvantaged castes* | | |
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 6777.398*** | 10004.730*** | 3998.802* |
| | (2307.956) | (2921.220) | (2301.629) |
| Observations | 1,381 | 1,381 | 1,381 |
| $R^2$ | 0.06 | 0.07 | 0.06 |
| clusters | 306 | 306 | 306 |
| mean | 13379.862 | 13379.862 | 13379.862 |
| | *Non institutional debt - Advantaged castes* | | |
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 1908.452 | 1179.376 | 284.127 |
| | (4785.960) | (4606.103) | (9030.898) |
| Observations | 386 | 386 | 386 |
| $R^2$ | 0.13 | 0.15 | 0.13 |
| clusters | 155 | 155 | 155 |
| mean | 18978.56 | 18978.56 | 18978.56 |

*Notes* : The dependent variable is farmer debt from non-institutional sources. All specifications control for mean altitude of district and household size. The last two specifications control for a geographic polynomial in the latitude and longitude of the district centroid and the first one does not. We include fixed effects for boundary segment and use sample weights provided in the NSS59. Errors are clustered at the nss cluster level. Robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## References


**Acemoglu, Daron and Alexander Wolitzky**, "The Economics of Labor Coercion," *Econometrica*, 2011, *79* (2), 555–600. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA8963.

_ **, Simon Johnson, and James A. Robinson**, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, December 2001, *91* (5), 1369–1401.

**Adamopoulos, Tasso and Diego Restuccia**, "Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data," *The Review of Economic Studies*, July 2022, *89* (4), 1629–1653.

**Ambedkar, BR**, "Castes in India, their mechanism, genesis and development," *Lecture at Columbia University*, 1916.

**Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad**, "Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform," 2021. World Bank Economic Review.

**Bandiera, Antonella, Lelys I. Dinarte Diaz, Juan Miguel Jimenez, Sandra V. Rozo, and Maria Micaela Sviatschi**, "Rebel Governance and Development: The Persistent Effects of Guerrillas in El Salvador," September 2022.

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

**Banerjee, Abhijit V. and Andrew F. Newman**, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, *101* (2), 274–298. Publisher: University of Chicago Press.

**Barreca, Alan I., Melanie Guldi, Jason M. Lindo, and Glen R. Waddell**, "Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification," *The Quarterly Journal of Economics*, 2011, *126* (4), 2117–2123. Publisher: Oxford University Press.

**Becker, Sascha O. and Luigi Pascali**, "Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years," *American Economic Review*, May 2019, *109* (5), 1764–1804.

**Benjamin, Thomas D. Fish David Eitelberg Mayala and Trinadh Dontamsetti**, "The DHS Program Geospatial Covariate Datasets Manual (Second Edition). Rockville, Maryland, USA: ICF," 2018.

**Bhanumathi, S**, *Land Revenue administration in the Nizam's dominions* 1992.

**Burgess, Robin and Rohini Pande**, "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment," *American Economic Review*, May 2005, *95* (3), 780–795.

**Cattaneo, Matias D, Nicolas Idrobo, and Rocıo Titiunik**, "A Practical Introduction to Regression Discontinuity Designs: Volume II," 2020.

**Chaudhary, Latika**, "Taxation and educational development: Evidence from British India," *Explorations in Economic History*, July 2010, *47* (3), 279–293.

_ **and Anand V. Swamy**, "Protecting the borrower: An experiment in colonial India," *Explorations in Economic History*, July 2017, *65*, 36–54.

_ **and Manuj Garg**, "Does history matter? Colonial education investments in India," *The Economic History Review*, 2015, *68* (3), 937–961. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ehr.12078.

_ **, Jared Rubin, Sriya Iyer, and Anand Shrivastava**, "Culture and colonial legacy: Evidence from public goods games," *Journal of Economic Behavior & Organization*, May 2020, *173*, 107–129.

**Chi, Guanghua, Han Fang, Sourav Chatterjee, and Joshua E Blumenstock**, "Microestimates of wealth for all low-and middle-income countries," *Proceedings of the National Academy of Sciences*, 2022, *119* (3), e2113658119.

**Colleoni, Marco**, "The long-term welfare effects of colonial institutions: Evidence from Central India," *Journal of Development Economics*, October 2023, p. 103170.

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

**Cramer, Kim Fe**, "Bank Presence and Health," *SSRN Electronic Journal*, 2021.

**Dell, Melissa**, "The Persistent Effects of Peru's Mining *Mita*," *Econometrica*, 2010, *78* (6), 1863–1903.

_ **and Benjamin A Olken**, "The Development Effects Of The Extractive Colonial Economy: The Dutch Cultivation System In Java," 2020, p. 51.

**Deshpande, Ashwini**, "Caste at Birth? Redefining Disparity in India," *Review of Development Economics*, 2001, *5* (1), 130–144. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9361.00112.

**Dhanagare, D.N.**, "Social origins of the peasant insurrection in Telangana (1946-51)," *Contributions to Indian Sociology*, January 1974, *8* (1), 109–134. Publisher: SAGE Publications India.

**GAEZ**, "FAO and IIASA. Global Agro Ecological Zones version 4 (GAEZ v4).," 2021.

**Galor, Oded and Joseph Zeira**, "Income Distribution and Macroeconomics," *The Review of Economic Studies*, January 1993, *60* (1), 35.

**Gelman, Andrew and Guido Imbens**, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, July 2019, *37* (3), 447–456.

**Grasse, Donald**, "State Terror and Long-Run Development: The Persistence of the Khmer Rouge," *American Political Science Review*, May 2023, pp. 1–18. Publisher: Cambridge University Press.

**Gupta, Akhil**, "Revolution in Telengana 1946–1951 (Part One)," *Comparative Studies of South Asia, Africa and the Middle East*, May 1984, *4* (1), 1–26.

**Henderson, J. Vernon, Adam Storeygard, and David N. Weil**, "A Bright Idea for Mesuring Economic Growth," *American Economic Review*, 2011.

**Heston, Alan W.**, "Official Yields Per Acre in India, 1886-1947; Some Questions of Interpretation," *The Indian Economic & Social History Review*, October 1973, *10* (4), 303–332. Publisher: SAGE Publications India.

**Hunter, Cotton James Sutherland Burn Richard Meyer William Great Britain. India Office. William Wilson**, *Imperial gazetteer of India ...*, Oxford: Clarendon Press, 1907.

**Iyengar, Kesava**, *Economic investigations in the Hyderabad State 1929-30*, Hyderabad: The Government Central Press, 1931.

**Iyer, Lakshmi**, "DIRECT VERSUS INDIRECT COLONIAL RULE IN INDIA: LONG-TERM CONSEQUENCES," *THE REVIEW OF ECONOMICS AND STATISTICS*, 2010, p. 21.

**Jain, Tarun**, "Common Tongue: The Impact of Language on Educational Outcomes," *The Journal of Economic History*, 2017, *77* (2), 473–510. Publisher: [Economic History Association, Cambridge University Press].

**Kelly, Morgan**, "Persistence, Randomization, and Spatial Noise," 2021.

**Labour, National Commission Rural**, "Report of the National Commission on Rural Labour, Vol. 2 Study Group on Bonded Labour," 1991.

**Lee, Alexander**, "Redistributive Colonialism: The Long Term Legacy of International Conflict in India," *Politics & Society*, June 2017, *45* (2), 173–224.

**Loury, Glenn C.**, "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 1981, *49* (4), 843–867. Publisher: [Wiley, Econometric Society].

**Lowes, Sara and Eduardo Montero**, "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State," October 2020, p. w27893.

**Mamdani, Mahmood**, *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*, Princeton University Press, 1996.

**Mendez, Esteban and Diana Van Patten**, "Multinationals, Monopsony, and Local Development: Evidence From the United Fruit Company," *Econometrica*, 2022, *90* (6), 2685–2721. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19514.

**Mookherjee, Dilip and Debraj Ray**, "Inequality and Markets: Some Implications of Occupational Diversity," *American Economic Journal: Microeconomics*, November 2010, *2* (4), 38–76.

**Moore, Barrington**, *Social origins of dictatorship and democracy*, Penguin, 1984.

**Mukherjee, Shivaji**, *Colonial Institutions and Civil War: Indirect Rule and Maoist Insurgency in India* Cambridge Studies in Contentious Politics, Cambridge University Press, 2021.

**Nayakwadi, Y.H.**, "Agrarian Relations in Hyderabad-Karnataka (1800-1948)," *Proceedings of the Indian History Congress*, 2006, *67*, 446–456. Publisher: Indian History Congress.

**Nunn, Nathan**, "THE LONG-TERM EFFECTS OF AFRICA'S SLAVE TRADES," *QUARTERLY JOURNAL OF ECONOMICS*, 2008, p. 38.

**of India Ministry of Rural Development, Government**, "Socio Economic and Caste Census," 2011.

**Pavier, Barry**, *The Telengana Movement 1944-51*, New Delhi: Vikas Publishing House, 1981.

**Peace, Foundation & National Labour Institute Gandhi**, *National survey on the incidence of bonded labour : preliminary report : an action research project of Gandhi Peace Foundation, National Labour Institute.*, New Delhi: The Institute, 1979.

**Rahim, Muhammad Abdur.**, "Lord Dalhousie's administration of the conquered and annexed states." PhD dissertation, Chand, Delhi 1963.

**Raju, Sarada**, "The Agrarian Structure of the Madras Presidency: Impact of British Administration," *Proceedings of the Indian History Congress*, 1974, *35*, 265–272. Publisher: Indian History Congress.

**Ratnoo, Vigyan D.**, "Persistent effects of colonial land tenure institutions: Village-level evidence from India," *Journal of Development Economics*, May 2024, *168*, 103247.

**Sharma, Ram Sharan**, *Sudras in Ancient India*, Motilal Banrassidas, 1958.

**Sherman, Taylor C.**, "Migration, Citizenship and Belonging in Hyderabad (Deccan), 1946–1956," *Modern Asian Studies*, January 2011, *45* (1), 81–107.

**Srivastava, Ravi S**, "Bonded Labour in India: Its Incidence and Pattern," 2005.

**Sundarayya, Puchalapalli**, *Telangana people's struggle and its lessons*, Communist Party of India (Marxist) [Calcutta], 1972.

**Thirumali, I.**, *Against Dora and Nizam: People's Movement in Telangana, 1939-1948*, Kanishka Publishers, Distributors, 2003.

**Thorat, Sukhadeo and Katherine S. Neuman**, *Blocked by Caste: Economic Discrimination in Modern India*, Oxford University Press, 2012.

**Voth, Hans-Joachim**, "Persistence: Myth and Mystery," 2020.

# Part I
# Online Appendix

## Table of Contents


**A Data Sources** **48**
- A.1 Geographic Data and Variables . . . 48
- A.2 DHS Survey Data and Variables . . . 48
- A.3 Banks Data . . . 48
- A.4 Public Infrastructure data . . . 49
- A.5 Actual Yield Data . . . 49
- A.6 Debt data . . . 49
- A.7 Data on agricultural wages . . . 49

**B Robustness checks for overall outcomes** **50**
- B.1 Overall outcomes with a donut hole . . . 50
- B.2 Conley Standard Errors . . . 51
- B.3 Overall outcomes without control variables . . . 52
- B.4 Results after excluding Berar . . . 53
- B.5 Results without river districts . . . 54
- B.6 Results with additional controls . . . 54
- B.7 Results from SHRUG wealth and Nightlights data . . . 55
- B.8 RD Plots for overall outcomes . . . 56
- B.9 Spatial Maps for overall outcomes . . . 56
- B.10 Robustness to RD Parameters . . . 58
- B.11 RD Plots for quadratic polynomial . . . 62

**C Robustness checks for social groups** **63**
- C.1 Estimates by caste groups with weights . . . 63
- C.2 Estimates by caste groups with a donut hole . . . 64
- C.3 Difference in difference by caste type . . . 66
- C.4 Results for caste groups without Berar . . . 67
- C.5 RD Plots for quadratic polynomial by caste groups . . . 68
- C.6 Simulations from spatial noise . . . 69
- C.7 Results with conley standard errors by caste . . . 70
- C.8 Robustness checks for different segments by caste groups . . . 72
- C.9 Robustness checks for different bandwidths by caste groups . . . 73
- C.10 Robustness checks for different parameters by caste groups . . . 74

**D Measures of Inequality** **74**
- D.1 Land inequality . . . 74
- D.2 Wealth inequality . . . 78

**E Intermediate Outcomes and Mechanisms** **81**
- E.1 Occupation in 1911 . . . 81
- E.2 Caste groups in 1931 . . . 82
- E.3 RD Plots for Banks . . . 83
- E.4 Yields in 1956-61 . . . 84
- E.5 Effect of Banks . . . 85

E.6 Public Infrastructure . . . 86
E.7 Women's education estimated for two different geographic granularity . . . 87

**F RD Plots of balance variables** **89**

**G Other princely states** **91**
G.1 Mysore . . . 91

**H Selective migration** **92**
H.1 Migration in 1911 . . . 92
H.2 Trimming for selective migration . . . 93
H.3 Migration to other parts of the Hyderabad . . . 94
H.4 Estimates for disadvantaged castes by gender . . . 95

**I A comment on using IVs in persistence regressions** **96**

**J Maps** **98**

# A Data Sources

## A.1 Geographic Data and Variables

- **Boundaries of *Hyderabad*:-** From the Imperial Gazetteer of India (Hunter (1907)
- **Geographic variables:** The DHS covariates manual (Mayala and Dontamsetti (2018)) captures information about the precipitation, average temperature, aridity, altitude, slope, episodes of droughts, vegetation at the DHS cluster level.
  - Rainfall: This is the average of precipitation across 2000, 2005, and 2010.
  - Temperature: We use the mean of average temperature across three years - 2000, 2005, and 2010.
  - Aridity: Aridity is the ratio of precipitation to transpiration. A higher level of aridity is indicative of higher moisture content. We use mean of average aridity across three years - 2000, 2005, and 2010.
  - Episodes of Drought - The average number of drought episodes (categorized between 1 (low) and 10 (high)) for the areas within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster location.
  - vegetation - The average vegetation index value within the 2 km (urban) or 10 km (rural) buffer surrounding the DHS survey cluster.
- **Soil Characteristics:** The soil rasters were obtained from the Global AgroEcological Zones version 4.0 (GAEZ v 4.0) project and are available at https://gaez.fao.org/.
- **Agro-Climatic Potential yield:** Agro-climatic yield rasters were obtained from the Global AgroEcological Zones version 4.0 (GAEZ v 4.0) project and are available at https://gaez.fao.org/. We used the 30-year average starting in 1961 of the crops with a significant share of the overall produce between 1956-2000. The data has a spatial resolution of 5 arc-minutes.

## A.2 DHS Survey Data and Variables

Survey data on development outcomes for individuals is provided by the 2015-16 NFHS-IV survey (also known as DHS). The survey data and detailed information on the sampling procedure and variable definitions is available at http://dhsprogram.com/data/ Data-Variables-and-Definitions.cfm. We describe the variables used below:

- **Years of Education:** The NFHS survey asks the individual the total number of years of education in single years.
- **Literacy:** Literacy is a 0 to 1 indicator variable for each individual where 0 is "cannot read at all", and 1 is "able to read only parts of a sentence" or "able to read a whole sentence".
- **Wealth Factor Score:** Wealth Factor Score is an index generated by the NFHS using principle component analysis on asset ownership for each individual. For the log of the wealth factor score, we define it as Log(Wealth Score) = Log(Wealth Factor Score + min(Wealth Factor Score) + 1).
- **Wealth Index:** Wealth Index is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Factor Score.

## A.3 Banks Data

The data on banks from 1980 onwards is taken from Asher et al. (2021). Data on banks in 1961 is compiled from the district census of India of 1961 separately for all districts within 200 km of the boundary. This data is publicly available.

### A.4 Public Infrastructure data

Data on presence of schools, and police stations in 1961 is compiled from the district census of India of 1961 separately for all districts within 200 km of the boundary. This data is publicly available. For the data on education infrastructure in 2016, we use public data from Ministry of rural development, India available at https://geosadak-pmgsy.nic.in/OpenData. This data captures all public educational facilities in the country.

### A.5 Actual Yield Data

Data on crops grown and yields is taken from the India Agriculture and Climate Data Set assembled by the World Bank.

### A.6 Debt data

We measure indebtedness of the farmers from the *Situation Assessment survey of Farmers* conducted as part of the $59^{th}$ round of the National Sample Survey in 2003. The survey captures the various sources of indebtedness of farmers across the country. We categorize a loan as non-institutional if it comes from either a moneylender or a trader.

### A.7 Data on agricultural wages

We use wage data from Agricultural Wages in India, an yearly survey conducted by Directorate of Economics and Statistics of India. The earliest survey was conducted in 1950-51. The survey covers most agricultural districts in India. In each district, the survey reports wages for men, women, and child in different agricultural (as well as non-agricultural) occupations, for at least one center. As districts are fairly large, data from just one center in the district would not give enough variation for an RD analysis. We use data from four years in the 1960s which have more than one center in most districts in our sample. We map the geographic location of each center mentioned in the dataset and use that for the RD analysis.

# B Robustness checks for overall outcomes

## B.1 Overall outcomes with a donut hole

Table B-1: Overall outcomes

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -1.008*** | -0.718*** | -0.948*** | -1.634*** | -0.073*** | -0.054*** | -0.081*** | -0.128*** |
| | (0.142) | (0.168) | (0.255) | (0.341) | (0.011) | (0.012) | (0.021) | (0.028) |
| Observations | 29,313 | 20,218 | 29,313 | 12,489 | 28,923 | 19,929 | 28,923 | 13,462 |
| Clusters | 1,101 | 760 | 1,101 | 465 | 1,101 | 760 | 1,101 | 509 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 29.66 | 75.00 | 50.00 | 75.00 | 32.41 |
| Mean Dep. Var. | 6.836 | 6.624 | 6.836 | 6.550 | 0.706 | 0.689 | 0.706 | 0.682 |
| SD Dep. Var. | 5.013 | 5.057 | 5.013 | 5.081 | 0.456 | 0.463 | 0.456 | 0.466 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.261*** | -0.187*** | -0.369*** | -0.756*** | -0.077*** | -0.055*** | -0.106*** | -0.223*** |
| | (0.049) | (0.058) | (0.083) | (0.123) | (0.016) | (0.018) | (0.026) | (0.038) |
| Observations | 18,212 | 12,621 | 18,217 | 7,039 | 18,212 | 12,621 | 18,217 | 7,323 |
| Clusters | 1,101 | 760 | 1,101 | 421 | 1,101 | 760 | 1,101 | 437 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 26.39 | 75.00 | 50.00 | 75.00 | 27.60 |
| Mean Dep. Var. | 3.129 | 3.081 | 3.129 | 3.053 | 12.304 | 12.288 | 12.304 | 12.279 |
| SD Dep. Var. | 1.194 | 1.195 | 1.194 | 1.189 | 0.387 | 0.392 | 0.387 | 0.388 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. We use a donut RD and remove observations within 10 km of the former *Hyderabad* boundary . Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

## B.2 Conley Standard Errors

Table B-2: Results with conley standard errors

| | *Years of Education* | | *Literacy* | |
|---|---|---|---|---|
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.786*** | -0.837** | -0.059** | -0.069** |
| | (0.266) | (0.338) | (0.024) | (0.029) |
| Observations | 30,095 | 15,866 | 29,639 | 18,343 |
| Bandwidth | 75.00 | 37.05 | 75.00 | 43.75 |
| Mean Dep. Var. | 6.668 | 6.439 | 0.696 | 0.680 |
| SD Dep. Var. | 5.027 | 5.047 | 0.460 | 0.467 |
| | *Wealth Index* | | *Log(wealth score)* | |
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.251*** | -0.271** | -0.074** | -0.084** |
| | (0.094) | (0.123) | (0.030) | (0.039) |
| Observations | 18,648 | 8,383 | 18,648 | 9,135 |
| Bandwidth | 75.00 | 31.24 | 75.00 | 34.09 |
| Mean Dep. Var. | 3.073 | 3.006 | 12.287 | 12.266 |
| SD Dep. Var. | 1.188 | 1.169 | 0.389 | 0.386 |

Standard errors are Conley (1999) spatial standard errors assuming a cut-off window of 50 kms. All regressions include a local linear specification estimated separately on each side of the concession boundary. Column 1 for each outcome uses a fixed bandwidth of 75 km and column 2 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

## B.3 Overall outcomes without control variables

Table B-3: Overall outcomes without control variables

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.745*** | -0.641*** | -1.157*** | -1.186*** | -0.054*** | -0.056*** | -0.106*** | -0.120*** |
| | (0.139) | (0.158) | (0.260) | (0.398) | (0.011) | (0.012) | (0.022) | (0.032) |
| Observations | 30,095 | 21,144 | 30,095 | 13,664 | 29,639 | 20,808 | 29,639 | 15,272 |
| Clusters | 1,125 | 786 | 1,125 | 505 | 1,125 | 786 | 1,125 | 576 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 31.69 | 75.00 | 50.00 | 75.00 | 36.11 |
| Mean Dep. Var. | 6.668 | 6.477 | 6.668 | 6.494 | 0.696 | 0.683 | 0.696 | 0.681 |
| SD Dep. Var. | 5.027 | 5.028 | 5.027 | 5.035 | 0.460 | 0.465 | 0.460 | 0.466 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.204*** | -0.206*** | -0.418*** | -0.370*** | -0.059*** | -0.063*** | -0.130*** | -0.136*** |
| | (0.047) | (0.052) | (0.082) | (0.127) | (0.015) | (0.017) | (0.027) | (0.040) |
| Observations | 18,648 | 13,113 | 18,648 | 8,237 | 18,648 | 13,113 | 18,648 | 8,940 |
| Clusters | 1,125 | 786 | 1,125 | 492 | 1,125 | 786 | 1,125 | 536 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 30.22 | 75.00 | 50.00 | 75.00 | 33.46 |
| Mean Dep. Var. | 3.073 | 3.012 | 3.073 | 3.001 | 12.287 | 12.268 | 12.287 | 12.266 |
| SD Dep. Var. | 1.188 | 1.174 | 1.188 | 1.166 | 0.389 | 0.389 | 0.389 | 0.385 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B.4 Results after excluding Berar

Table B-4: Overall outcomes without Berar

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.607*** | -0.531*** | -1.252*** | -1.306*** | -0.053*** | -0.058*** | -0.120*** | -0.133*** |
| | (0.177) | (0.203) | (0.318) | (0.482) | (0.014) | (0.017) | (0.027) | (0.041) |
| Observations | 23,234 | 16,519 | 23,234 | 10,996 | 22,838 | 16,225 | 22,838 | 11,835 |
| Clusters | 875 | 616 | 875 | 409 | 875 | 616 | 875 | 448 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.40 | 75.00 | 50.00 | 75.00 | 35.36 |
| Mean Dep. Var. | 6.559 | 6.441 | 6.559 | 6.442 | 0.679 | 0.670 | 0.679 | 0.668 |
| SD Dep. Var. | 5.138 | 5.150 | 5.138 | 5.176 | 0.467 | 0.470 | 0.467 | 0.471 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.244*** | -0.235*** | -0.409*** | -0.206 | -0.074*** | -0.073*** | -0.123*** | -0.094** |
| | (0.058) | (0.063) | (0.094) | (0.153) | (0.018) | (0.020) | (0.030) | (0.047) |
| Observations | 14,465 | 10,261 | 14,468 | 5,940 | 14,465 | 10,261 | 14,468 | 6,496 |
| Clusters | 875 | 616 | 875 | 355 | 875 | 616 | 875 | 389 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 27.63 | 75.00 | 50.00 | 75.00 | 30.82 |
| Mean Dep. Var. | 3.114 | 3.074 | 3.114 | 3.037 | 12.302 | 12.290 | 12.302 | 12.286 |
| SD Dep. Var. | 1.169 | 1.161 | 1.169 | 1.151 | 0.377 | 0.379 | 0.377 | 0.374 |

We do our analysis after removing districts in the erstwhile state of Berar (outside *Hyderabad*) as well as bordering districts in *Hyderabad*. Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender , household size and boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B.5 Results without river districts

Table B-5: Overall outcomes without river districts

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.793*** | -0.661*** | -1.023*** | -0.836* | -0.055*** | -0.052*** | -0.095*** | -0.114*** |
| | (0.166) | (0.183) | (0.291) | (0.483) | (0.013) | (0.014) | (0.024) | (0.039) |
| Observations | 24,365 | 17,413 | 24,365 | 10,284 | 23,979 | 17,133 | 23,979 | 11,513 |
| Clusters | 887 | 632 | 887 | 371 | 887 | 632 | 887 | 424 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 26.56 | 75.00 | 50.00 | 75.00 | 31.85 |
| Mean Dep. Var. | 6.797 | 6.558 | 6.797 | 6.481 | 0.711 | 0.694 | 0.711 | 0.699 |
| SD Dep. Var. | 4.938 | 4.944 | 4.938 | 4.924 | 0.453 | 0.461 | 0.453 | 0.459 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.151*** | -0.175*** | -0.395*** | -0.100 | -0.041** | -0.051*** | -0.120*** | -0.043 |
| | (0.055) | (0.060) | (0.088) | (0.169) | (0.018) | (0.019) | (0.029) | (0.052) |
| Observations | 14,807 | 10,589 | 14,810 | 4,632 | 14,807 | 10,589 | 14,810 | 5,191 |
| Clusters | 887 | 632 | 887 | 277 | 887 | 632 | 887 | 310 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 18.65 | 75.00 | 50.00 | 75.00 | 20.95 |
| Mean Dep. Var. | 2.991 | 2.922 | 2.991 | 2.899 | 12.259 | 12.238 | 12.259 | 12.237 |
| SD Dep. Var. | 1.200 | 1.177 | 1.200 | 1.127 | 0.397 | 0.394 | 0.397 | 0.377 |

We do our analysis after removing districts that are not along the river. Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B.6 Results with additional controls

Table B-6: Overall outcomes with additional controls

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.930*** | -0.920*** | -1.186*** | -1.215*** | -0.070*** | -0.074*** | -0.108*** | -0.122*** |
| | (0.162) | (0.174) | (0.261) | (0.390) | (0.012) | (0.014) | (0.022) | (0.033) |
| Observations | 30,095 | 21,144 | 30,095 | 14,425 | 29,639 | 20,808 | 29,639 | 14,683 |
| Clusters | 1,125 | 786 | 1,125 | 534 | 1,125 | 786 | 1,125 | 554 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 33.42 | 75.00 | 50.00 | 75.00 | 34.55 |
| Mean Dep. Var. | 6.668 | 6.477 | 6.668 | 6.484 | 0.696 | 0.683 | 0.696 | 0.684 |
| SD Dep. Var. | 5.027 | 5.028 | 5.027 | 5.039 | 0.460 | 0.465 | 0.460 | 0.465 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.238*** | -0.266*** | -0.433*** | -0.224* | -0.071*** | -0.082*** | -0.135*** | -0.111*** |
| | (0.053) | (0.056) | (0.081) | (0.134) | (0.017) | (0.018) | (0.027) | (0.042) |
| Observations | 18,645 | 13,109 | 18,648 | 7,304 | 18,645 | 13,109 | 18,648 | 8,237 |
| Clusters | 1,125 | 786 | 1,125 | 436 | 1,125 | 786 | 1,125 | 492 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 26.05 | 75.00 | 50.00 | 75.00 | 30.23 |
| Mean Dep. Var. | 3.073 | 3.012 | 3.073 | 2.973 | 12.287 | 12.268 | 12.287 | 12.265 |
| SD Dep. Var. | 1.188 | 1.174 | 1.188 | 1.153 | 0.389 | 0.389 | 0.389 | 0.383 |

We do our analysis with additional controls for cluster altitude, rainfall and temperature. Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender, household size, cluster altitude, rainfall , temperature and boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

## B.7 Results from SHRUG wealth and Nightlights data

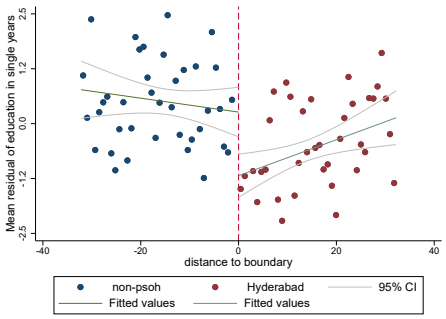
Table B-7: Wealth and nightlights

| | *Facebook wealth measure* | | | | *Nightlights from VIIRS* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.033*** | -0.022** | -0.016* | -0.023* | -0.060** | -0.053* | -0.038* | -0.000 |
| | (0.009) | (0.010) | (0.009) | (0.013) | (0.027) | (0.029) | (0.021) | (0.019) |
| Observations | 27,725 | 19,926 | 27,725 | 9,163 | 27,799 | 19,980 | 27,799 | 8,873 |
| Clusters | 614 | 448 | 614 | 238 | 614 | 448 | 614 | 229 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 19.67 | 75.00 | 50.00 | 75.00 | 18.93 |
| Mean Dep. Var. | -0.184 | -0.201 | -0.184 | -0.215 | 0.447 | 0.421 | 0.447 | 0.363 |
| SD Dep. Var. | 0.247 | 0.238 | 0.247 | 0.229 | 0.736 | 0.749 | 0.736 | 0.534 |

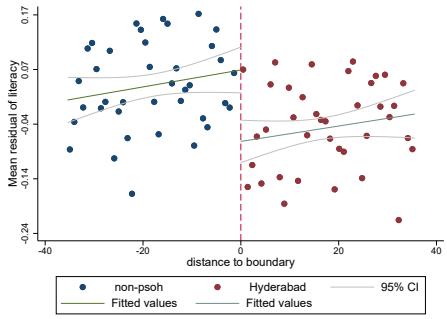
We do our analysis using the wealth and nightlights data from Asher et al. (2021). Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. RD regressions include a local linear specification estimated separately on each side of the boundary and uses a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for boundary segment fixed effects. All regressions include controls for total population. We calculate the nightlights variable by taking the annual mean for the years 2012-2021. * p < 0.10, ** p <0.05, *** p<0.01.

## B.8 RD Plots for overall outcomes

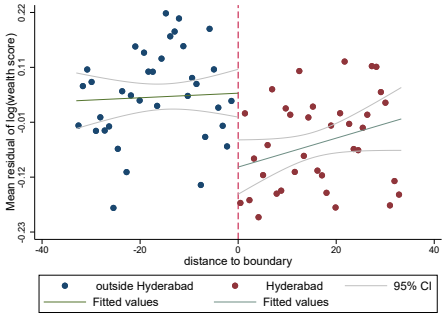
Figure B-1: RD plots of outcomes



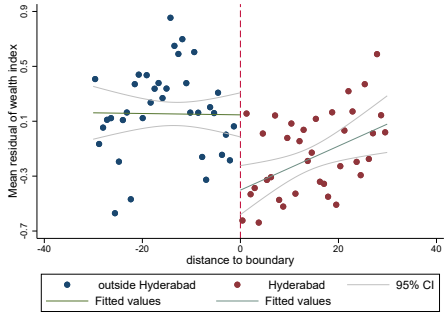

(a) Years of education




(b) Literacy




(c) log(wealth)

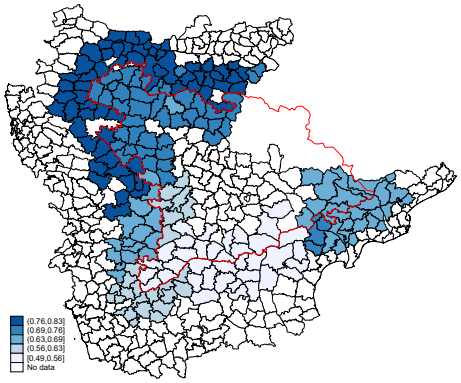



(d) Wealth Index

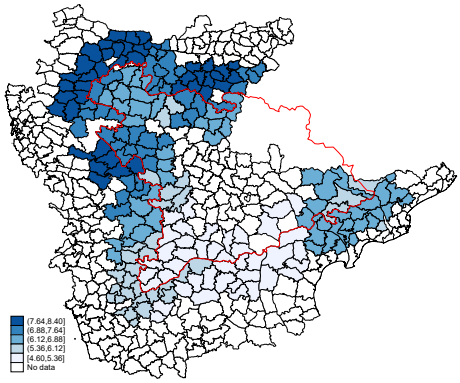
Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 1 km bin along the running variable as well as with a local linear trend estimated separately on each side of the discontinuity. Each regression is estimated using the optimal bandwidth chosen through the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

## B.9 Spatial Maps for overall outcomes

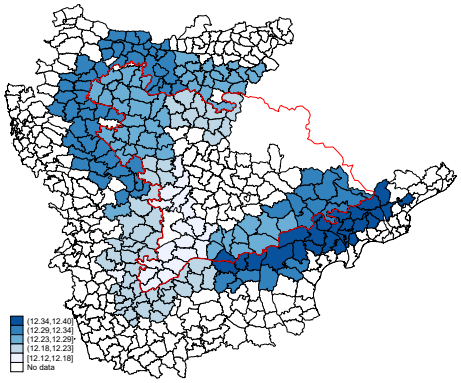
Figure B-2: RD Maps of outcomes



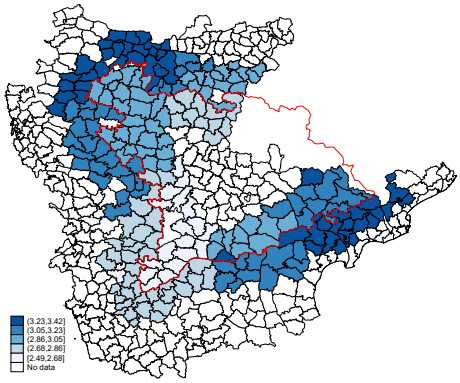

(a) Literacy




(b) years of education




(c) log of wealth score

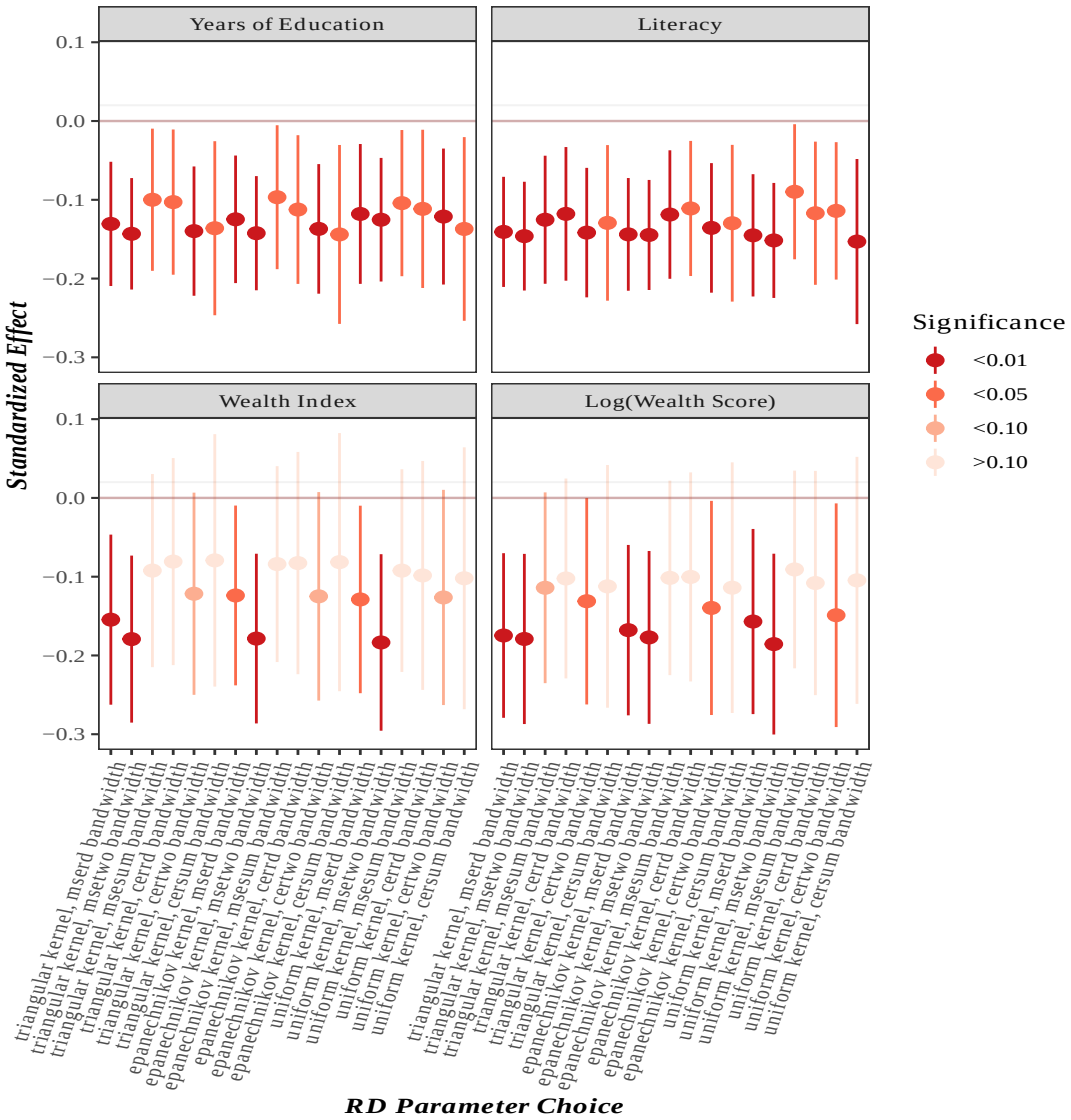



(d) Wealth Index

Notes: The figure presents RD maps for our main outcomes. We estimate a regression of mean outcomes at block level on linear polynomial in coordinates with boundary segment fixed effects within 75 km of the boundary. The plot colour represents predicted outcomes of the regression. Standard errors are clustered at the district level.

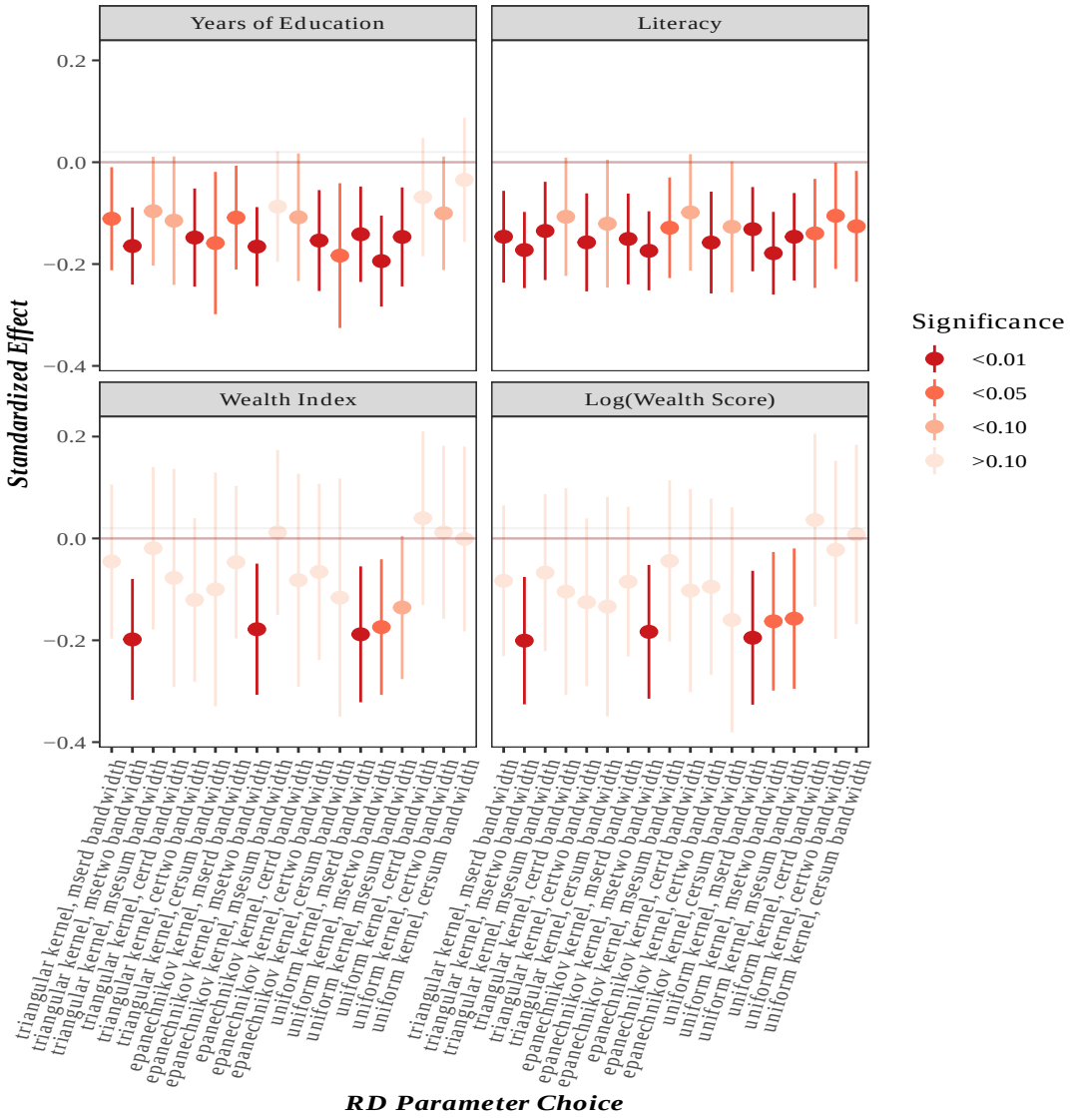
## B.10 Robustness to RD Parameters

Figure B-3: Robustness to RD Parameters



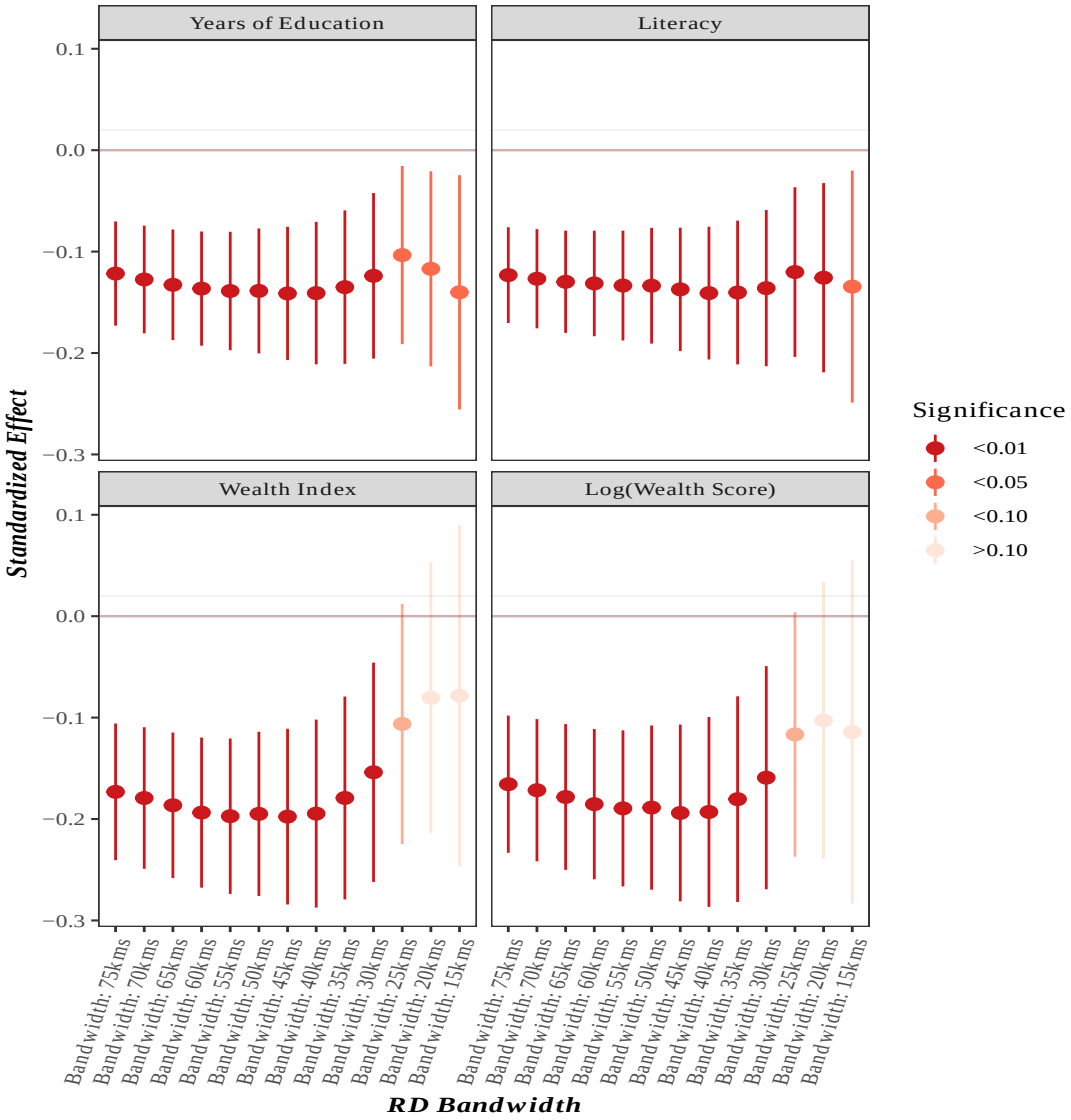

Notes: The figure presents RD coefficients for our main outcomes for different values of the RD Parameters. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. Bandwidth Type represents the optimal bandwidth selection procedure used for each regression: mserd choses one common MSE-optimal bandwidth; msetwo choses two different MSE-optimal bandwidths (below and above the cutoff); msesum choses one common MSE-optimal bandwidth selector for the sum of regression estimates (instead of the difference); cerrd choses one common CER-optimal bandwidth; certwo two different CER-optimal bandwidths (below and above the cutoff); cersum choses one common CER-optimal bandwidth for the sum of regression estimates. Figure shows point estimates and 95% confidence intervals.

Figure B-4: Robustness to RD Parameters : Quadratic polynomial



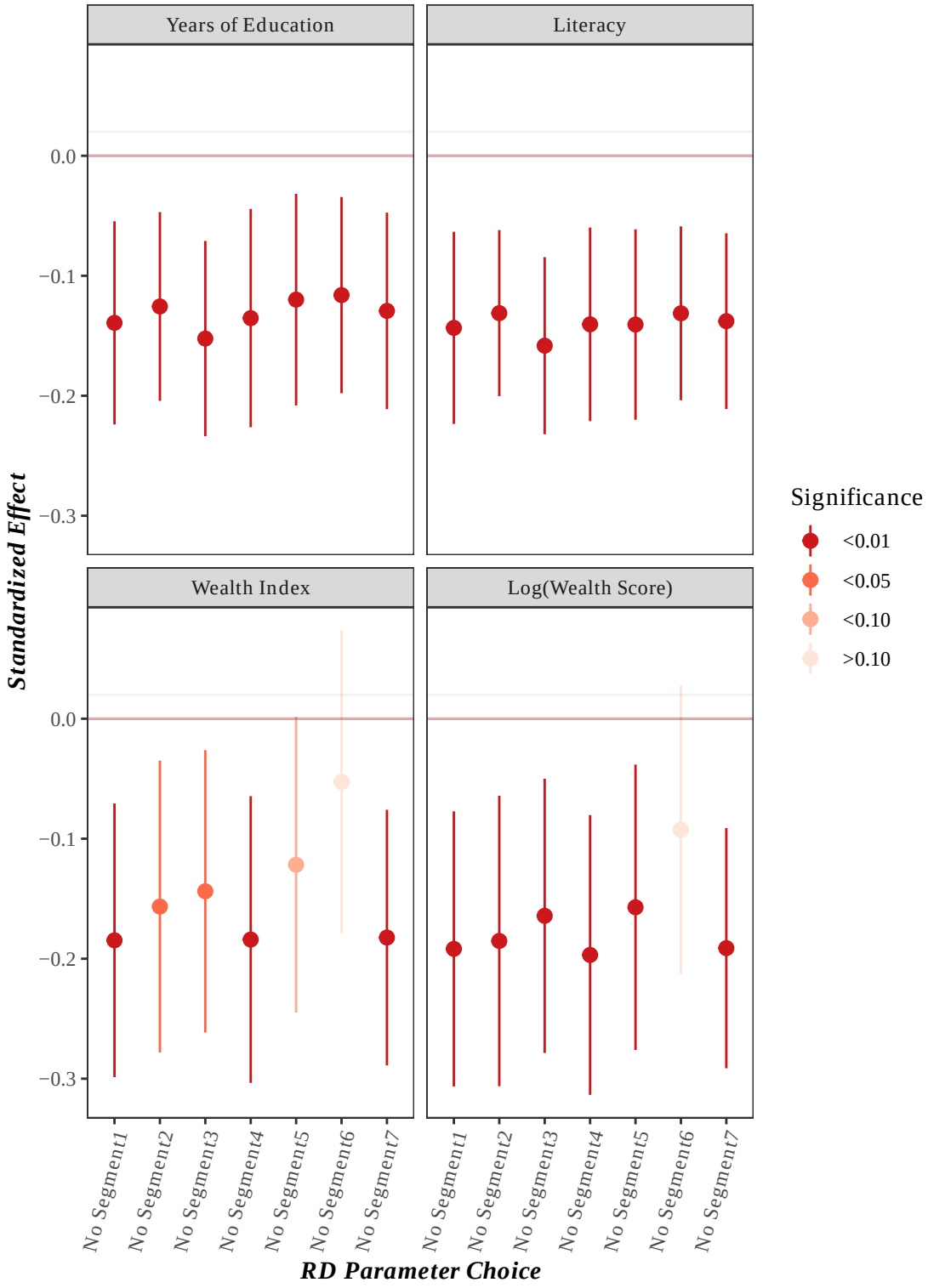

Notes: The figure presents RD coefficients for our main outcomes for different values of the RD Parameters for a local quadratic polynomial. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. Bandwidth Type represents the optimal bandwidth selection procedure used for each regression: mserd choses one common MSE-optimal bandwidth; msetwo choses two different MSE-optimal bandwidths (below and above the cutoff); msesum choses one common MSE-optimal bandwidth selector for the sum of regression estimates (instead of the difference); cerrd choses one common CER-optimal bandwidth; certwo two different CER-optimal bandwidths (below and above the cutoff); cersum choses one common CER-optimal bandwidth for the sum of regression estimates. Figure shows point estimates and 95% confidence intervals.

Figure B-5: Robustness to Distance to boundary




Notes: The figure presents RD coefficients for our main outcomes for different bandwidths. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. The figure shows result for a triangular kernel, and optimal bandwith chosen using procedure suggested by Cattaneo et al. (2020). Figure shows point estimates with 95% confidence intervals.

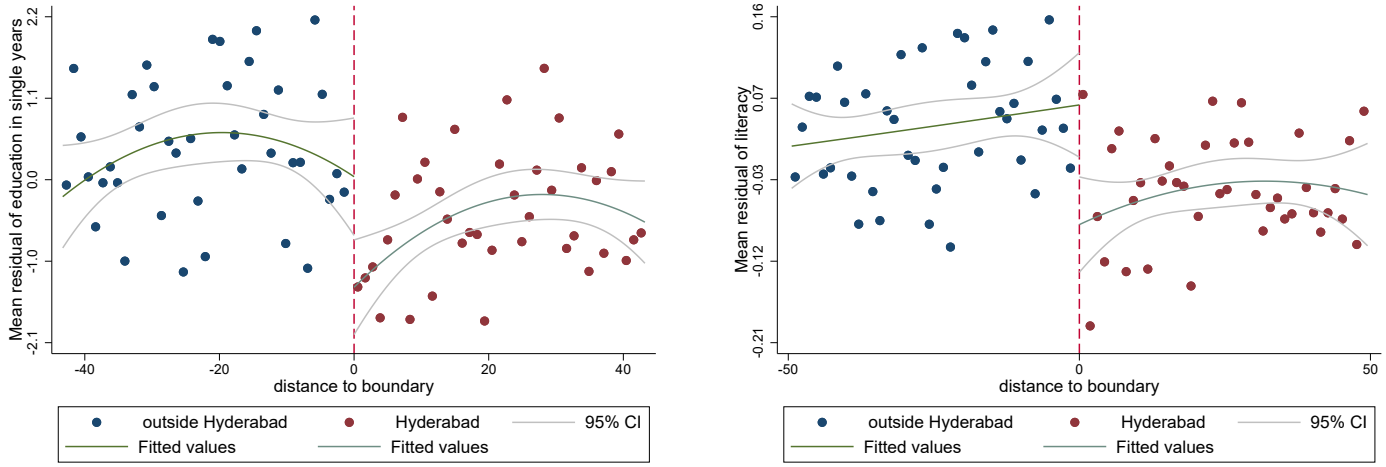
Figure B-6: Robustness to segments




Notes: The figure presents RD coefficients for our main outcomes after excluding different segments. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. The figure shows result for a triangular kernel, and optimal bandwith chosen using procedure suggested by Cattaneo et al. (2020). Figure shows point estimates with 95% confidence intervals.

## B.11 RD Plots for quadratic polynomial

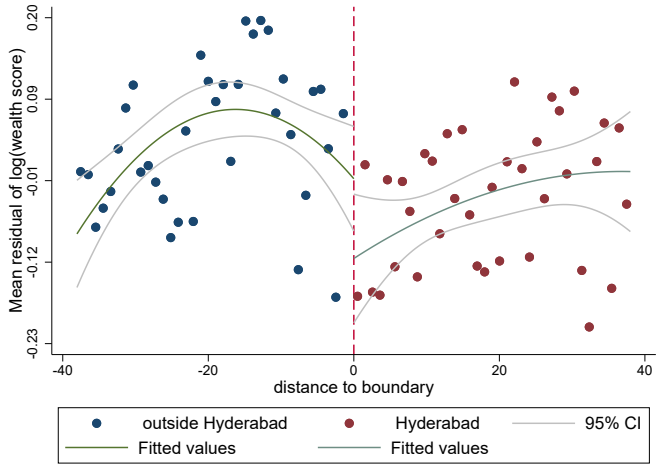
Figure B-7: RD plots of outcomes



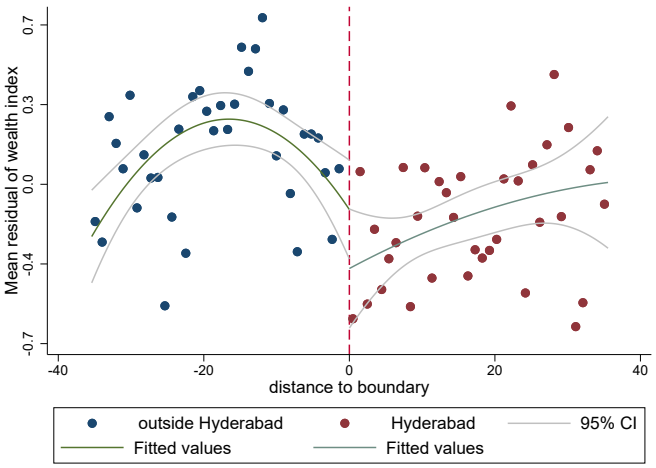

(a) Years of education




(b) Literacy




(c) log(wealth)




(d) Wealth Index

Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 1 km bin along the running variable with a local quadratic trend estimated separately on each side of the discontinuity. Each regression is estimated using the optimal bandwidth chosen through the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

# C Robustness checks for social groups

## C.1 Estimates by caste groups with weights

Table C-1: Estimates by caste groups

| | *Disadvantaged Castes* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.753*** | -0.625*** | -1.134*** | -1.291*** | -0.064*** | -0.063*** | -0.106*** | -0.120*** |
| | (0.164) | (0.193) | (0.307) | (0.471) | (0.013) | (0.015) | (0.027) | (0.037) |
| Observations | 21,245 | 15,042 | 21,245 | 10,053 | 20,920 | 14,783 | 20,920 | 11,770 |
| Clusters | 1,113 | 781 | 1,113 | 522 | 1,113 | 781 | 1,113 | 618 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.59 | 75.00 | 50.00 | 75.00 | 39.57 |
| Mean Dep. Var. | 6.334 | 6.108 | 6.334 | 6.051 | 0.665 | 0.651 | 0.665 | 0.647 |
| SD Dep. Var. | 5.088 | 5.066 | 5.088 | 5.043 | 0.472 | 0.477 | 0.472 | 0.478 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.187*** | -0.160** | -0.369*** | -0.322** | -0.058*** | -0.054*** | -0.119*** | -0.120** |
| | (0.056) | (0.063) | (0.103) | (0.161) | (0.018) | (0.021) | (0.035) | (0.054) |
| Observations | 13,220 | 9,337 | 13,223 | 5,726 | 13,220 | 9,337 | 13,223 | 5,897 |
| Clusters | 1,113 | 781 | 1,113 | 476 | 1,113 | 781 | 1,113 | 493 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 29.40 | 75.00 | 50.00 | 75.00 | 30.82 |
| Mean Dep. Var. | 2.980 | 2.921 | 2.980 | 2.908 | 12.259 | 12.240 | 12.259 | 12.237 |
| SD Dep. Var. | 1.168 | 1.152 | 1.168 | 1.140 | 0.387 | 0.386 | 0.387 | 0.380 |
| | *Advantaged Castes* | | | | | | | |
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | 0.288 | 0.277 | -0.126 | -0.361 | 0.024 | 0.017 | 0.001 | 0.005 |
| | (0.246) | (0.293) | (0.434) | (0.569) | (0.016) | (0.020) | (0.036) | (0.051) |
| Observations | 6,052 | 4,215 | 6,052 | 2,872 | 6,000 | 4,188 | 6,000 | 3,318 |
| Clusters | 767 | 530 | 767 | 362 | 764 | 530 | 764 | 425 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.59 | 75.00 | 50.00 | 75.00 | 39.57 |
| Mean Dep. Var. | 7.765 | 7.650 | 7.765 | 7.799 | 0.781 | 0.771 | 0.781 | 0.773 |
| SD Dep. Var. | 4.853 | 4.871 | 4.853 | 4.935 | 0.413 | 0.420 | 0.413 | 0.419 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | 0.066 | 0.055 | -0.113 | 0.058 | 0.024 | 0.023 | -0.029 | 0.007 |
| | (0.079) | (0.089) | (0.132) | (0.170) | (0.024) | (0.027) | (0.040) | (0.051) |
| Observations | 3,808 | 2,669 | 3,811 | 1,629 | 3,808 | 2,669 | 3,811 | 1,700 |
| Clusters | 767 | 530 | 767 | 332 | 767 | 530 | 767 | 344 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 29.40 | 75.00 | 50.00 | 75.00 | 30.82 |
| Mean Dep. Var. | 3.336 | 3.276 | 3.336 | 3.286 | 12.364 | 12.347 | 12.364 | 12.356 |
| SD Dep. Var. | 1.219 | 1.216 | 1.219 | 1.208 | 0.390 | 0.396 | 0.390 | 0.390 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). We use survey weights provided in the DHS. Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p < 0.10, ** p <0.05, *** p<0.01.

## C.2 Estimates by caste groups with a donut hole

Table C-2: Estimates by caste groups

*Disadvantaged Castes*

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -1.097*** | -0.860*** | -1.078*** | -1.822*** | -0.088*** | -0.074*** | -0.109*** | -0.145*** |
| | (0.156) | (0.180) | (0.271) | (0.347) | (0.012) | (0.015) | (0.024) | (0.028) |
| Observations | 20,536 | 14,072 | 20,536 | 9,613 | 20,267 | 13,863 | 20,267 | 13,473 |
| Clusters | 1,089 | 752 | 1,089 | 513 | 1,089 | 752 | 1,089 | 730 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.84 | 75.00 | 50.00 | 75.00 | 47.67 |
| Mean Dep. Var. | 6.534 | 6.238 | 6.534 | 6.117 | 0.675 | 0.655 | 0.675 | 0.656 |
| SD Dep. Var. | 5.089 | 5.100 | 5.089 | 5.075 | 0.468 | 0.476 | 0.468 | 0.475 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.291*** | -0.219*** | -0.351*** | -0.612*** | -0.091*** | -0.069*** | -0.106*** | -0.186*** |
| | (0.053) | (0.061) | (0.087) | (0.106) | (0.017) | (0.020) | (0.028) | (0.034) |
| Observations | 12,877 | 8,851 | 12,882 | 7,160 | 12,877 | 8,851 | 12,882 | 7,270 |
| Clusters | 1,089 | 752 | 1,089 | 611 | 1,089 | 752 | 1,089 | 619 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 38.97 | 75.00 | 50.00 | 75.00 | 39.42 |
| Mean Dep. Var. | 3.042 | 2.984 | 3.042 | 2.952 | 12.277 | 12.259 | 12.277 | 12.250 |
| SD Dep. Var. | 1.179 | 1.169 | 1.179 | 1.159 | 0.387 | 0.388 | 0.387 | 0.389 |

*Advantaged Castes*

| | *Years of Education* | | | | *Literacy* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.261 | 0.075 | -0.078 | -0.752 | -0.006 | 0.018 | 0.003 | -0.013 |
| | (0.260) | (0.296) | (0.461) | (0.593) | (0.017) | (0.019) | (0.031) | (0.036) |
| Observations | 5,972 | 4,295 | 5,972 | 3,001 | 5,923 | 4,264 | 5,923 | 4,153 |
| Clusters | 757 | 522 | 757 | 370 | 753 | 522 | 753 | 506 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 32.84 | 75.00 | 50.00 | 75.00 | 47.67 |
| Mean Dep. Var. | 7.873 | 7.792 | 7.873 | 7.765 | 0.789 | 0.780 | 0.789 | 0.782 |
| SD Dep. Var. | 4.831 | 4.895 | 4.831 | 4.946 | 0.408 | 0.414 | 0.408 | 0.413 |

| | *Wealth Index* | | | | *Log(wealth score)* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.111 | -0.032 | -0.280** | -0.534*** | -0.025 | 0.003 | -0.073* | -0.157*** |
| | (0.084) | (0.096) | (0.138) | (0.164) | (0.027) | (0.031) | (0.044) | (0.051) |
| Observations | 3,736 | 2,705 | 3,738 | 2,202 | 3,736 | 2,705 | 3,738 | 2,225 |
| Clusters | 757 | 522 | 757 | 426 | 757 | 522 | 757 | 431 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 38.97 | 75.00 | 50.00 | 75.00 | 39.42 |
| Mean Dep. Var. | 3.380 | 3.354 | 3.380 | 3.313 | 12.378 | 12.369 | 12.378 | 12.361 |
| SD Dep. Var. | 1.222 | 1.236 | 1.222 | 1.227 | 0.388 | 0.397 | 0.388 | 0.396 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. We use a donut RD and remove observations within 10 km of the former *Hyderabad* boundary. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## C.3 Difference in difference by caste type

Table C-3: Difference in difference Results

| | *Years of Education* | | *Literacy* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Hyderabad | -0.632** | -0.346 | -0.032* | -0.016 |
| | (0.249) | (0.292) | (0.017) | (0.020) |
| SC/ST/OBC | -1.016*** | -0.953*** | -0.061*** | -0.045*** |
| | (0.185) | (0.225) | (0.014) | (0.017) |
| DiD | -0.320 | -0.532* | -0.048** | -0.071*** |
| | (0.249) | (0.296) | (0.019) | (0.022) |
| Observations | 27,297 | 19,257 | 26,920 | 18,971 |
| Clusters | 1,124 | 786 | 1,124 | 786 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 50.00 |
| Mean Dep. Var. | 6.668 | 6.477 | 0.696 | 0.683 |

| | *Wealth Index* | | *Log(wealth score)* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Hyderabad | -0.186** | -0.172** | -0.043* | -0.039 |
| | (0.077) | (0.087) | (0.025) | (0.028) |
| SC/ST/OBC | -0.318*** | -0.301*** | -0.090*** | -0.086*** |
| | (0.054) | (0.063) | (0.018) | (0.021) |
| DiD | -0.049 | -0.054 | -0.031 | -0.035 |
| | (0.075) | (0.085) | (0.024) | (0.028) |
| Observations | 17,007 | 11,996 | 17,007 | 11,996 |
| Clusters | 1,123 | 786 | 1,123 | 786 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 50.00 |
| Mean Dep. Var. | 3.073 | 3.012 | 12.287 | 12.268 |

Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude . Regressions control for age, age squared, gender, household size, rural/urban sector and boundary segment fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## C.4 Results for caste groups without Berar

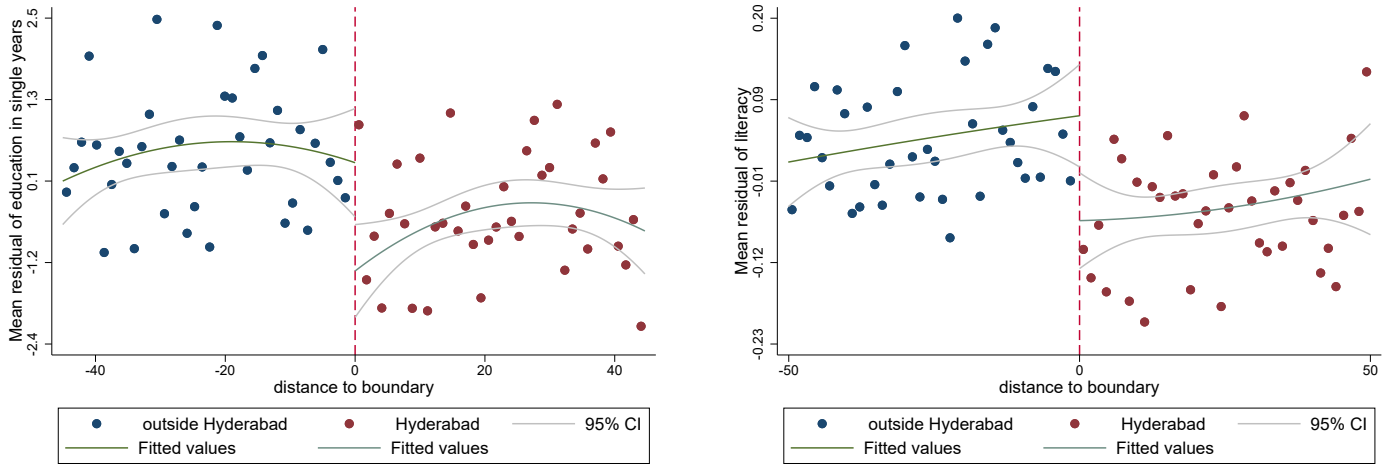
Table C-4: Results for caste groups without Berar

| | *Disadvantaged Castes* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.766*** | -0.680*** | -1.380*** | -1.326*** | -0.074*** | -0.079*** | -0.147*** | -0.148*** |
| | (0.191) | (0.212) | (0.347) | (0.506) | (0.016) | (0.019) | (0.031) | (0.044) |
| Observations | 17,135 | 12,190 | 17,135 | 9,016 | 16,842 | 11,958 | 16,842 | 9,302 |
| Clusters | 868 | 613 | 868 | 449 | 868 | 613 | 868 | 473 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.81 | 75.00 | 50.00 | 75.00 | 38.17 |
| Mean Dep. Var. | 6.141 | 5.974 | 6.141 | 5.891 | 0.642 | 0.631 | 0.642 | 0.630 |
| SD Dep. Var. | 5.169 | 5.161 | 5.169 | 5.141 | 0.479 | 0.482 | 0.479 | 0.483 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.253*** | -0.235*** | -0.421*** | -0.273* | -0.079*** | -0.075*** | -0.127*** | -0.104** |
| | (0.062) | (0.066) | (0.102) | (0.165) | (0.020) | (0.021) | (0.033) | (0.052) |
| Observations | 10,677 | 7,545 | 10,680 | 4,624 | 10,677 | 7,545 | 10,680 | 5,051 |
| Clusters | 868 | 613 | 868 | 375 | 868 | 613 | 868 | 413 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 29.89 | 75.00 | 50.00 | 75.00 | 33.07 |
| Mean Dep. Var. | 3.015 | 2.968 | 3.015 | 2.961 | 12.274 | 12.259 | 12.274 | 12.254 |
| SD Dep. Var. | 1.144 | 1.133 | 1.144 | 1.121 | 0.372 | 0.373 | 0.372 | 0.365 |
| | *Advantaged Castes* | | | | | | | |
| | *Years of Education* | | | | *Literacy* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | 0.344 | 0.296 | -0.142 | -0.549 | 0.021 | 0.007 | -0.003 | -0.016 |
| | (0.331) | (0.387) | (0.557) | (0.782) | (0.022) | (0.026) | (0.042) | (0.064) |
| Observations | 4,138 | 2,938 | 4,138 | 1,948 | 4,098 | 2,918 | 4,098 | 2,188 |
| Clusters | 560 | 391 | 560 | 256 | 557 | 391 | 557 | 292 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 31.34 | 75.00 | 50.00 | 75.00 | 35.80 |
| Mean Dep. Var. | 8.090 | 8.114 | 8.090 | 8.149 | 0.795 | 0.793 | 0.795 | 0.791 |
| SD Dep. Var. | 4.894 | 4.906 | 4.894 | 5.022 | 0.404 | 0.405 | 0.404 | 0.407 |
| | *Wealth Index* | | | | *Log(wealth score)* | | | |
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.087 | -0.092 | -0.157 | -0.188 | -0.026 | -0.026 | -0.049 | -0.069 |
| | (0.096) | (0.110) | (0.158) | (0.197) | (0.029) | (0.034) | (0.050) | (0.062) |
| Observations | 2,622 | 1,877 | 2,627 | 1,417 | 2,622 | 1,877 | 2,627 | 1,412 |
| Clusters | 560 | 391 | 560 | 292 | 560 | 391 | 560 | 290 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.71 | 75.00 | 50.00 | 75.00 | 35.46 |
| Mean Dep. Var. | 3.460 | 3.447 | 3.460 | 3.435 | 12.404 | 12.402 | 12.404 | 12.396 |
| SD Dep. Var. | 1.196 | 1.189 | 1.196 | 1.208 | 0.377 | 0.379 | 0.377 | 0.385 |

We do our analysis for the sample which excludes data in Berar and nearby districts in the princely state. Standard errors clustered at the NFHS cluster level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. Local Linear RD regressions include a local linear specification estimated separately on each side of the boundary (Cattaneo et al. (2020)) and use a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## C.5 RD Plots for quadratic polynomial by caste groups

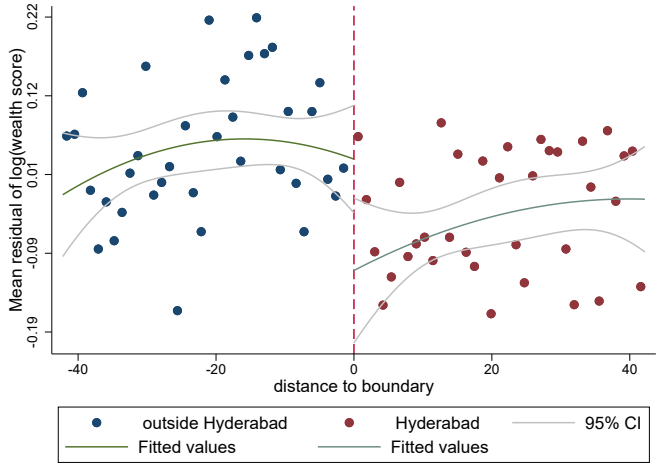
Figure C-1: RD plots of outcomes for Disadvantaged caste



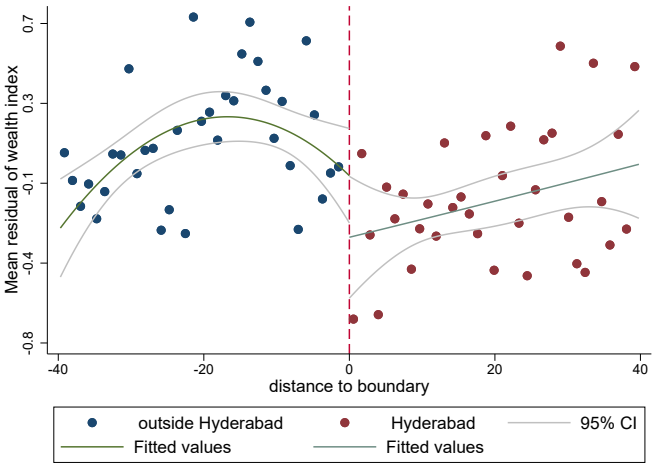

(a) Years of education




(b) Literacy



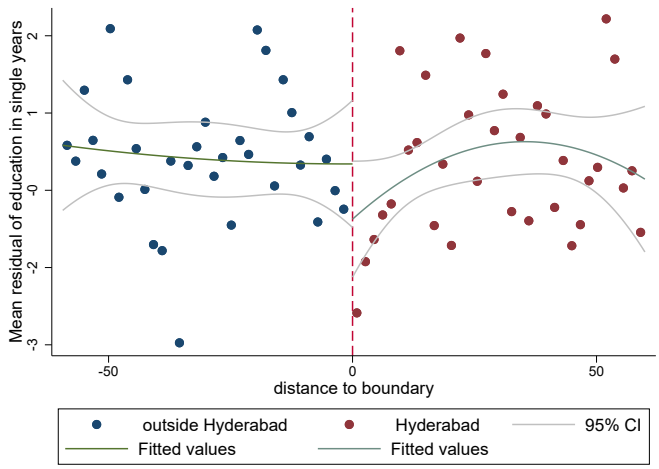

(c) log(wealth)



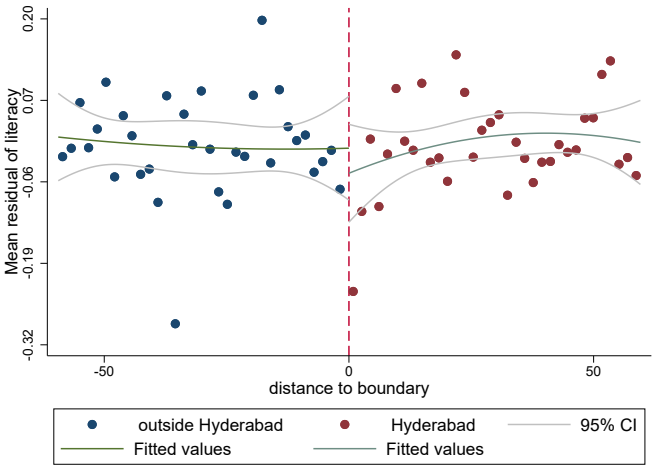

(d) Wealth Index

Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 1 km bin along the running variable as well as with a local quadratic trend estimated separately on each side of the discontinuity. We restrict the sample to disadvantaged groups. Each regression is estimated using the optimal bandwidth chosen through the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

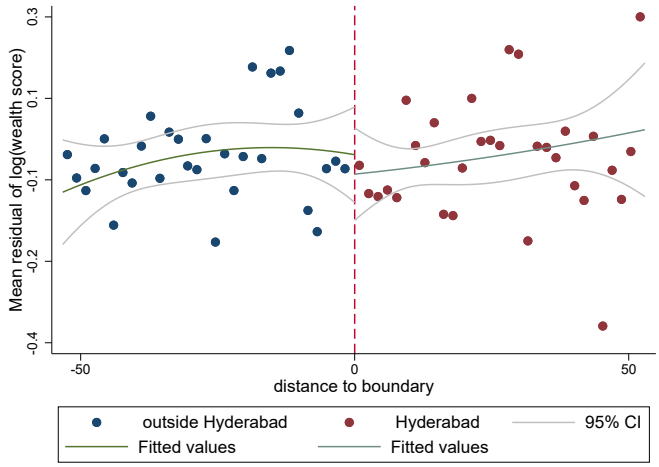
Figure C-2: RD plots of outcomes for Advantaged caste



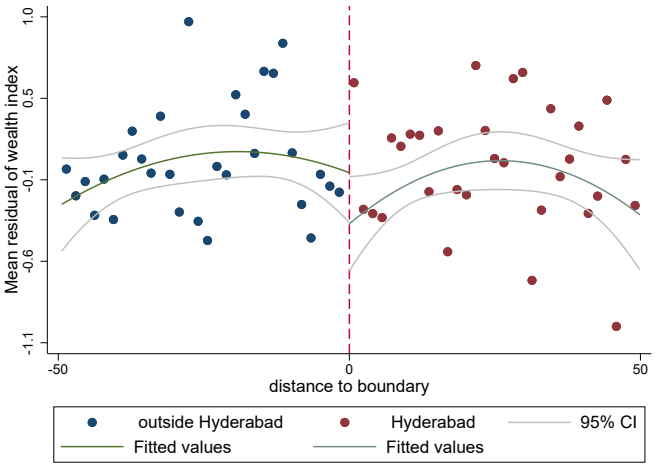

(a) Years of education




(b) Literacy




(c) log(wealth)




(d) Wealth Index

Notes: The figure presents RD plots for our main outcomes and the mean value of each outcome variable at each 1 km bin along the running variable as well as with a local quadratic trend estimated separately on each side of the discontinuity. We restrict the dataset to advantaged groups. Each regression is estimated using the optimal bandwidth chosen through the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Standard errors are clustered at the DHS cluster level and the figures show 95% confidence intervals.

## C.6 Simulations from spatial noise

Table C-5: P-values from placebo regressions as per Kelly (2021)

| | Disadvantaged Castes | | | | Advantaged Castes | | | |
|---|---|---|---|---|---|---|---|---|
| | Placebo p-value | Original p-value | Coeff | Bandwidth | Placebo p-value | Original p-value | Coeff | Bandwidth |
| Literacy | 0.003 | 0 | -0.090 | 46.77 | 0.806 | 0.598 | 0.007 | 46.77 |
| Years of education | 0.000 | 0 | -0.958 | 36.35 | 0.988 | 0.972 | -0.006 | 36.35 |
| Wealth Index | 0.002 | 0 | -0.263 | 35.40 | 0.565 | 0.140 | -0.083 | 35.40 |
| Log(wealth score) | 0.004 | 0 | -0.081 | 37.05 | 0.781 | 0.490 | -0.013 | 37.05 |

The table shows p-values from spatial placebo regressions using spatial noise as the independent variable (Voth (2020)) as suggested by Kelly (2021). Placebo p values indicate the fraction of times (out of 10000) when the t-statistic of the placebo regression was higher than the original specification. We use an average of optimal bandwidth for disadvantaged and advantaged caste using methods from Cattaneo et al. (2020) for each variable. Note that these estimates are based on the linear polynomial results which control for latitude and longitude and not the local linear RD results based on Cattaneo et al. (2020). * p < 0.10, ** p <0.05, *** p<0.01.

## C.7 Results with conley standard errors by caste

Table C-6: Results with conley standard errors - Disadvantaged castes

| | *Years of Education* | | *Literacy* | |
|---|---|---|---|---|
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.983*** | -0.960*** | -0.088*** | -0.088*** |
| | (0.288) | (0.303) | (0.027) | (0.028) |
| Observations | 21,245 | 12,617 | 20,920 | 17,023 |
| Bandwidth | 75.00 | 41.78 | 75.00 | 58.74 |
| Mean Dep. Var. | 6.334 | 6.062 | 0.665 | 0.654 |
| SD Dep. Var. | 5.088 | 5.053 | 0.472 | 0.476 |
| | *Wealth Index* | | *Log(wealth score)* | |
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.275*** | -0.299** | -0.085*** | -0.092** |
| | (0.094) | (0.118) | (0.030) | (0.037) |
| Observations | 13,223 | 6,047 | 13,223 | 6,603 |
| Bandwidth | 75.00 | 31.98 | 75.00 | 34.85 |
| Mean Dep. Var. | 2.980 | 2.909 | 12.259 | 12.235 |
| SD Dep. Var. | 1.168 | 1.140 | 0.387 | 0.379 |

Standard errors are Conley (1999) spatial standard errors assuming a cut-off window of 50 kms. All regressions include a local linear specification estimated separately on each side of the concession boundary. Column 1 for each outcome uses a fixed bandwidth of 75 km and column 2 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

Table C-7: Results with conley standard errors - Advantaged castes

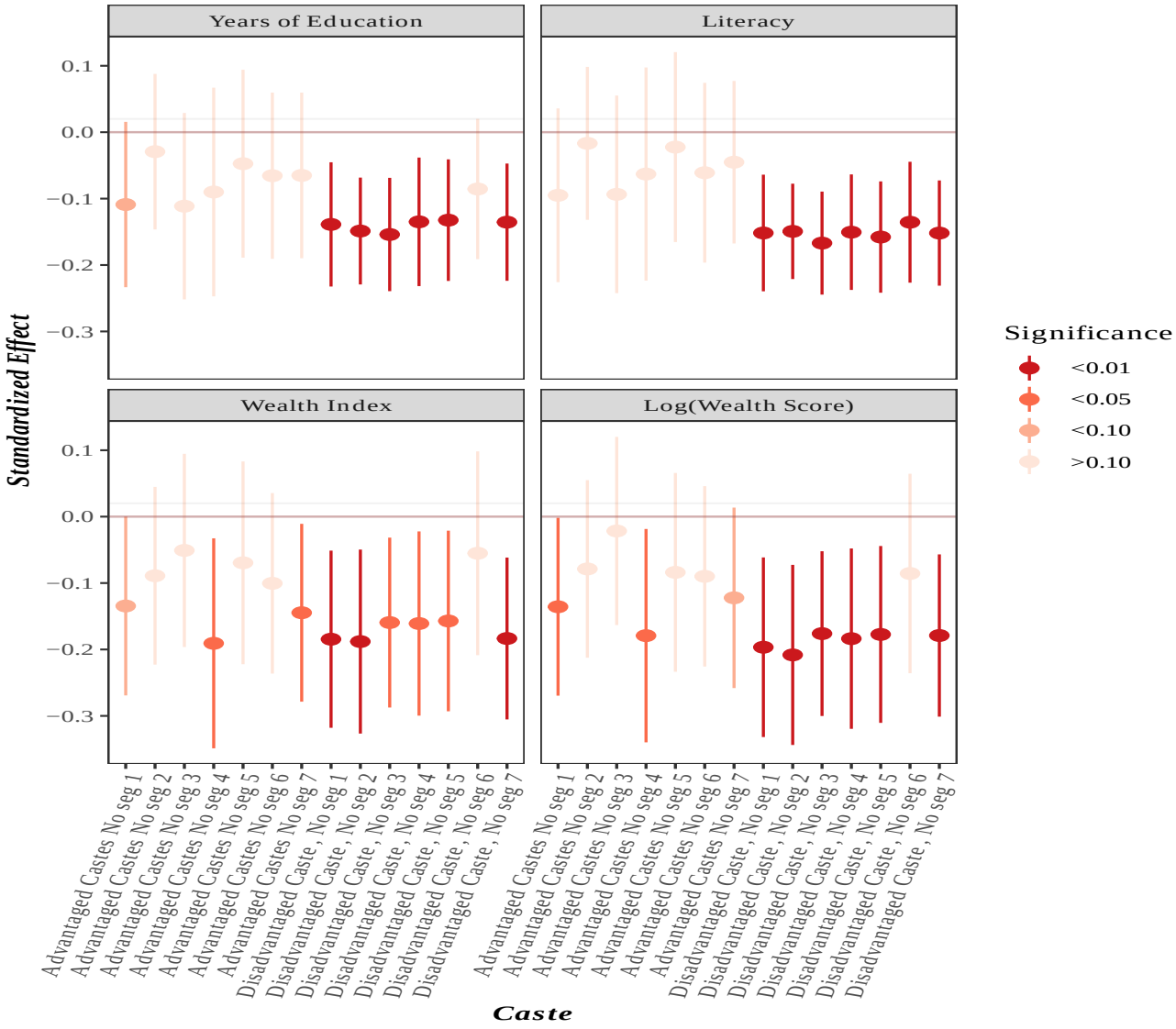
| | *Years of Education* | | *Literacy* | |
|---|---|---|---|---|
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.481 | -0.262 | -0.014 | -0.014 |
| | (0.345) | (0.467) | (0.023) | (0.031) |
| Observations | 6,052 | 2,664 | 6,000 | 3,060 |
| Bandwidth | 75.00 | 31.04 | 75.00 | 34.90 |
| Mean Dep. Var. | 7.765 | 7.802 | 0.781 | 0.776 |
| SD Dep. Var. | 4.853 | 4.946 | 0.413 | 0.417 |

| | *Wealth Index* | | *Log(wealth score)* | |
|---|---|---|---|---|
| | Local Linear RD | | Local Linear RD | |
| | (1) | (2) | (3) | (4) |
| **Hyderabad** | -0.306** | -0.280 | -0.082* | -0.071 |
| | (0.140) | (0.183) | (0.044) | (0.057) |
| Observations | 3,811 | 2,005 | 3,811 | 1,975 |
| Bandwidth | 75.00 | 36.14 | 75.00 | 35.91 |
| Mean Dep. Var. | 3.336 | 3.295 | 12.364 | 12.354 |
| SD Dep. Var. | 1.219 | 1.221 | 0.390 | 0.393 |

Standard errors are Conley (1999) spatial standard errors assuming a cut-off window of 50 kms. All regressions include a local linear specification estimated separately on each side of the concession boundary. Column 1 for each outcome uses a fixed bandwidth of 75 km and column 2 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). Regressions control for age, age squared, gender for educational outcomes and household size for wealth outcomes. We also use boundary segment fixed effects. Literacy is an indicator variable equal to 0 if the respondent cannot read at all and 1 otherwise. Wealth Score is an index generated by the NFHS using principle component of asset ownership. Wealth Index is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile from the Wealth Score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## C.8 Robustness checks for different segments by caste groups

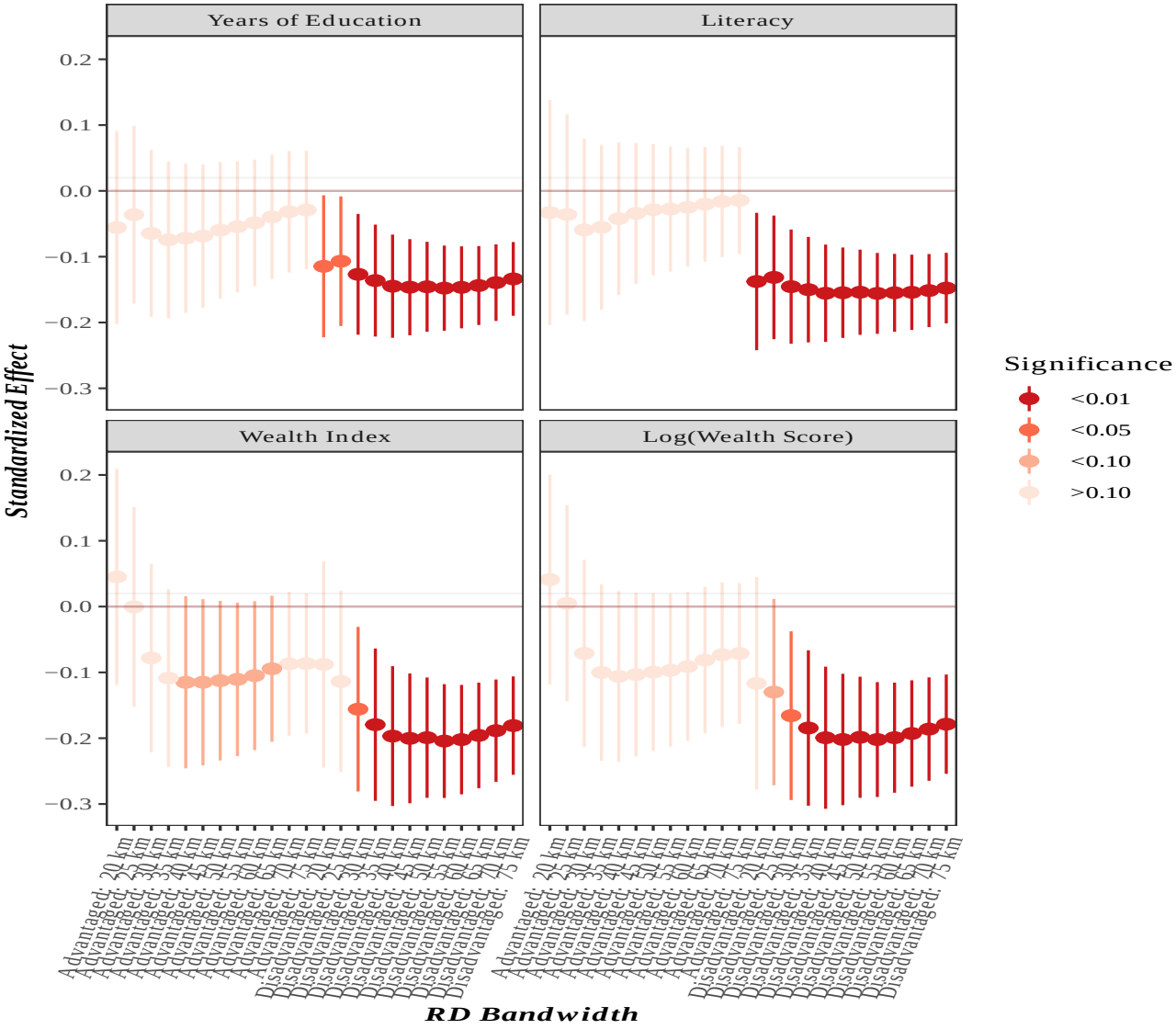
Figure C-3: Robustness checks for different segments by caste groups




Notes: The figure presents standardized RD coefficients for our main outcomes for different caste groups after excluding each segment. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. The figure shows result for a triangular kernel, and optimal bandwith chosen using procedure suggested by Cattaneo et al. (2020). Figure shows point estimates with 95% confidence intervals.

## C.9 Robustness checks for different bandwidths by caste groups

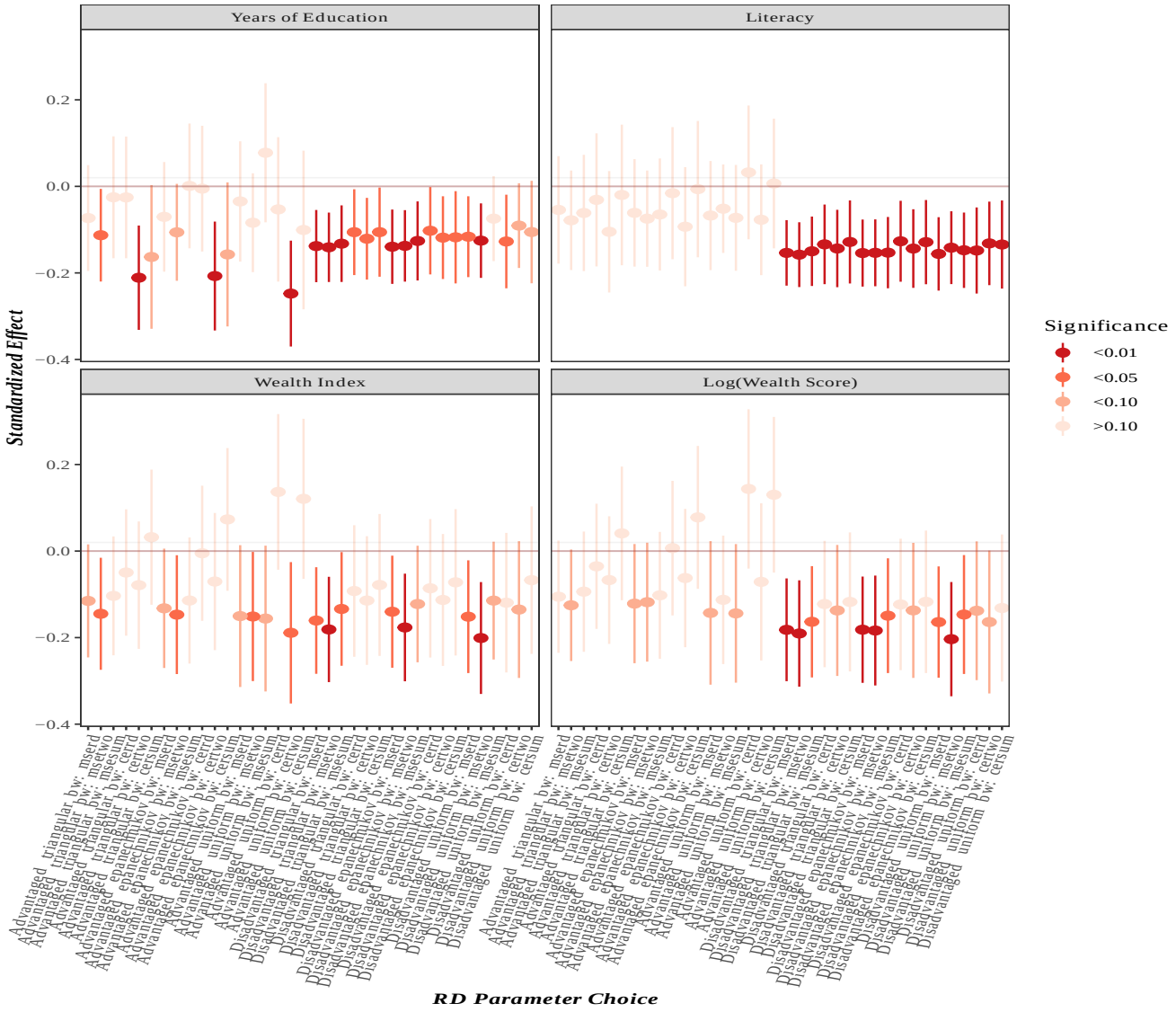
Figure C-4: Robustness checks for different bandwidths by caste groups




Notes: The figure presents RD coefficients for our main outcomes for different bandwidths by caste groups. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. The figure shows result for a triangular kernel, and optimal bandwith chosen using MSE minimizing procedure suggested by Cattaneo et al. (2020). Figure shows point estimates with 95% confidence intervals.

## C.10 Robustness checks for different parameters by caste groups

Figure C-5: Robustness checks for different parameters for different castes

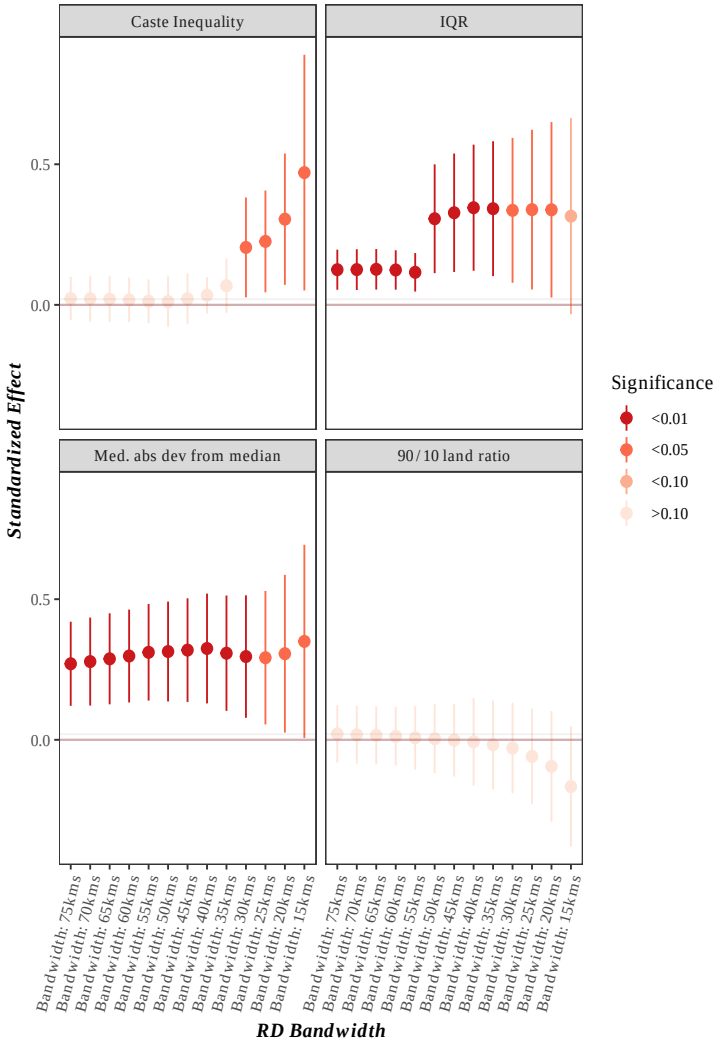



Notes: The figure presents RD coefficients for our main outcomes for different values of the RD Parameters separated by caste. Regressions control for age, age squared and gender for educational outcomes and household size for wealth outcomes as well as boundary segment fixed effects for all outcomes. Bandwidth Type represents the optimal bandwidth selection procedure used for each regression: mserd choses one common MSE-optimal bandwidth; msetwo choses two different MSE-optimal bandwidths (below and above the cutoff); msesum choses one common MSE-optimal bandwidth selector for the sum of regression estimates (instead of the difference); cerrd choses one common CER-optimal bandwidth; certwo two different CER-optimal bandwidths (below and above the cutoff); cersum choses one common CER-optimal bandwidth for the sum of regression estimates. Figure shows point estimates and 95% confidence intervals.

# D Measures of Inequality

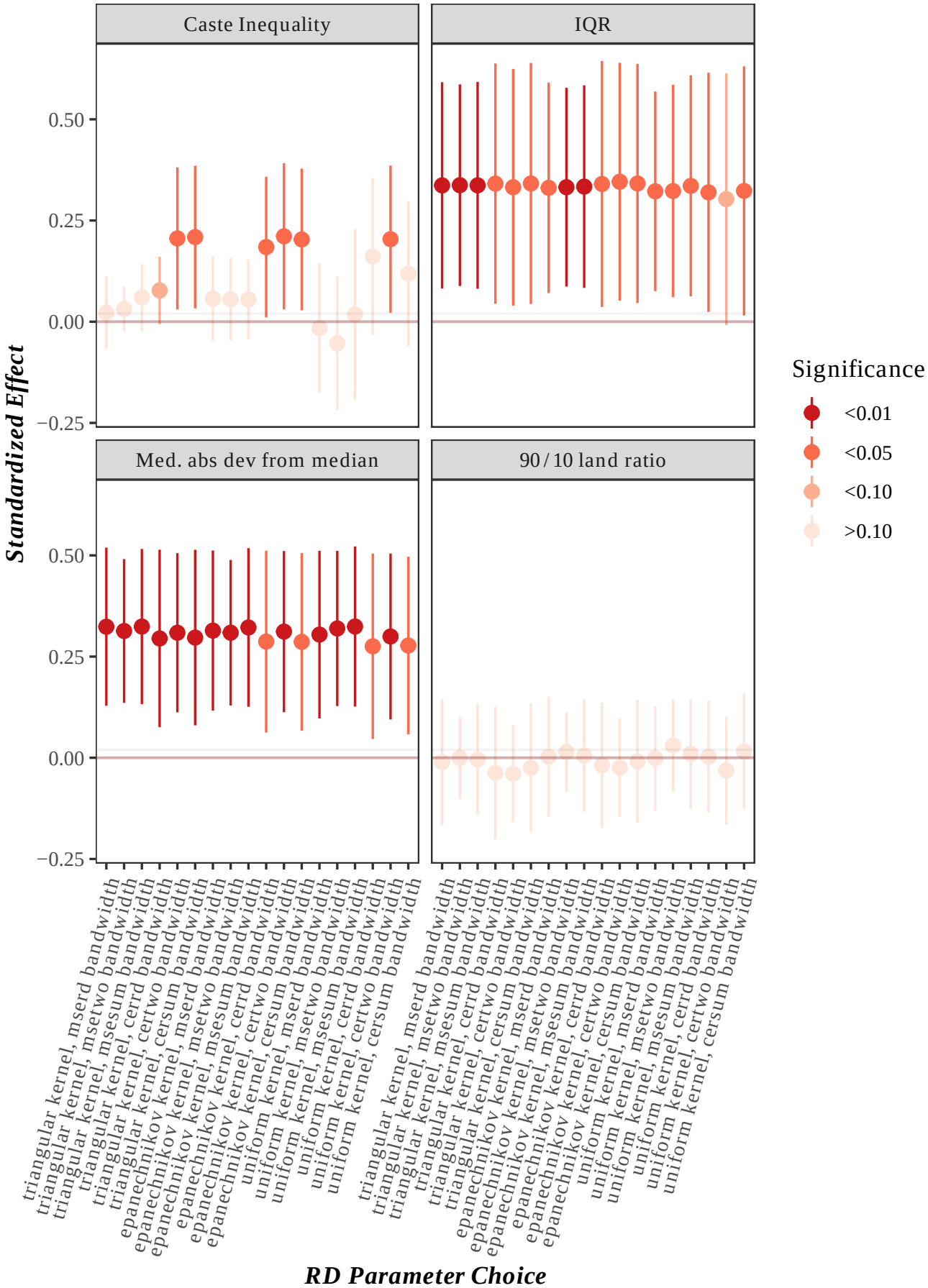
## D.1 Land inequality

Figure D-1: Robustness of Land inequality to distance



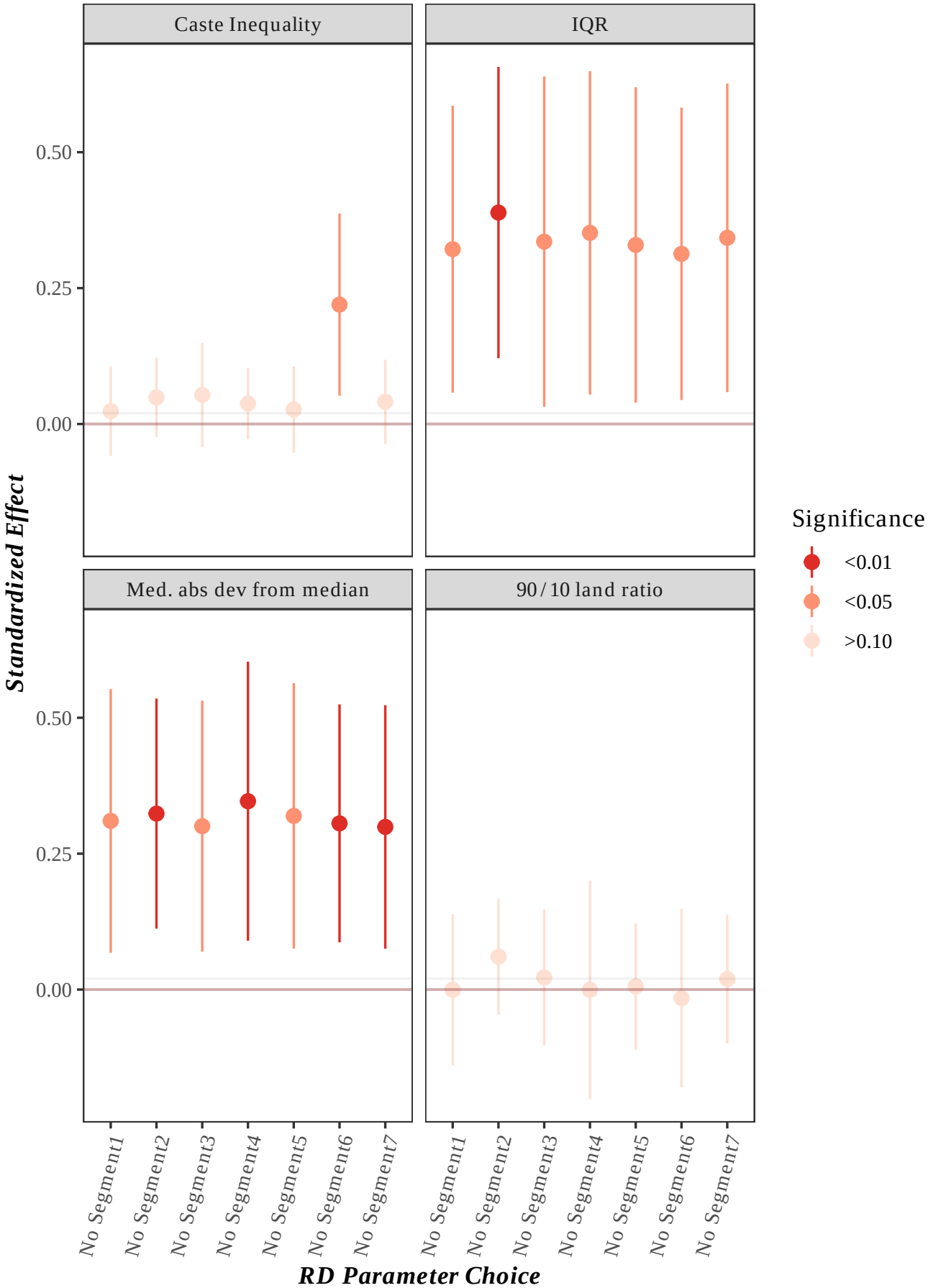

Notes: The figure presents standardized RD coefficients for our land inequality outcomes for different distances to the boundary. Regressions control for boundary segment fixed effects. Caste inequality is difference between median land of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in land at the cluster level. 90 to 10 land ratio is the difference in land between the 90th percentile and the 10th percentile at the cluster level. Figure shows point estimates and 95% confidence intervals.

Figure D-2: Land inequality with different parameters for a linear polynomial




Notes: The figure presents standardized RD coefficients for our land inequality outcomes for different parameters of the Linear RD . Regressions control for boundary segment fixed effects. Caste inequality is difference between median land of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in land at the cluster level. 90 to 10 land ratio is the difference in land between the 90th percentile and the 10th percentile at the cluster level. Figure shows point estimates and 95% confidence intervals.

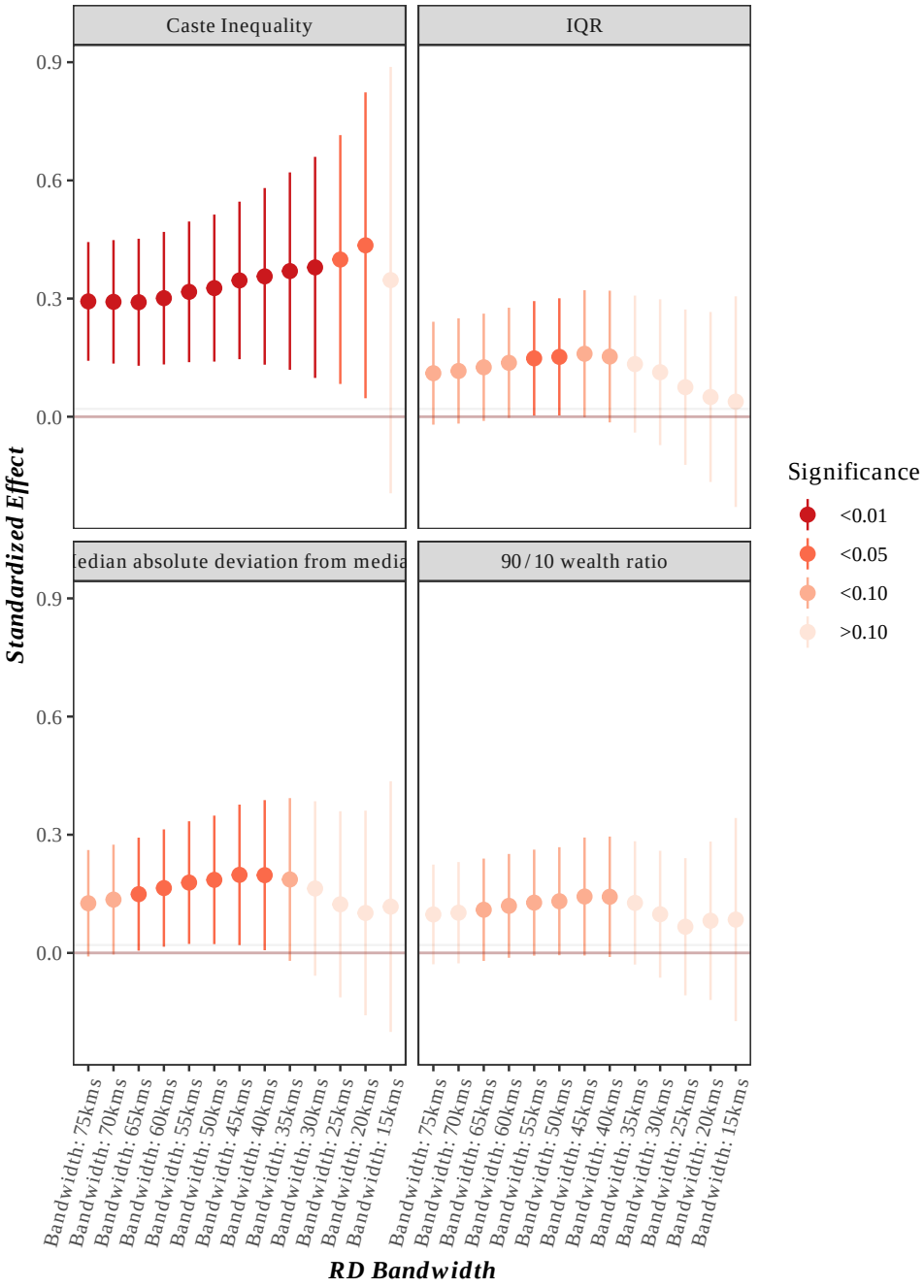
Figure D-3: Land inequality after removing segments




Notes: The figure presents standardized RD coefficients for our land inequality outcomes after removing different segments. Regressions control for boundary segment fixed effects. Caste inequality is difference between median land of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in land at the cluster level. 90 to 10 land ratio is the difference in land between the 90th percentile and the 10th percentile at the cluster level. Figure shows point estimates and 95% confidence intervals.

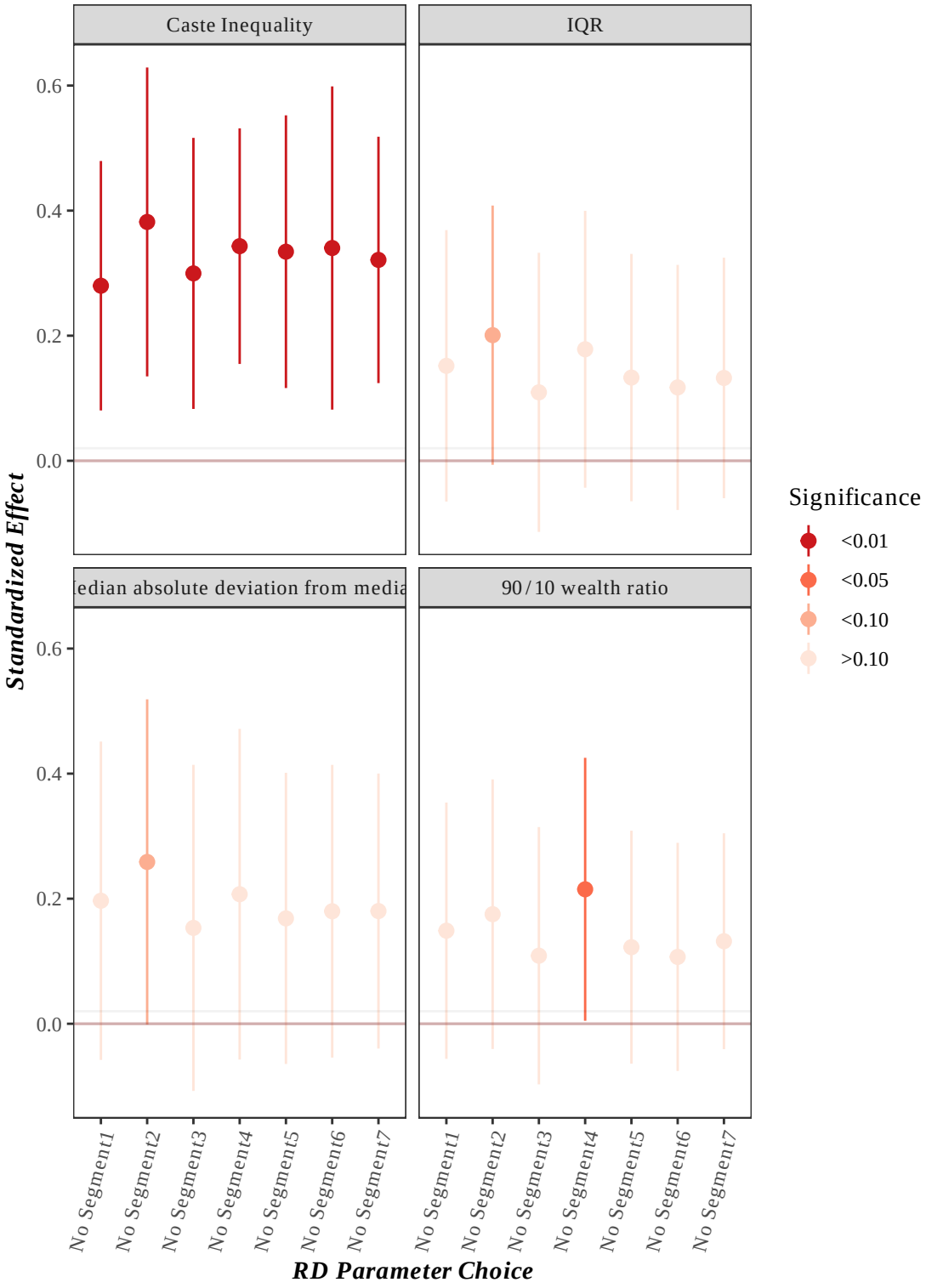
## D.2 Wealth inequality

Figure D-4: Robustness to distance



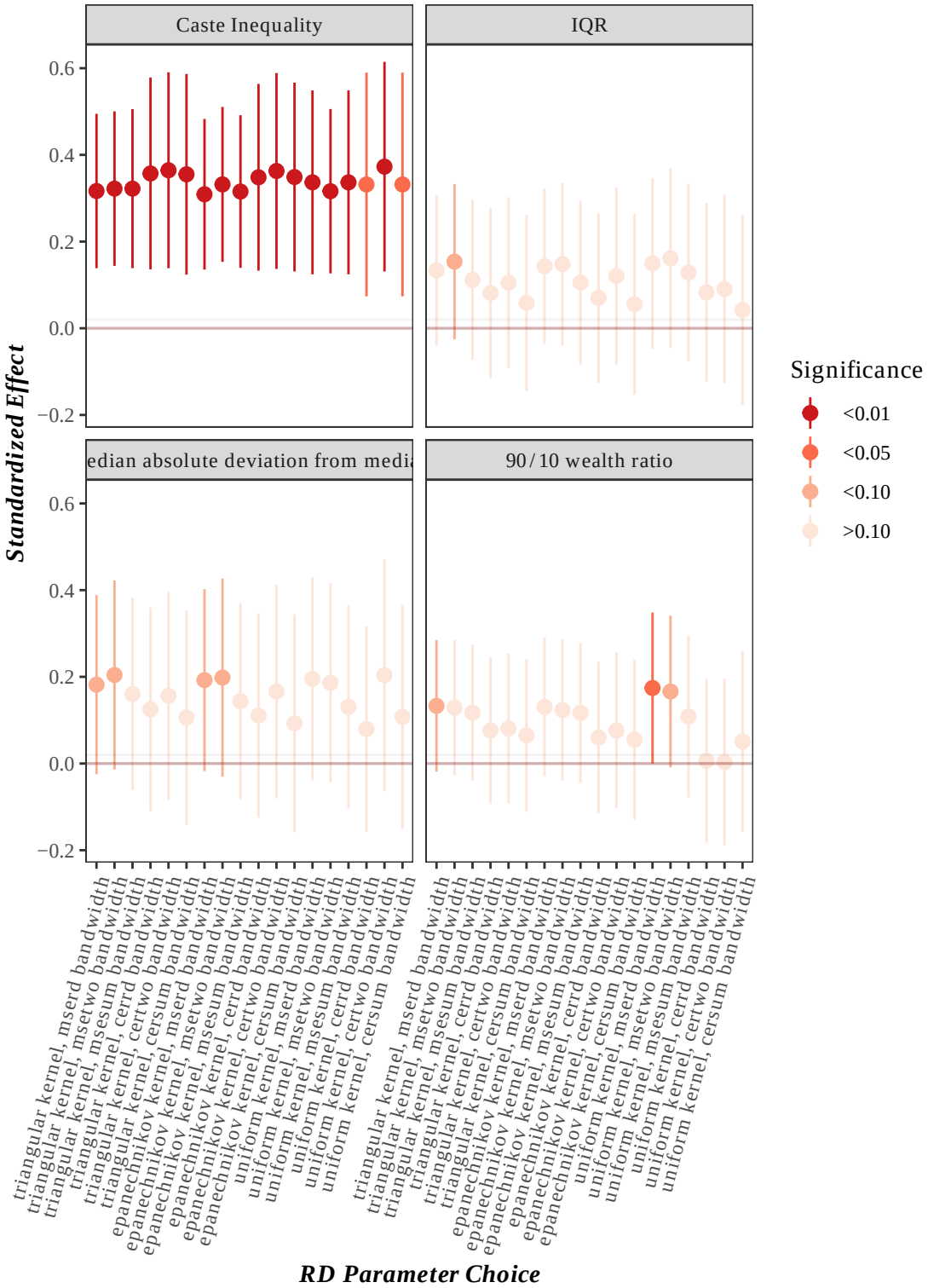

Notes: The figure presents standardized RD coefficients for our wealth inequality outcomes. Regressions control for boundary segment fixed effects. Caste inequality is difference between median wealth of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in wealth at the cluster level. 90 to 10 wealth ratio is the ratio of the wealth of the 90th percentile to the 10th percentile at the cluster level.

Figure D-6: Robustness to removing segments




Notes: The figure presents standardized RD coefficients for our wealth inequality outcomes after removing different boundary segments . Regressions control for boundary segment fixed effects and do not include clusters within 10 km of the former *Hyderabad* boundary. Caste inequality is difference between median wealth of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in wealth at the cluster level. 90 to 10 wealth ratio is the ratio of the wealth of the 90th percentile to the 10th percentile at the cluster level.

Figure D-5: Robustness to different parameters for a linear polynomial




Notes: The figure presents standardized RD coefficients for our wealth inequality outcomes for different parameters in the RD. Regressions control for boundary segment fixed effects. Caste inequality is difference between median wealth of disadvantaged and advantaged caste group at cluster level. IQR is the interquartile range in wealth at the cluster level. 90 to 10 wealth ratio is the ratio of the wealth of the 90th percentile to the 10th percentile at the cluster level.

# E Intermediate Outcomes and Mechanisms

## E.1 Occupation in 1911

Table E-1: Occupation in 1911

| | *Occupation in 1911* | | | | | |
|---|---|---|---|---|---|---|
| | *% Landlords* | | *% Laborers* | | *% Moneylender*1000* | |
| | None | Linear coordinates | None | Linear coordinates | None | Linear coordinates |
| psoh | 0.032*** | 0.032*** | 0.000 | 0.002 | 0.033 | 0.034 |
| | (0.005) | (0.005) | (0.022) | (0.022) | (0.027) | (0.027) |
| Observations | 28 | 28 | 28 | 28 | 28 | 28 |
| $R^2$ | 0.73 | 0.74 | 0.43 | 0.46 | 0.14 | 0.18 |
| mean | .023 | .023 | .221 | .221 | .114 | .114 |

The dependent variable is % landlords, labourers and moneylenders in the population. The linear coordinate specification controls for a linear polynomial in the latitude and longitude of the district centroid and the first one does not. We include fixed effects for boundary segment. Robust standard errors are in parenthesis. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

## E.2 Caste groups in 1931

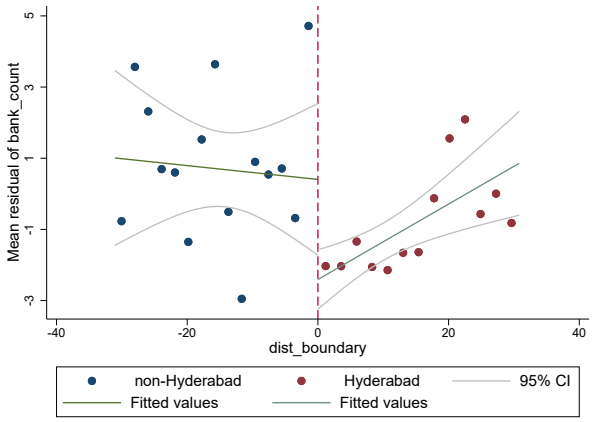
Table E-2: Major caste groups in 1931

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Advantaged Castes | | Disdvantaged Castes | |
| Segment | Indirect rule | Direct rule | Indirect rule | Direct rule |
| 1 | Maratha | Maratha | Lingayat | Lingayat |
| | Brahman | Brahman | Mala, Mahar | Yadava |
| | Rajput | Rajput | Madiga | Bedar |
| | | | Yadava | Madiga |
| | | | Bedar | Mahar |
| 2 | Brahman | Brahman | Bedar | Lingayat |
| | Komati | Maratha | Lingayat | Boya |
| | Maratha | Komati | Madiga | Yadava |
| | | | Yadava | Madiga |
| | | | Kunbi | Bedar |
| 3 | Maratha | Maratha | Lingayat | Mala, Mahar |
| | Brahman | Brahman | Mala, Mahar | Yadava |
| | Rajput | Rajput | Madiga | Mali |
| | | | Yadava | Lingayat |
| | | | Mali | Madiga |
| 4 | Maratha | Maratha | Mala, Mahar | Mala, Mahar |
| | Brahman | Brahman | Yadava | Kunbi |
| | Komati | Rajput | Madiga | Bhil |
| | | | Mali | Mali |
| | | | Kunbi | Yadava |
| 5 | Maratha | Maratha | Mala, Mahar | Mala, Mahar |
| | Brahman | Brahman | Madiga | Mali |
| | Komati | Rajput | Lingayat | Madiga |
| | | | Yadava | Kunbi |
| | | | Lambadi | Lambadi |
| 6 | Kapu | Kapu | Madiga | Boya |
| | Komati | Komati | Yadava | Madiga |
| | Brahman | Brahman | Mahar | Yadava |
| | | | Mala | Mala |
| | | | Eadiga,Goundla | Mahar |
| 7 | Kapu | Kapu | Madiga | Mala |
| | Komati | Komati | Yadava | Mahar |
| | Brahman | Brahman | Eadiga,Goundla | Madiga |
| | | | Mala | Yadava |
| | | | Mahar | Dhobi |

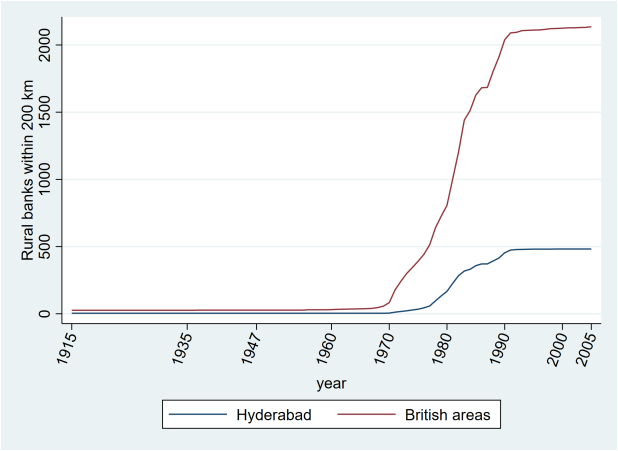
The table shows the biggest caste groups (in descending order of size) in direct and indirect rule areas for the districts along the 7 segments in our boundary. The data is taken from census of 1931. We show the 5 largest caste groups for disadvantaged castes and the two largest for advantaged castes. As the census was done by different authorities in different states, the categorization is not uniform across states. Some caste groups were lumped into one in some areas and therefore it is difficult to disaggregate the population for the lumped caste groups. Note that while we have listed *Lingayats* under disadvantaged caste, there are some subgroups within Lingayats who are categorized as advantaged castes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## E.3 RD Plots for Banks

Figure E-1: Number of banks




(a) RD Plot of banks with distance to the boundary




(b) Rural banks within 200 km of the boundary

## E.4 Yields in 1956-61

Table E-3: Crop yields in 1956-61

| | None | Linear coordinates polynomial | Cubic distance polynomial |
|---|---|---|---|
| | (1) | (2) | (3) |
| *wheat* | -0.101*** | -0.089*** | -0.078*** |
| | (0.018) | (0.021) | (0.019) |
| Observations | 147 | 147 | 147 |
| Mean | 0.346 | 0.346 | 0.346 |
| *rice* | -0.100 | -0.169** | -0.110 |
| | (0.072) | (0.074) | (0.079) |
| Observations | 180 | 180 | 180 |
| Mean | 0.945 | 0.945 | 0.945 |
| *cotton* | -0.039*** | -0.035*** | -0.041*** |
| | (0.006) | (0.005) | (0.006) |
| Observations | 167 | 167 | 167 |
| Mean | 0.078 | 0.078 | 0.078 |
| *sorghum* | -0.032 | -0.024 | 0.004 |
| | (0.027) | (0.027) | (0.025) |
| Observations | 180 | 180 | 180 |
| Mean | 0.513 | 0.513 | 0.513 |
| *groundnut* | -0.140*** | -0.155*** | -0.125*** |
| | (0.044) | (0.041) | (0.043) |
| Observations | 172 | 172 | 172 |
| Mean | 0.674 | 0.674 | 0.674 |
| *pigeonpea* | -0.205*** | -0.173*** | -0.187*** |
| | (0.035) | (0.030) | (0.031) |
| Observations | 180 | 180 | 180 |
| Mean | 0.442 | 0.442 | 0.442 |
| *sugarcane* | -0.161 | -0.761* | -0.287 |
| | (0.439) | (0.446) | (0.398) |
| Observations | 165 | 165 | 165 |
| Mean | 6.098 | 6.098 | 6.098 |
| *pearl_millet* | -0.057** | -0.065** | -0.058** |
| | (0.027) | (0.030) | (0.029) |
| Observations | 165 | 165 | 165 |
| Mean | 0.345 | 0.345 | 0.345 |
| *sesamum* | -0.061*** | -0.070*** | -0.062*** |
| | (0.009) | (0.008) | (0.009) |
| Observations | 150 | 150 | 150 |
| Mean | 0.198 | 0.198 | 0.198 |

The dependent variable is yield in tonnes per hectare between 1956 and 1961 at the district level. We use districts along the boundary of the princely state of Hyderabad. All specifications control for mean altitude. We include fixed effects for boundary segment. The second column controls for a linear polynomial in coordinates and the third columns controls for a cubic polynomial in distance to the boundary. We only show results for crops which have a substantial share of cropping area. Robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## E.5 Effect of Banks

Table E-4: Change in yields after rural bank policy

| | None | Cubic distance polynomial | Linear coordinates |
|---|---|---|---|
| | (1) | (2) | (3) |
| DID estimate of wheat yield | -0.224* | -0.226* | -0.227* |
| | (0.118) | (0.122) | (0.117) |
| Observations | 52 | 52 | 52 |
| Mean | 0.980 | 0.980 | 0.980 |
| DID estimate of rice yield | 0.084 | 0.084 | 0.084 |
| | (0.190) | (0.178) | (0.188) |
| Observations | 62 | 62 | 62 |
| Mean | 1.539 | 1.539 | 1.539 |
| DID estimate of cotton yield | -0.028 | -0.029 | -0.028 |
| | (0.027) | (0.026) | (0.027) |
| Observations | 59 | 59 | 59 |
| Mean | 0.208 | 0.208 | 0.208 |
| DID estimate of sorghum yield | -0.109 | -0.109 | -0.109 |
| | (0.088) | (0.088) | (0.081) |
| Observations | 62 | 62 | 62 |
| Mean | 0.785 | 0.785 | 0.785 |
| DID estimate of groundnut yield | -0.017 | -0.017 | -0.017 |
| | (0.089) | (0.086) | (0.089) |
| Observations | 62 | 62 | 62 |
| Mean | 0.926 | 0.926 | 0.926 |
| DID estimate of pigeonpea yield | -0.004 | -0.004 | -0.004 |
| | (0.086) | (0.080) | (0.086) |
| Observations | 62 | 62 | 62 |
| Mean | 0.531 | 0.531 | 0.531 |
| DID estimate of sugarcane yield | -0.696 | -0.696 | -0.696 |
| | (0.684) | (0.637) | (0.697) |
| Observations | 62 | 62 | 62 |
| Mean | 7.851 | 7.851 | 7.851 |
| DID estimate of pearl_millet yield | -0.100 | -0.100* | -0.100 |
| | (0.066) | (0.059) | (0.064) |
| Observations | 60 | 60 | 60 |
| Mean | 0.628 | 0.628 | 0.628 |
| DID estimate of sesamum yield | -0.005 | -0.005 | -0.005 |
| | (0.021) | (0.020) | (0.021) |
| Observations | 62 | 62 | 62 |
| Mean | 0.252 | 0.252 | 0.252 |

The dependent variable is mean yield in tonnes per hectare before and after 1971. All specifications control for mean altitude of districts and include fixed effects for boundary segment. Robust standard errors are in parenthesis. The sample includes districts which border *Hyderabad*. * p < 0.10, ** p <0.05, *** p<0.01.

Table E-5: Agricultural debt

| | *Non institutional debt* | | |
|---|---|---|---|
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 5795.219*** | 7763.970*** | 3028.326 |
| | (2061.382) | (2329.964) | (4895.768) |
| Observations | 1,767 | 1,767 | 1,767 |
| $R^2$ | 0.05 | 0.06 | 0.02 |
| clusters | 347 | 347 | 347 |
| mean | 16784.154 | 16784.154 | 16784.154 |
| | *Non institutional debt - Disadvantaged castes* | | |
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 6777.398*** | 10004.730*** | 3998.802* |
| | (2307.956) | (2921.220) | (2301.629) |
| Observations | 1,381 | 1,381 | 1,381 |
| $R^2$ | 0.06 | 0.07 | 0.06 |
| clusters | 306 | 306 | 306 |
| mean | 13379.862 | 13379.862 | 13379.862 |
| | *Non institutional debt - Advantaged castes* | | |
| | None | Cubic distance polynomial | Linear coordinates polynomial |
| | (1) | (2) | (3) |
| Hyderabad | 1908.452 | 1179.376 | 284.127 |
| | (4785.960) | (4606.103) | (9030.898) |
| Observations | 386 | 386 | 386 |
| $R^2$ | 0.13 | 0.15 | 0.13 |
| clusters | 155 | 155 | 155 |
| mean | 18978.56 | 18978.56 | 18978.56 |

*Notes* : The dependent variable is farmer debt from non-institutional sources in INR. All specifications control for mean altitude of district and household size. The last two specifications control for a geographic polynomial in the latitude and longitude of the district centroid and the first one does not. We include fixed effects for boundary segment and use sample weights provided in the NSS59. The sample includes districts which border *Hyderabad*. Errors are clustered at the nss cluster level. Robust standard errors are in parenthesis. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

## E.6 Public Infrastructure

Table E-6: Public Infrastructure

| | *Primary Schools in 1961* | | | | *High Schools in 1961* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -43.762*** | -44.061*** | -44.600** | -38.107 | -2.809*** | -2.967*** | -3.145** | -3.259 |
| | (10.584) | (11.522) | (17.601) | (24.587) | (0.549) | (0.690) | (1.574) | (2.610) |
| Observations | 157 | 113 | 157 | 91 | 195 | 141 | 195 | 106 |
| Clusters | 26 | 21 | 26 | 21 | 33 | 26 | 33 | 26 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 39.31 | 75.00 | 50.00 | 75.00 | 34.93 |
| Mean Dep. Var. | 146.994 | 143.442 | 146.994 | 141.253 | 6.600 | 6.525 | 6.600 | 5.877 |
| SD Dep. Var. | 65.585 | 60.761 | 65.585 | 60.249 | 5.979 | 5.970 | 5.979 | 5.121 |
| Sample | No Karnataka | No Karnataka | No Karnataka | No Karnataka | Complete | Complete | Complete | Complete |

| | *Police Stations in 1961* | | | | *All educational institutions in 2016* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | 0.152 | -0.224 | -0.177 | -0.179 | -0.680 | 1.675 | 3.879 | 3.924 |
| | (0.372) | (0.464) | (0.544) | (1.555) | (1.629) | (1.714) | (2.877) | (3.128) |
| Observations | 195 | 141 | 195 | 79 | 563 | 389 | 563 | 468 |
| Clusters | 33 | 26 | 33 | 26 | 559 | 385 | 559 | 464 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 24.00 | 75.00 | 50.00 | 75.00 | 62.77 |
| Mean Dep. Var. | 3.487 | 3.525 | 3.487 | 3.089 | 21.665 | 22.121 | 21.790 | 22.878 |
| SD Dep. Var. | 2.871 | 2.963 | 2.871 | 2.709 | 23.578 | 23.079 | 23.744 | 24.698 |
| Sample | Complete | Complete | Complete | Complete | Complete | Complete | Complete | Complete |

For 2016, the dependent variable is number of educational facilities at the block level as per public data from Ministry of rural development, India (https://geosadak-pmgsy.nic.in/OpenData). For 1961, observations are at the block level and compiled from district census of each district. Standard errors are clustered at the district level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. RD regressions include a local linear specification estimated separately on each side of the boundary and uses a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications include fixed effects for boundary segment. Robust standard errors are in parenthesis. * p < 0.10, ** p <0.05, *** p<0.01.

Table E-7: Roads and Hospitals in 2001

| | *Paved approach road* | | | | *Hospitals* | | | |
|---|---|---|---|---|---|---|---|---|
| | Linear polynomial | | Local Linear RD | | Linear polynomial | | Local Linear RD | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *Hyderabad* | -0.030* | -0.038** | -0.062*** | -0.061*** | 0.004 | 0.007** | 0.011*** | 0.010* |
| | (0.017) | (0.019) | (0.020) | (0.023) | (0.003) | (0.003) | (0.004) | (0.006) |
| Observations | 27,464 | 19,751 | 27,464 | 14,841 | 27,470 | 19,754 | 27,470 | 11,639 |
| Clusters | 612 | 447 | 612 | 332 | 612 | 447 | 612 | 278 |
| Bandwidth | 75.00 | 50.00 | 75.00 | 35.35 | 75.00 | 50.00 | 75.00 | 26.16 |
| Mean Dep. Var. | 0.796 | 0.794 | 0.796 | 0.792 | 0.023 | 0.021 | 0.023 | 0.021 |
| SD Dep. Var. | 0.403 | 0.404 | 0.403 | 0.406 | 0.150 | 0.144 | 0.150 | 0.145 |

The data is taken from Asher et al. (2021). Standard errors are clustered at the subdistrict level. Linear polynomial regressions include a linear polynomial of cluster latitude, longitude. RD regressions include a local linear specification estimated separately on each side of the boundary and uses a triangular kernel. Column 3 for each outcome uses a fixed bandwidth of 75 km and column 4 shows results for an optimal bandwidth chosen using the MSE minimizing procedure of Cattaneo et al. (2020). All specifications include fixed effects for boundary segment and control for total population. Robust standard errors are in parenthesis. * p < 0.10, ** p <0.05, *** p<0.01.

## E.7 Women's education estimated for two different geographic granularity

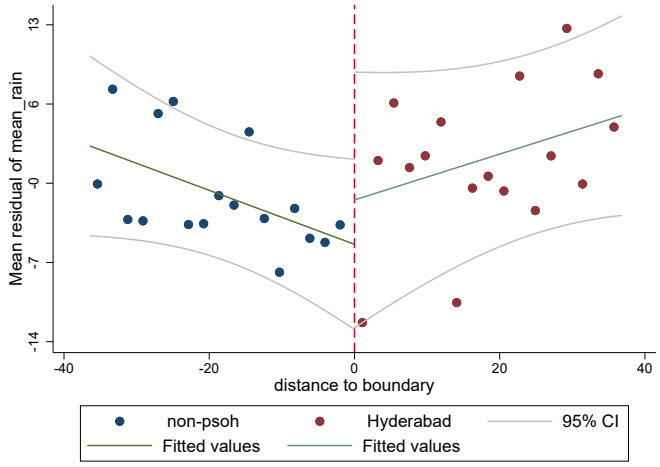
Table E-8: Years of education of women in 2015

| | None | cubic polynomial in distance | | latitude and longitude | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | none | cluster | district | cluster | district |
| psoh | -0.595*** | -0.614*** | -0.629*** | -0.576*** | -0.561*** |
| | (0.153) | (0.153) | (0.153) | (0.154) | (0.154) |
| $N$ | 30816 | 30816 | 30816 | 30816 | 30816 |
| $R^2$ | 0.220 | 0.223 | 0.222 | 0.221 | 0.222 |
| geographic_observations | | 1372 | 36 | 1372 | 36 |

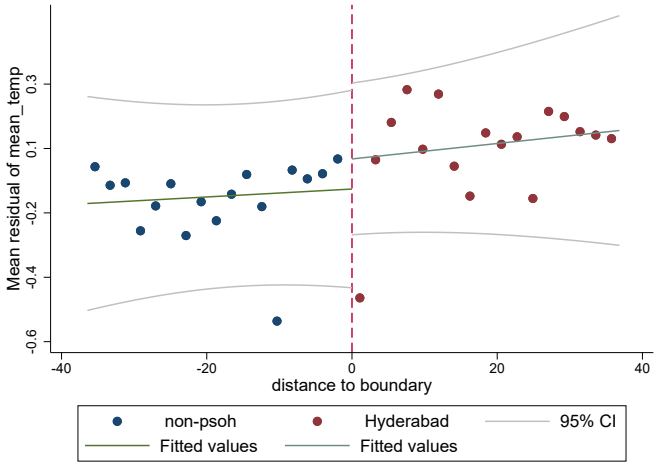
The dependent variable is years of education of women. All specifications control for mean altitude of the cluster and age of women. The cluster columns control for a geographic polynomial of the centroid of the sample cluster and the district column estimates the same specification except it uses a function of the latitude and longitude of the district. We include fixed effects for boundary segment and use sample weights provided in the data. Errors are clustered at the district level. Clustered standard errors are in parenthesis. * p $<$ 0.10, ** p $<$0.05, *** p$<$0.01.

# F RD Plots of balance variables

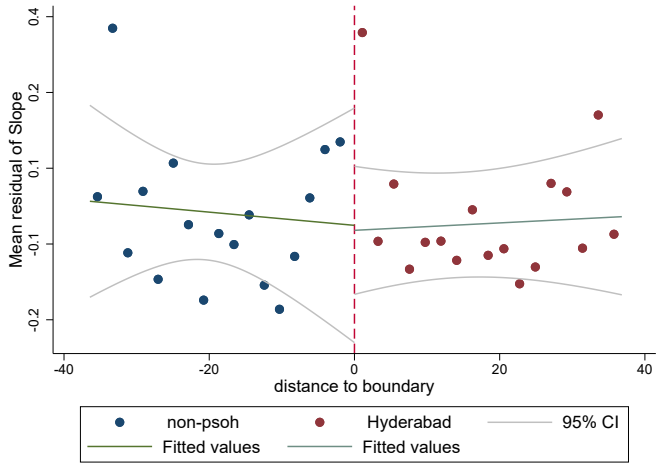
Figure F-1: RD plots of balance variables



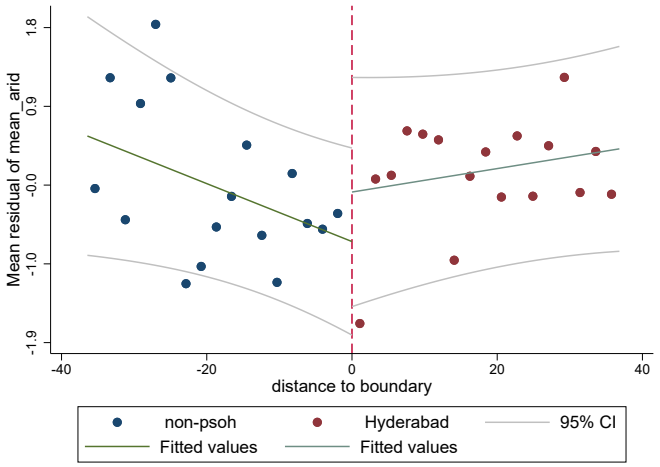

(a) Mean rainfall



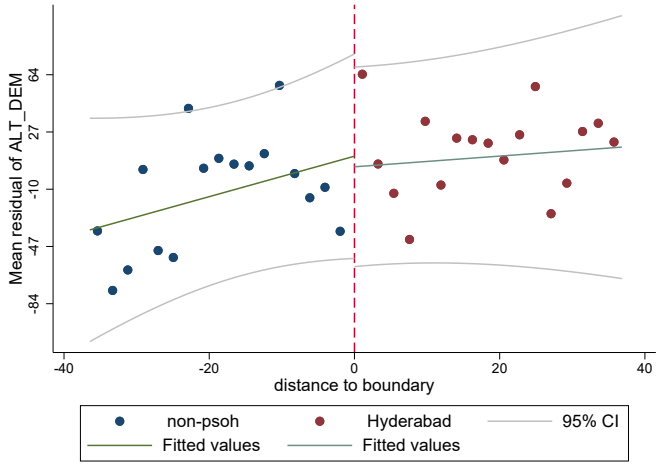

(b) Mean temperature



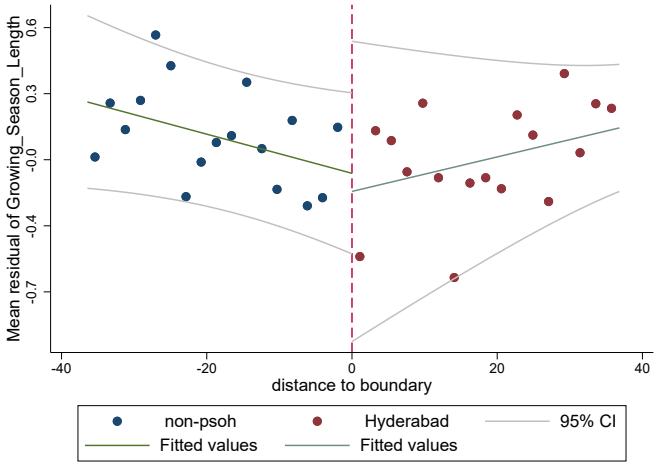

(c) Slope




(d) Aridity



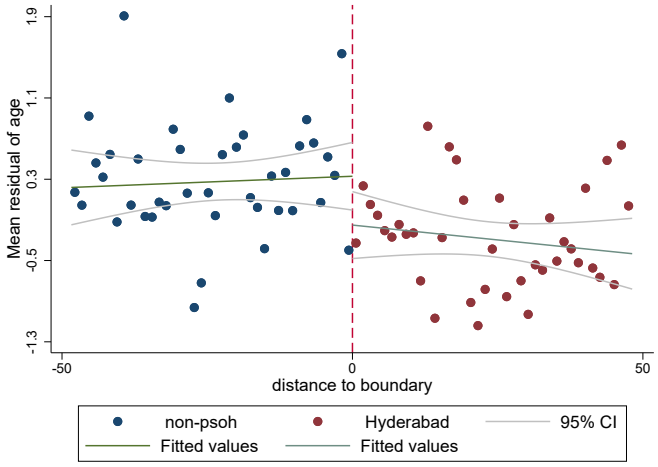

(e) Altitude



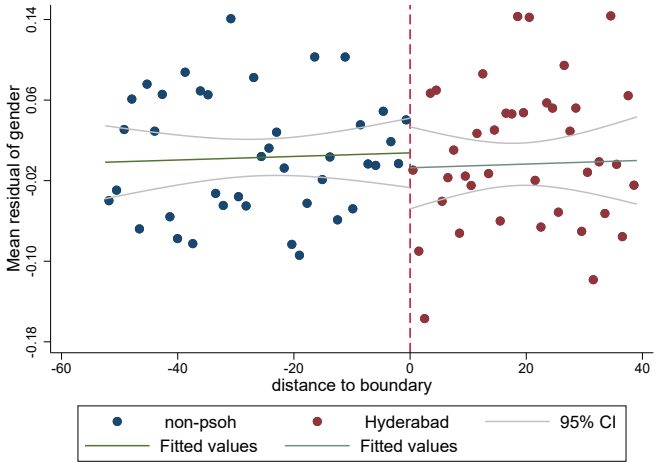

(f) Growing season length

Notes: The figure presents RD plots for our balance variables and the mean value of each outcome variable at each 1 km bin along the running variable with a local linear trend estimated separately on each side of the discontinuity. Regressions controls for boundary segment fixed effect. Standard errors are clustered at the cluster level and the figures show 95% confidence intervals.

Figure F-2: RD plots of covariates




(a) Age

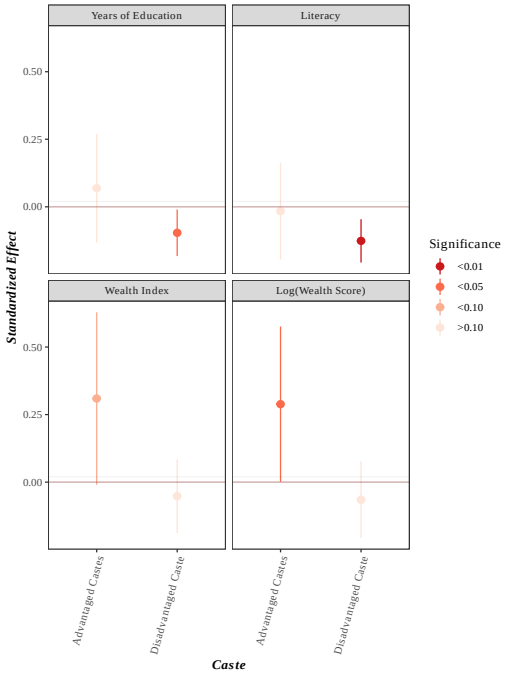



(b) gender

# G Other princely states

## G.1 Mysore

Figure G-1: Standardized estimates for princely state of Mysore




Notes: The figure presents standardized RD coefficients for our main outcomes for difference caste groups in *Mysore*. Regressions control for age, age squared , household size and gender. We use the optimal bandwidth for estimation.

# H Selective migration

Differential migration of disadvantaged caste individuals could be a possible channel of persistence if the most capable disadvantaged caste individuals migrated to either direct rule areas or to internal parts of *Hyderabad* which are not in our sample. Appendix Table H.1 shows that migration in 1911 was small. However, differential migration post independence might be consistent with all our results. In this section we conduct two exercises which suggest that differential migration cannot not explain all our results.

Firstly, we conduct a trimming exercise. Assuming that the most well-off disadvantaged group people in indirectly ruled areas migrated to directly ruled areas, we estimate the percentage of individuals which would have to migrate for our estimates to disappear. For this we restrict our sample to the disadvantaged caste group which lie within 75 km of the boundary. Then we remove the top x% individuals/households on the direct rule side of the sample. Figure H-1 suggests that this percentage would have to be somewhere between 15-25%. This would imply that for selective migration to fully explain the results, only the "best" disadvantaged caste people from indirect rule areas are leaving and that the top one-fifth of the individuals we observe in directly ruled areas came from indirectly ruled areas. This is a very high percentage of differential migration, and is unlikely given the fact that we observe no difference in total population in contemporary data. Another possibility which might be consistent with our results is that the poorest individuals in the direct rules migrate and settle in indirect rule areas, thus bringing down the average education and wealth level in the area. We conduct another trimming test, where we remove the bottom x% households on the indirect rule side of the sample. For this form of selective migration to fully explain the results, around 10-15% of the poorest population in indirect rule areas must have migrated from direct rule areas (see Figure H-2). This level of differntial migration is unlikely.

Secondly, we look at possible migration to other parts of the Hyderabad state. For a moment, let us assume that disadvantaged caste individuals near the boundary migrated to other parts of the state which are not in our sample and this migration explains all our results. In that case, if we extend the boundary to those areas, we should see no difference. In figure H-3 we show results after we extend our sample to 200 km from the boundary, which includes the entire state of Hyderabad. The results still remain strong and significant for all outcomes.

As we use data from DHS survey, it might be possible that some individuals were not present in the house at the time of the survey and might have gone out of town for work. If the pattern of migration were such that the most educated individuals were more likely to migrate from areas in *Hyderabad*, then our estimate is larger than what it should be. We utilize a characteristic of work-related migration in India, its male-dominated nature, to test whether work-related migration is a concern. If our estimate is due to the selective migration of best disadvantaged caste individuals in *Hyderabad* then due to the male-dominated nature of migration, the effect should be higher for men and lower for women. In Figure H-4 we decompose the effect size for disadvantaged caste by gender. The figure shows that the effect is similar for men and women, and if anything it is larger for women. Therefore work related migration cannot explain our results.

## H.1 Migration in 1911

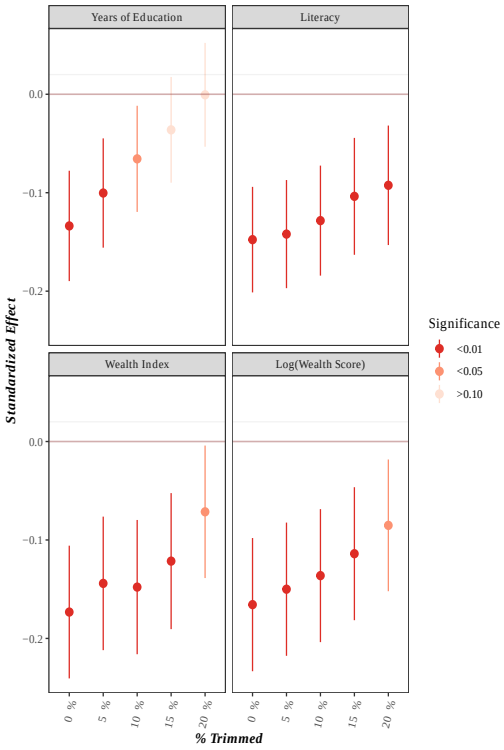
Table H-1: Migration in 1911

| | (1) Hyderabad | (2) Bombay | (3) Madras | (4) Central Provinces |
|---|---|---|---|---|
| Hyderabad (total migration) | . | 140990 | 60692 | 92731 |
| Bombay (total migration) | 118653 | . | 17304 | 86003 |
| Madras (total migration) | 67821 | 35484 | . | 10219 |
| Central Provinces (total migration) | 20945 | 34764 | 6291 | . |

This table shows the number of individuals born in a state or presidency(rows) who are residing in a neighbouring state(columns) in 1911. The column names indicate the place of residence in 1911 and row names indicate place of origin.

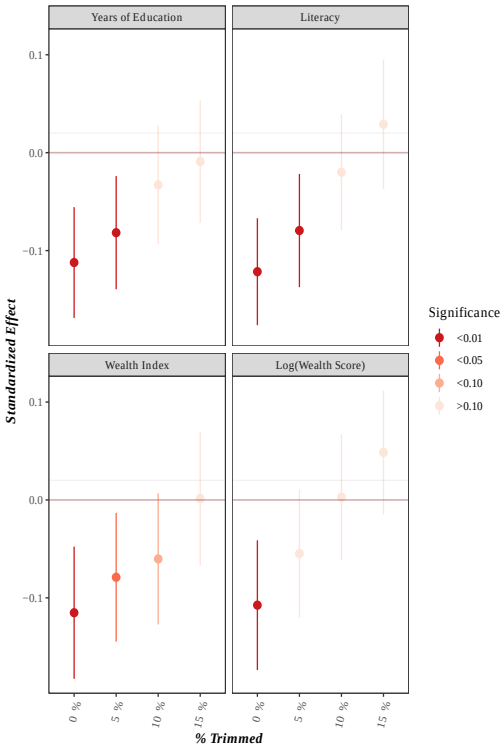
## H.2 Trimming for selective migration

Figure H-1: Results after trimming for selective migration




Notes: We restrict our sample to disadvantaged caste and then trim it by dropping the top x% of observations which lie within 75 km of the boundary under direct rule and then estimate our specification for each dependent variable. All regressions include a local linear specification within 75 km of the boundary and uses a triangular kernel. Regressions control for age, age squared, gender, and boundary segment fixed effects. Standard errors are clustered at the NFHS cluster level. Figure plots 95% confidence intervals.

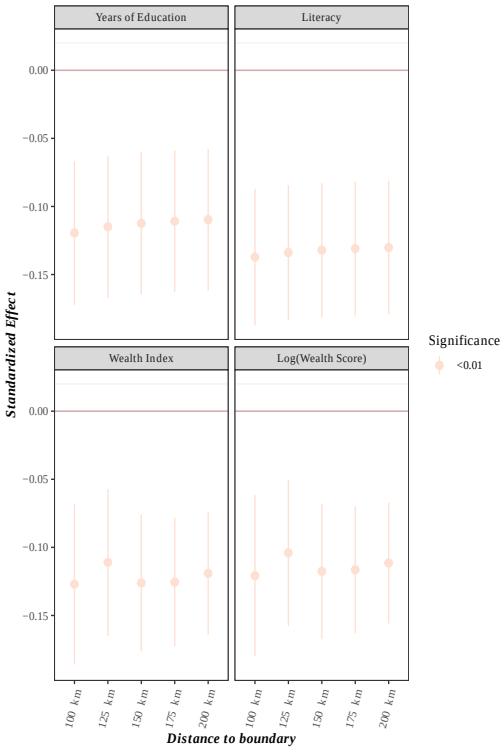
Figure H-2: Results after trimming for selective migration




Notes: We restrict our sample to disadvantaged caste and then trim it by dropping the bottom x% of observations which lie within 75 km of the boundary under indirect rule and then estimate our specification for each dependent variable. All regressions include a local linear specification within 75 km of the boundary and uses a triangular kernel. Regressions control for age, age squared, gender, and boundary segment fixed effects. Standard errors are clustered at the NFHS cluster level. Figure plots 95% confidence intervals.

## H.3 Migration to other parts of the Hyderabad

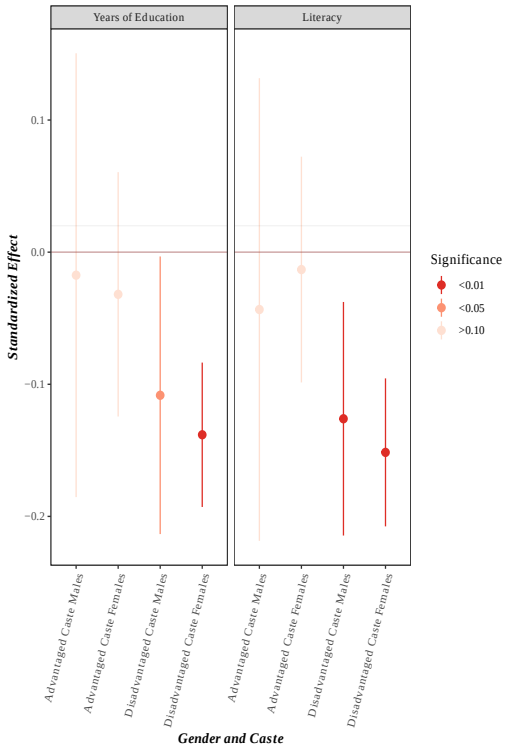
Figure H-3: Results after extending our sample to 200 km




Notes: We restrict our sample to disadvantaged caste and extend it till 200 km and then estimate our specification for each dependent variable. All regressions include a local linear specification and uses a triangular kernel. Regressions control for age, age squared, gender, and boundary segment fixed effects. Standard errors are clustered at the NFHS cluster level. Figure plots 95% confidence intervals.

## H.4 Estimates for disadvantaged castes by gender

Figure H-4: Results by gender




Notes: We restrict our sample to 50 km within the boundary and then decompose the effect by gender and caste. All regressions include a local linear specification and uses a triangular kernel. Regressions control for age, age squared, gender, and boundary segment fixed effects. Standard errors are clustered at the NFHS cluster level. Figure plots 95% confidence intervals.

# I A comment on using IVs in persistence regressions

The primary motive of most colonial powers that acquire land is extraction and oppression. Therefore, when comparing areas with different types of colonial institutions, concerns arise regarding selection bias on the productive characteristics of the annexed area. Previous studies in the literature on persistence (Acemoglu et al. (2001), Iyer (2010), Banerjee and Iyer (2005) and others) have utilized instrumental variables (IVs) to address this selection bias. While the instrument may be exogenous, estimating the effects of colonial institutions on contemporary outcomes may produce inaccurate results if systematic differences in institutions that emerged later are not taken into account.

Let Equation (2) be the specification we want to estimate, where $i$ indicates a district/county. As there are concerns regarding selection, the author may instrument for $colonial$, our independent variable, with $Z$. For the exclusion restriction to be satisfied, we need that: $\mathbf{E}[\epsilon_i Z_i | X_i] = 0$. Let us assume that the dependent variable $y_i$, and therefore $\epsilon_i$ is taken after the former colonial state became independent. Let district $i$ be in state $s_i$ of the independent country. Therefore, outcomes in district $i$ are going to be affected by (policies of the) state $s_i$ of which the district $i$ is a part after independence. This is formally stated in equation (3). For the exclusion restriction to hold, it must be that $\mathbf{E}[(\nu_i + s_i)Z_i] = 0$. If there are infinitely many states, and instrument $Z_i$ was randomly assigned before independence then this restriction should be satisfied. In the infinite sample, the overlap condition would also hold for all states. That is each state would have a district $i$, such that $colonial_i = 1, Z_i = 1$ and a district $j$, such that $colonial_j = 0, Z_j = 0$. However, in the finite sample not all states may have a district in which the instrument satisfies the overlap condition. To get the correct estimates, we should restrict the sample to states where the overlap condition holds.

$$y_i = \alpha + \beta colonial_i + \gamma X_i + \epsilon_i \tag{I.1}$$

$$y_i = \alpha + \beta colonial_i + \gamma X_i + s_i + \nu_i \tag{I.2}$$

To make the point more salient, let us take the example of Iyer (2010). To address concerns regarding selection in areas that the British took over, Iyer (2010) uses the British policy of *Doctrine of Lapse*, under which the British took over princely states where the ruler died without a natural heir, as an instrument. Iyer (2010) argues that since the death of a ruler without a natural heir depends on natural circumstances, the instrument satisfies the exclusion restriction. Figure I-1 (b) clearly shows that the lapsed districts are located only in four states, and hence the first stage can be valid only in these four states. Therefore, the correct estimation strategy should include only those states where a *lapsed* district was overtaken by

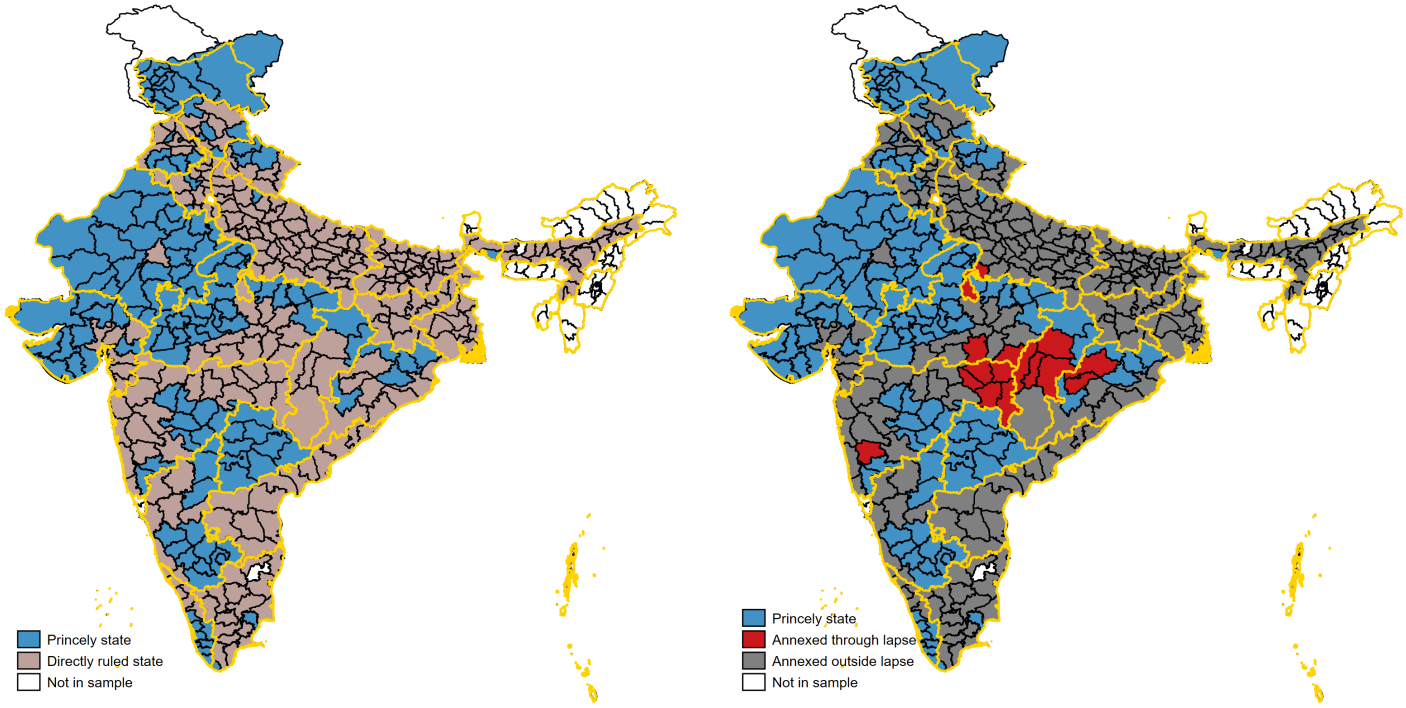
the british (where the overlap condition holds).

To see how the results differ we reanalyze Iyer (2010) using the alternate specification. Results are in Table I-1. In column2 of I-1 we restrict the sample to those Indian states where the britishers annexed a princely state which had *lapsed*. This reduces our sample to the states of Madhya Pradesh (including Chhattisgarh), Uttar Pradesh, Maharasthra and Odisha. The alternate specification suggests that the picture on direct and indirect rule is not as clear as Iyer (2010)'s main specification would suggest. Firstly, districts under British seem to be doing better in terms of area irrigated, yield, canals and literacy. The log of total yield of 15 different crops, log rice yield are larger (p value of 0.147 and 0.104 respectively). Secondly, while Iyer (2010) finds large and significant positive effects of indirect rule on public goods, health and poverty, the alternate specifications suggest a different story. For public goods, a majority of the results still suggest that indirect rule districts have more public goods, though the estimates are much smaller in size and most of them are not significant. For health and poverty, the estimates are much smaller and not statistically significant. Therefore, caution must be exercised in interpreting results in the persistence literature when there are systematic differences in institutions that emerged later but are not accounted for in the empirical specification.

Table I-1: Iyer 2010

| | *Main specification* (1) | *Restricting to Lapse states* (2) |
|---|---|---|
| Proportion of area irrigated | 0.059 | 0.174* |
| | (0.065) | (0.093) |
| Fertilizer usage | -3.145 | 4.507 |
| | (3.765) | (4.125) |
| Proportion of area sown with HYV | 0.051 | -0.000 |
| | (0.038) | (0.056) |
| Log total yield | 0.087 | 0.266 |
| | (0.127) | (0.161) |
| Log rice yield | -0.107 | 0.308 |
| | (0.215) | (0.183) |
| Log wheat yield | -0.184 | 0.002 |
| | (0.243) | (0.212) |
| Primary school | -0.011 | -0.006 |
| | (0.041) | (0.047) |
| Middle school | -0.091** | -0.053 |
| | (0.037) | (0.035) |
| High school | -0.065 | -0.035* |
| | (0.042) | (0.017) |
| Primary health center | -0.031** | -0.009 |
| | (0.013) | (0.006) |
| Primary health subcenter | -0.053** | -0.017 |
| | (0.021) | (0.022) |
| Canals | -0.043 | 0.010** |
| | (0.028) | (0.005) |
| Roads | -0.198*** | -0.033 |
| | (0.066) | (0.047) |
| Literacy rate (1961, 1971, 1981, 1991) | 0.014 | 0.046 |
| | (0.043) | (0.029) |
| Infant mortality rate 1981 | 37.349** | 10.570 |
| | (14.197) | (13.670) |
| Infant mortality rate 1991 | 26.869** | 2.691 |
| | (10.544) | (10.827) |
| Poverty 1983, 1987, 1993 | 0.161*** | 0.044 |
| | (0.054) | (0.055) |
| Inequality 1983, 1987, 1993 | -0.063*** | -0.056** |
| | (0.018) | (0.021) |

Figure I-1: Direct and indirect rule areas in India



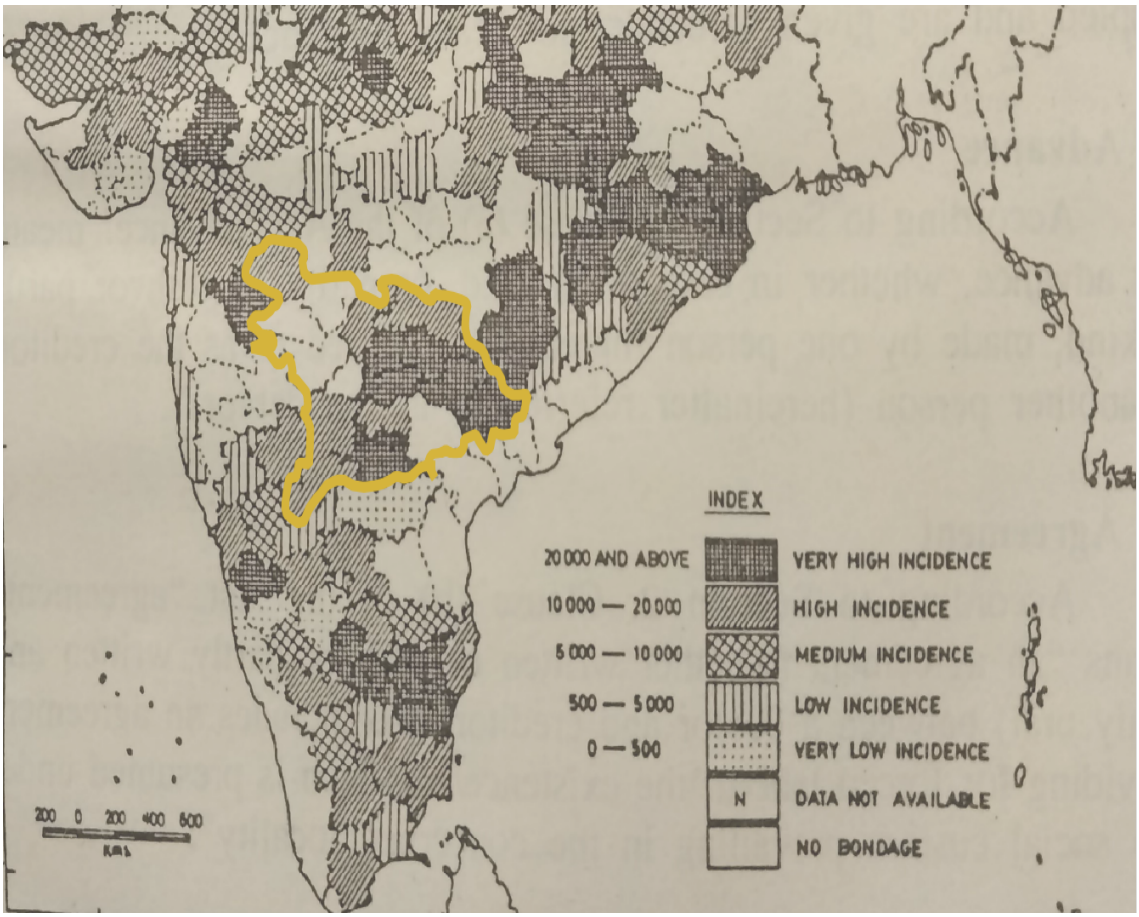

(a) Sample of Iyer (2010) for OLS

(b) Sample of Iyer (2010) for IV regression

Notes: The map shows the districts which are in the sample of Iyer (2010)

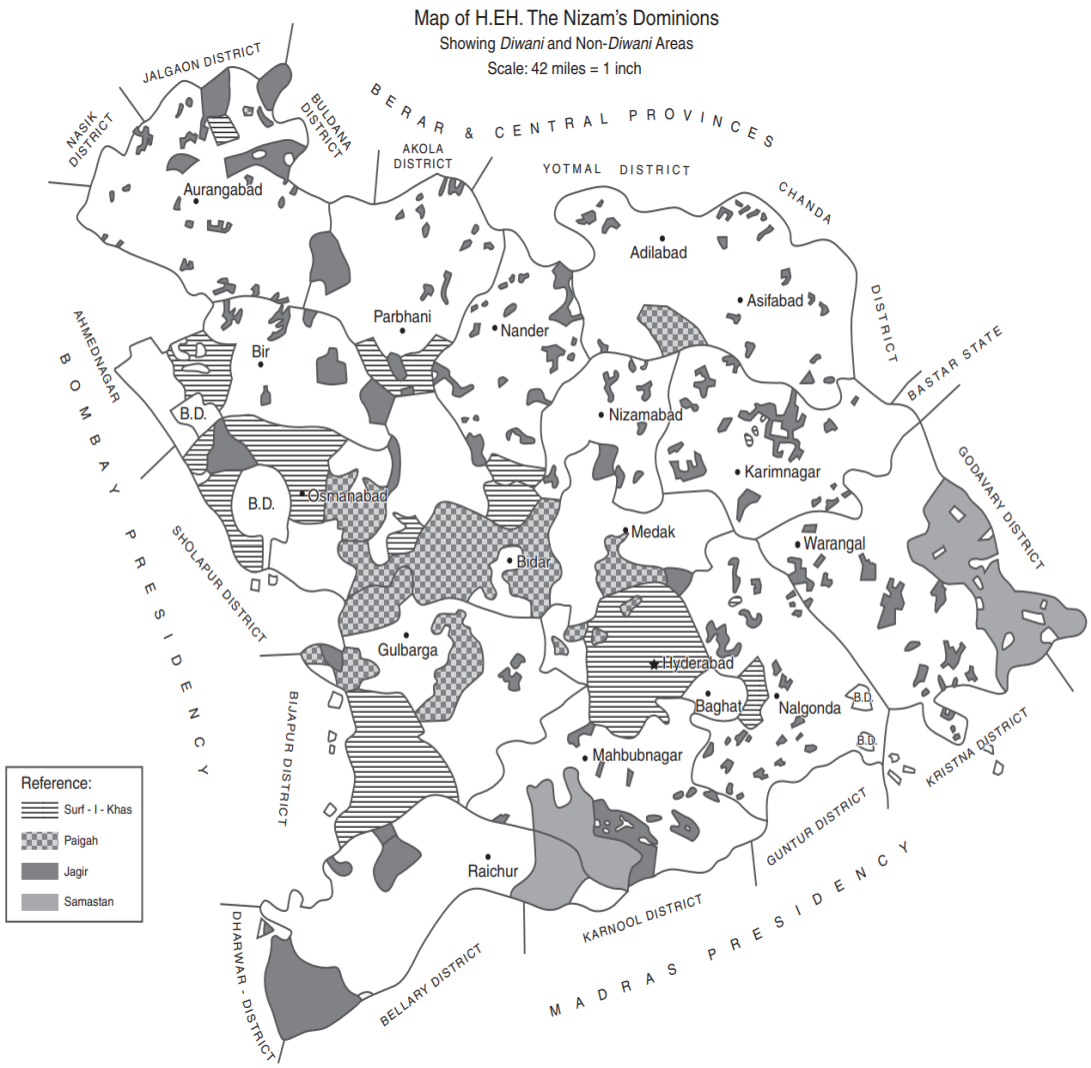
# J Maps

Figure J-1: Bonded Labor in India in 1979



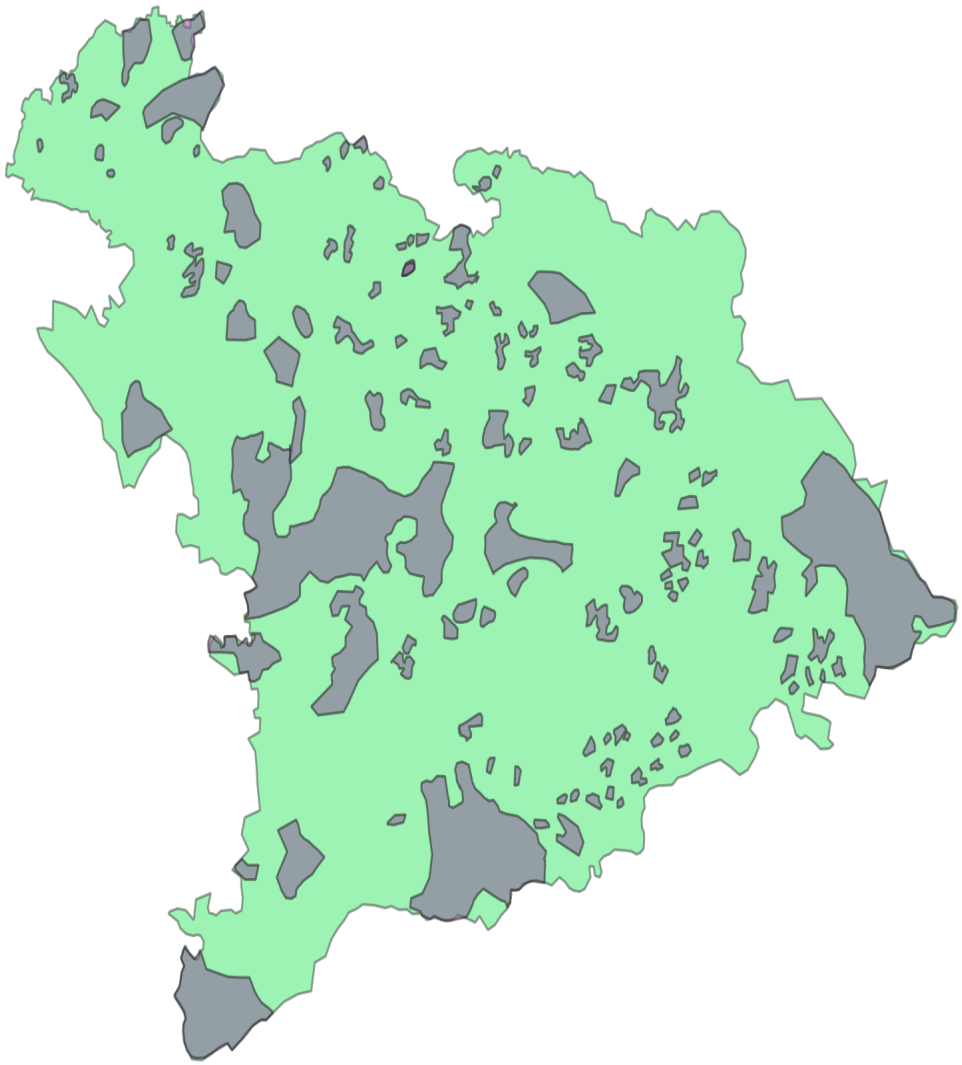

Notes: The map shows the incidence of bonded labor in India in 1979. The boundary of the state of *Hyderabad* is in brown. Source Gandhi Peace (1979).

Figure J-2: Map of *Hyderabad* showing Jagirs




Notes: The map shows the administrative divisions in *Hyderabad*. Source Mukherjee (2021)

Figure J-3: Map of *Hyderabad* showing labor-repressive areas



Notes: The map shows the state of *Hyderabad*. Labor-repressive areas are colored in gray.