

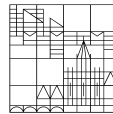
**Three Essays on the Economics of Government
Intervention: Education, Jurisdiction, and Infrastructure**

**Doctoral thesis for obtaining the
academic degree Doctor of
Economics**

submitted by
Naghavi, Arash

at the

Universität
Konstanz



Faculty of Politics, Law and Economics

Department of Economics

Konstanz, 2018

Date of the oral examination: 2 July 2018

1. Reviewer: Professor Dr. Heinrich Ursprung

2. Reviewer: Professor Dr. Guido Schwerdt

3. Reviewer: Professor Dr. Michael Becher

Contents

Acknowledgements	i
Summary	ii
Zusammenfassung	iv
Author Contribution	vi
1 Education reform and Education Gaps	1
1.1. Introduction	1
1.2. Data	2
1.3. Empirical Strategy	2
1.4. Results	3
1.5. Conclusion	4
References	6
2 Colonial Indirect Rule and Pre-colonial Ethnic Institutions: Evidence from British India	7
2.1. Introduction	7
2.2. Related Literature	10
2.3. Historical and Institutional Background	11
2.4. Data	13
2.5. Empirical Strategy	16
2.6. Results	19
2.7. Cultural legacy of the FCR	24
2.8. Conclusion	26
Reference	28
3 Was Railroad Expansion a Success Story of the British Colonial Rule in India? A Long-Term Perspective	31
3.1. Introduction	31
3.2. Related Literature	33
3.3. Historical background	34
3.4. Data	36
3.5. Empirical Strategy and Estimation Results	39

3.6. Conclusion.....	44
Reference.....	44
Bibliography.....	48

Acknowledgements

To my wife, Mansa Sabaghian, because I owe it all to you. Many Thanks for all your love, constant support, and patience.

I would like to express the deepest appreciation to my supervisor, Prof. Heinrich W. Ursprung, for affording me the opportunity to freely choose the topics of my research, for continually and unhesitatingly supporting me not only in my academic carrier but also in my personal life, and for being instrumental in helping me to develop my approach to economic analysis and academic writing. Without his guidance and persistent help, this dissertation would not have been possible.

I would like to thank my committee members, Prof. Guido Schwerdt and Prof. Michael Becher, for their insightful comments, encouragement, and the hard questions which motivated me to develop my research.

I would also like to thank Prof. Michael Becher and Prof. Susanne Goldlücke for the opportunity they gave me to work as a research/teaching assistant.

I am grateful to my life-coaches, my parents who provided me both moral and emotional support in my life. I am also grateful to my parents-in-law for supporting my family.

Finally, I am also grateful to my friends and fellow PhD students, especially to Masoud Farhadi, Daniela Behr, Katarina Zigova, Emilia Oljemark, Jana Mareckova, Philip Heiler, and Tommy Krieger, for their help and keeping my spirits up during the last four years.

Summary

Institutions are a major determinant of vast differences in total factor productivity, human capital, and physical capital across and within countries today. My dissertation consists of three stand-alone empirical studies on the economic effects of three institutional designs: the Italian 2009 “Gelmini” education reform (chapter 1), indirect ruling system in the north-west of British India (chapter 2), and railroad construction in British India (chapter 3). The institutions are not exogenous and change in tandem with other determinants of economic development. To address the endogeneity issue, I apply three microeconomic techniques to estimate the causal effect of the institutions: a difference-in-difference (chapter 1), a regression discontinuity (chapter 2), and an instrumental variable (chapter 3) identification strategy. In the following I briefly summarize the main features and results of each study.

Chapter 1 is a reprint of a joint article with Babak Jahanshahi (Università degli studi di Pavia). The article *Education Reform and Education Gaps* appeared in the *Applied Economics Letters*, 24(19), 1385-1388. In this article, we estimate the causal effect of the Italian 2009 “Gelmini” education reform on three academic performance gaps. The reform mainly consisted of reducing the number of teachers and daily instruction time. Using lags in implementing the reform across grades to specify a difference-in-difference identification strategy, we find that the reform had a statistically and economically significant effect on the immigrant-native gap and on the gender gap. The reform increased the gap between immigrant and native students and changed the gender gap in both mathematics and Italian language skills to the advantage of boys. We, however, did not find a statistically significant effect of the reform on the gap between students with low-skilled parents and students with high-skilled parents.

Chapter 2, entitled *Colonial Indirect Rule and Pre-colonial Ethnic Institutions: Evidence from British India*, is a joint study with Shujaat Farooq (Pakistan Institute of Development Economics). In this study, we contribute to a growing literature on historical determinants of comparative development by examining the long-term economic effect of institutional design in the north-west of British India (present-day Pakistan). Enacting the Punjab Frontier Regulation in 1872 that integrated a pre-colonial institution, called Jirga or council of elders, into colonial administration, the British colonial state indirectly ruled the northwest of colonial India. This colonial policy regulated and administrated justice in cooperation with local elites and transferred Jirga from a council of all adult male members of a community to a council composed of three or more state-appointed members of local elites. We examine the long-term effect of this colonial policy on development and public goods provision. Using a regression discontinuity design across the boundary that separated indirectly from directly ruled areas, we find that the areas exposed to indirect rule are less developed and have inferior access to those public goods which began to be provided in the colonial era. We conclude that the divergence started in that

time and persists today and show that one mechanism underlying the long-term effect is the effect of the indirect rule on social norms, specifically interpersonal trust and corruption.

Chapter 3, entitled Was Railroad Expansion a Success Story of the British Colonial Rule in India? A Long-Term Perspective, is a joint study with Amin Z. Ashtiani (Luiss Guido Carli University). In this project, we investigate lasting economic effect of access to railroads constructed by the colonial British state in India, focusing on the Northwest of the colony (present-day Pakistan). Railroad construction in British India began in 1853. It expanded rapidly in the second half of the 19th century and by the 1900s became the fourth largest railway network in the world. It had far-reaching economic effect at the time. Using both OLS and instrumental variable specifications, we find a statistically and economically significant lasting effect of railroad access on local development. Literacy rate, and schooling attainment, are higher in villages that are closer to the railroad network. Access to railroads also improved living standard, measured by access to piped water, access to electricity, and housing quality.

Zusammenfassung

Institutionen sind ein wesentlicher Faktor für beträchtliche Unterschiede in der totalen Faktorproduktivität, im Human- und Sachkapital zwischen und innerhalb vieler Länder. Meine Dissertation besteht aus drei eigenständigen, empirischen Studien zu den ökonomischen Auswirkungen von drei verschiedenen Institutionen: der italienischen Bildungsreform "Gelmini" (Kapitel 1), dem System der sogenannten indirect rule im Nordwesten von Britisch-Indien (Kapitel 2) und dem Eisenbahnbau in Britisch-Indien (Kapitel 3). Institutionen sind nicht exogen, sondern verändern sich fortlaufend, insbesondere im Zusammenspiel mit anderen Determinanten der wirtschaftlichen Entwicklung. Um das Problem der Endogenität zu adressieren, wende ich in den verschiedenen Kapiteln drei mikroökonomische Techniken an, um den kausalen Effekt der Institutionen zu bewerten: das Difference-in-Difference Design (Kapitel 1), die Regressions-Diskontinuitäts-Analyse (Kapitel 2) und den Instrumentalvariablen Ansatz (Kapitel 3). Im Folgenden fasse ich kurz die wichtigsten Merkmale und Ergebnisse dieser drei Studien zusammen.

Kapitel 1 ist der Abdruck des gemeinsamen Artikels *Education Reform and Education Gaps* mit Babak Jahanshahi (Università degli studi di Pavia), der in den *Applied Economics Letters*, 24 (19), 1385-1388 erschienen ist. In diesem Artikel schätzen wir den kausalen Effekt der italienischen Bildungsreform „Gelmini“ 2009 auf drei akademische Leistungslücken. Die „Gelmini“ Bildungsreform bestand hauptsächlich darin, die Anzahl der Lehrer zu reduzieren und die tägliche Unterrichtszeit zu verkürzen. Durch die zeitverzögerte Einführung der Reform über die verschiedenen Klassen hinweg, ergibt sich die Möglichkeit einer Difference-in-Difference-Identifizierungsstrategie. Wir stellen fest, dass die Reform statistisch und ökonomisch signifikante Auswirkungen auf die Einwanderungslücke und die geschlechtsspezifische Diskrepanz hat: die Reform vergrößerte die Kluft zwischen Einwanderern und einheimische Schülern und erhöht die Gender-Diskrepanz in den Bereichen Mathematik und Italienisch Sprachkenntnisse zum Vorteil von Jungen. Wir fanden jedoch keine statistisch signifikanten Auswirkungen der Reform auf die Kluft zwischen Schülern mit gering qualifizierten Eltern und hochqualifizierte Eltern.

Kapitel 2, das den Titel trägt *Colonial Indirect Rule and Pre-colonial Ethnic Institutions: Evidence from British India*, ist eine gemeinsame Studie mit Shujaat Farooq (Pakistan Institute of Development Economics). In dieser Studie tragen wir zu der wachsenden Literatur zu historischen Determinanten der vergleichenden Entwicklung bei. Wir untersuchen die langfristigen, ökonomischen Auswirkungen von Institutionen im Nordwesten von Britisch-Indien (dem heutigen Pakistan). Mit der Verabschiedung der sogenannten Punjab Frontier Regulation im Jahr 1872, die eine vorkoloniale Institution namens Jirga oder Ältestenrat in die Kolonialverwaltung integrierte, regierte der britische Kolonialstaat den Nordwesten des kolonialen Indiens unter sogenannter indirect rule. Diese

Kolonialpolitik regulierte und verwaltete die Justiz in Zusammenarbeit mit lokalen Eliten und änderte die Komposition von Jirga. Statt von allen erwachsenen, männlichen Mitgliedern einer Gemeinde zusammengesetzt, konstituierte sich der Rat nun aus drei oder mehr vom Staat ernannten Mitgliedern lokaler Eliten. Wir untersuchen die langfristigen Auswirkungen dieser Kolonialpolitik auf das Entwicklungsniveau und auf die öffentliche Güterversorgung. Wir führen ein Regressionsdiskontinuitätsdesigns unter Bezugnahme der Grenze zwischen indirekt von direkt regierten Gebieten durch. Dabei zeigen wir, dass Gebiete unter indirect rule weniger entwickelt sind und einen schlechteren Zugang zu jenen öffentlichen Gütern haben, die in der Kolonialzeit angeboten wurden. Wir schließen daraus, dass die Divergenz in dieser Zeit begann und bis heute andauert. Zudem zeigen wir Mechanismen auf, die diesem Effekt zu Grunde liegen: Die Auswirkung der indirect rule auf soziale Normen, insbesondere zwischenmenschliches Vertrauen und Korruption.

Kapitel 3 mit dem Titel *Was Railroad Expansion a Success Story of the British Colonial Rule in India? A Long-Term Perspective*, ist eine gemeinsame Studie mit Amin Z. Ashtiani (Universität Luiss Guido Carli). In diesem Projekt untersuchen wir die dauerhaften, ökonomischen Auswirkungen des Zugangs zu Eisenbahnlinien, die vom britischen Kolonialstaat in Indien gebaut wurden und sich auf den Nordwesten der Kolonie (das heutige Pakistan) konzentrieren. Der Eisenbahnbau in Britisch-Indien begann 1853. In der zweiten Hälfte des 19. Jahrhunderts expandierte er rasch und wurde im 20. Jahrhundert zum viertgrößten Eisenbahnnetz der Welt und hatte damals weitreichende wirtschaftliche Auswirkungen. Unter Verwendung sowohl von OLS- als auch von Instrumentalvariablen-Schätzungen finden wir einen statistischen und signifikanten Zusammenhang zwischen dem Eisenbahnzugang und lokaler Entwicklung; so sind Alphabetisierungsrate und Schulbildung in Dörfern, die näher am Eisenbahnnetz liegen, höher. Der Zugang zu Eisenbahnen steigert auch den Lebensstandard, gemessen am Zugang zu Leitungswasser, Elektrizität und Wohnqualität.

Author Contribution

Chapter 1 is a joint article with Babak Jahanshahi. I together with Babak developed the idea. I collected and cleaned data, performed analysis, and wrote the manuscript. Babak also helped in writing the manuscript and interpreting the results.

Chapter 2 is a joint study with Shujaat Farooq. I developed the idea. Shujaat obtained data. I cleaned data, performed analysis, and wrote the manuscript.

Chapter 3 is a joint study with Amin Z. Ashtiani. We developed the idea together. I cleaned data. We geolocated villages and railroads together. I performed analysis and wrote the manuscript.

Chapter 1

Education reform and Education Gaps

1.1. Introduction

Education reforms are often motivated by poor student achievements and budgetary considerations. Just so in Italy, where a need for cutting public spending and the poor performance of students in some international standardized tests (such as PISA, PIRLS, and TIMSS) led the Berlusconi government to propose in 2008 a reform of the primary school system. The reform stipulated a reduction in the number of teachers and an increase in the maximum (minimum) class size from 25 to 27 (10 to 15). The reform also abolished collaborative teaching and allowed parents to opt for a reduced weekly instruction time of 24 hours.¹ Starting in the school year 2009/10, collaborative teaching was abolished in all grades; the other reforms were implemented for the first grade, and subsequently phased in throughout all grades of primary school.²

We evaluate the causal effect of this reform on the academic achievement gaps between (i) native students and students with an immigration background, (ii) boys and girls, (iii) students from high and low social status families, and (iv) students with high- and low-educated parents. Exploiting the lags in implementing the reform for different grades, a difference-in-difference (DD) identification strategy is used. We focus on academic achievements in grades 2 and 5 in the year 2009/10 and 2010/11. In 2009/10, the reform affected neither of these grades, while in 2010/11 grade 2 was affected, whereas grade 5 was not.

In 2015, the *Research Institute for the Evaluation of Public Policies* (IRVAPP) investigated the implementation of the 2009 reform and found an 11% reduction in the number of teachers and a significant increase in class size; but not many parents appeared to have opted for the offered reduction in instruction time. This suggests that any of the reform's effects are likely to work through the increase in class size and the reduction in the number of teachers.

¹ Before the reform, the minimum instruction time was 27 hours/week.

² Decreto del Presidente della Repubblica 81/2009 and 89/2009.

Our study contributes to the literature on causal effects of education policies (Schlotter et al. 2011). The DD identification strategy is by now well established in education economics. Meghir and Palme (2005), for example, use the DD strategy to identify the causal effect of an education reform in Sweden. Our results, moreover, contribute to the literature exploring how academic performance differs across native and immigrant students (Contini 2013, Entorf and Minoiu 2005).

1.2. Data

We use data provided by the *Istituto Nazionale per la Valutazione del Sistema Dell'Istruzione* (INVALSI). The dataset contains the test scores of a nation-wide standardized test measuring academic achievements in mathematics and Italian language skills of second and fifth graders in primary schools. The dataset also records student specific demographic information that, in particular, shows no significant difference in the average share of immigrants in second grade and fifth grade in 2009/10 and 2010/11.

1.3. Empirical Strategy

Apart from abolishing collaborative teaching, the other reform items were not implemented for all grades at once. In the first year, these items were only implemented for grade 1, and were then phased in grade by grade in yearly steps. Second and fifth graders in 2009/10 were therefore subject to the old system, while in 2010/11 second graders were subject to the new system, whereas fifth graders remained subject to the old system. Because of this implementation lag, we can use a DD strategy to estimate the effect of the reform:

$$\Delta Y = (Y_{2,2010/11} - Y_{2,2009/10}) - (Y_{5,2010/11} - Y_{5,2009/10}) \quad (1)$$

where ΔY is the DD estimator of the entire student population and $Y_{g,t}$ denotes the academic achievement of pupils in grade g and school year t , measured by the average test scores in either mathematics or Italian language skills.

Since we cannot control for the difficulty of the tests that might differ across school years and grades,³ we standardize the test scores (zero mean and unit standard deviation) by subject, school year, and grade. When standardizing one can, of course, no longer estimate the effect of the reform on the accomplishments of the entire student population. Focusing on dichotomic subgroups of students, we can however estimate the relative causal effect of the reform on these subgroups. Let ΔY^i denote the DD estimator for subgroup i and let $-i$ denote the complement of subgroup i . Then,

³ The difference in the difficulty of the tests across school years and grades is elaborated upon in IRVAPP 2015.

$$\Delta Y_r = \Delta Y^i - \Delta Y^{-i} = \{(Y_{2,2010/11}^i - Y_{2,2010/11}^{-i}) - (Y_{2,2009/10}^i - Y_{2,2009/10}^{-i})\} - \{(Y_{5,2010/11}^i - Y_{5,2010/11}^{-i}) - (Y_{5,2009/10}^i - Y_{5,2009/10}^{-i})\} \quad (2)$$

where ΔY_r is the reform-induced change in the gap, i.e. the effect of the reform on the academic achievement gap between subgroups i and $-i$ in the treatment group (grade 2) *as compared to the gap in the control group* (grade 5). Applying a triple interaction term, ΔY_r is estimated with the help of the following regression:

$$y_{igt} = \alpha + \beta g_2 + \gamma t_{10} + \mu D_i + \delta(g_2 \cdot t_{10}) + \eta(g_2 \cdot D_i) + \theta(t_{10} \cdot D_i) + \varphi(g_2 \cdot t_{10} \cdot D_i) + X'_{igt} \cdot \tau + \varepsilon_{igt} \quad (3)$$

where y_{igt} is the standardized test score of pupil i in grade g at school year t , g_2 is a dummy for second graders, t_{10} is a time dummy switching on for observations in school year 2010/11, D_i indicates the respective partition of the population, and X'_{igt} is a vector of covariates. The parameter φ estimates ΔY_r , i.e. the reform-induced change in the gap. The coefficient of the triple interaction term can be interpreted to be causal under a fairly weak identifying assumption. The assumption only requires that, apart from the reform, no contemporaneous shock affects the grade-specific trends in the test scores of students in subgroups i relative to student in subgroups $-i$ (Gruber, 1994).

1.4. Results

Partitioning the student population, we estimated the effect of the reform on the academic achievement gap relating to immigration status, gender, parental social status, and parental education.⁴ Table 1 reports the estimates of the triple interaction term φ in equation 3, when controlling for gender, the share of immigrants in class, the share of girls in class, and parental education and socio-economic status. IRVAPP (2015) documented that the implementation of the reform varied across provinces. To control for potential heterogeneity across provinces, we add province fixed effect into equation 3 and report the regressions with province fixed effect in tables 1 and 2.

We find a statistically significant effect of the reform on the immigrant-native gap (table 1, panel A). Since in Italy (like almost everywhere) native students outperform immigrant students, the negative signs of the estimates in panel A imply that the reform-induced change increases the gap between immigrant and native students. In Italian language skills and mathematics, the reform-induced change in the gap amounts to approximately 7.5 percent and 9.5 percent of the standard deviation in the students' test scores. It thus transpires that because of the reform immigrant students lost substantial additional ground against their native fellow students. Moreover, we find a significant effect of the

⁴ Immigrant students are children of foreign-born non-Italian parents. The parental social-status gap measures the gap between students from low social-status families (unemployed and blue color workers) and other students. The parental education gap measures the gap between students from low-educated parents (less than high school diploma) and other students.

reform on the gender gap (table 1, panel B). Since in Italy, boys outperform girls in mathematics and girls outperform boys in Italian language skills, the results reported in panel B indicate that the reform-induced change in the gender gap is in both subjects to the advantage of the boys. In Italian language skills, boys caught up on the girls by 3.5 percent of the standard deviation, and in mathematics girls performed relatively worse by 9.5 percent. The results reported in panel C indicate a rather small change of less than 1.5 percent (to the advantage of students from high social status parents), and the results reported in panel D show that the reform did not affect the gap relating to parental education.

Students with an immigration background thus appear to be the main losers from the Gelmini reform. We therefore also estimated the immigrant-native gap for the potentially most vulnerable subgroups: girls, students with low-educated and low social-status parents. Table 2 shows the effect of the reform on the immigrant-native gap for these subgroups. Comparing the estimates for the general reform-induced change in the immigrant-native gap (table 1, panel A) with the estimates reported in table 3 indicates that the selected subgroups are indeed especially vulnerable: the absolute values of all estimates exceed the respective estimates when taking the entire group of immigrant students as a basic. The difference is especially large for immigrant girls in Italian language skills and for immigrant students with low-educated parents in both subjects.

To interpret the results as causal effect, one needs to assume (in addition to the usual DD identifying assumption) that the abolishment of collaborative teaching has the same marginal effect on the academic gaps for second and fifth graders. This boils down to the assumption that the different cohort-specific experiences with collaborative teaching have no long-term effect.

1.5. Conclusion

In this paper, we estimate the causal effect of the Italian 2009 primary school reform on (i) the academic achievement gap between Italian students and students with a migration background, (ii) the gender gap, and (iii) the gap between students with different household backgrounds in terms of parental education and social status. The reform reduced the number of teachers, increased class size, and abolished collaborative teaching. Using a difference-in-difference identification strategy and estimating the coefficient of a triple interaction term, we find that the reform disadvantaged girls as compared to boys and students with a migration background as compared to Italian students. The study thus suggests that even though the reform might have been successful in meeting the objective of reducing government spending, the reform has had adverse effects on education gaps.

Table 1: The effect of the reform on the academic achievement gaps

		1 Italian	2 Math
<i>A. Immigrant-native gap</i>			
	Reform effect	-0.076*** (0.012)	-0.095*** (0.012)
<i>B. Gender gap</i>			
	Reform effect	-0.035*** (0.006)	-0.063*** (0.006)
<i>C. Parental social status gap</i>			
	Reform effect	-0.012* (0.007)	-0.014** (0.007)
<i>D. Parental education gap</i>			
	Reform effect	-0.001 (0.006)	-0.005 (0.006)
Observations		1,480,829	1,463,807

Notes: Standard errors (clustered by schools) in parentheses, *=10%, **=5%, ***=1%. Source: INVALSI data for academic years 2009-10 and 2010-11.

Table 2: The effect of the reform on immigrant-native gap in academic achievement across various subgroups

		1 Italian	2 Math
<i>A. Girls</i>			
	Reform effect	-0.112*** (0.017)	-0.101*** (0.018)
	Observations	731,243	722,343
<i>B. Low-educated parents</i>			
	Reform effect	-0.115*** (0.021)	-0.123*** (0.020)
	Observations	529,346	522,685
<i>C. Low social status parents</i>			
	Reform effect	-0.084*** (0.018)	-0.103*** (0.017)
	Observations	409,466	404,363

Notes: Standard errors (clustered by schools) in parentheses, *=10%, **=5%, ***=1%. Source: INVALSI data for academic years 2009-10 and 2010-11.

References

- Contini, D. (2013). “*Immigrant background peer effects in Italian schools.*” *Social Science Research*, 42(4), 1122-1142.
- Entorf, H., and Minoiu, N. (2005). “*What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration.*” *German Economic Review*, 6(3), 355-376.
- Gruber, J. (1994). “*The incidence of mandated maternity benefits.*” *The American Economic Review*, 622-641.
- IRVAPP (2015). “*La riforma del maestro unico,*” *Il Mulino*, forthcoming.
- Meghir, C. and Palme M. (2005). “*Educational reform, ability, and family background.*” *American Economic Review*, 414-424.
- Schlotter, M., Schwerdt, G., and Woessmann, L. (2011). “*Econometric methods for causal evaluation of education policies and practices: a non-technical guide.*” *Education Economics*, 19(2), 109-137.

Chapter 2

Colonial Indirect Rule and Pre-colonial Ethnic Institutions: Evidence from British India

2.1. Introduction

In the past two decades, economists exploring historical determinants of comparative development have emphasized the long-term effects of colonial rule. Several studies focus on spatial variation in colonial institutions to show that the experience of colonial rule varies both across and within colonies⁵. We contribute to this growing literature by examining the long-term economic effect of institutional design in the north-west of British India (present-day Pakistan). Following the doctrine of indirect rule, the British colonial state instituted the Punjab Frontier Regulation in 1872 to give a substantial degree of legal autonomy to local elites in the north-west frontier districts. Neighboring districts, however, were ruled directly and enjoyed no legal autonomy. We investigate whether areas under direct and indirect rule experienced different development paths.

Two decades after annexing present-day Pakistan, the British colonizers acknowledged the importance of cooperation with local elites in the frontier districts to ensure public order. The colonial state enacted the Punjab Frontier Regulation in 1872 and the revised Frontier Crimes Regulation in 1887 and 1901 (henceforth FCR) to incorporate a pre-colonial institution, called Jirga or council of elders, into colonial administration. Jirga was an ethnic method of conflict resolution, essentially an egalitarian platform that allowed every adult male member of a community to participate, at least nominally, in the decision making process. The FCR adopted the ethnic method to supplement the formal legal codes that applied to all of British India, known as the Anglo-Indian Codes. According to the FCR, the colonial state could send both civil and criminal cases to an officially appointed Jirga to make a verdict. The

⁵ Studying the variation across colonies, for example, Acemoglu, Johnson, and Robinson (2001) focus on property rights institutions and La Porta et al. (1998) examine legal systems. Studying within colony variation, for instance, Banerjee and Iyer (2005) focus on different systems of collecting land revenue in British India, Dell (2010) studies a forced labor system instituted by the Spanish colonial state, and Iyer (2010) examines the lasting effect of indirect British rule compared with direct British rule in British India.

FCR-Jirga was ruled to consist of three or more state appointed members of the local elites (Nichols, 2013). The FCR, therefore, introduced a legalized and newly designed Jirga that empowered the local elites and remained unchanged until independence in 1947. The post-independence state of Pakistan abolished in the 1950s and 1960s in several steps the FCR in most areas except for some small tribal areas along the Afghan border (Berry, 1966). In the following, we refer to the FCR also as indirect rule.

The FCR split up the north-west of colonial India into two neighboring regions, both of which were administered by the colonial state but under two different legal codes. We use colonial records to identify to which districts the FCR applied and then geo-locate the FCR boundary that separated the indirectly ruled districts from those that were directly ruled. The FCR boundary (Figure 1) imposed a deterministic and discontinuous variation in the implementation of the FCR. We use a sharp Regression Discontinuity Design (RDD) to estimate the long-lasting effect of the FCR treatment on contemporary economic outcomes. To ensure validity of the RDD, we exclude three segments of the FCR boundary: the segment that lay on the border between the two provinces of Sind and Baluchistan⁶ (illustrated by a dashed black line in figure 1) and two segments that correspond to ethnic borders⁷ (illustrated by a solid black line in figure 1). We exclude these segments to focus exclusively on boundaries with identical ethnic distributions on both sides and boundaries that pass through the 1849 province of Punjab (illustrated by solid red lines in figure 1). We will show that also geographical characteristics and distances to the closest rivers vary smoothly at the boundary.

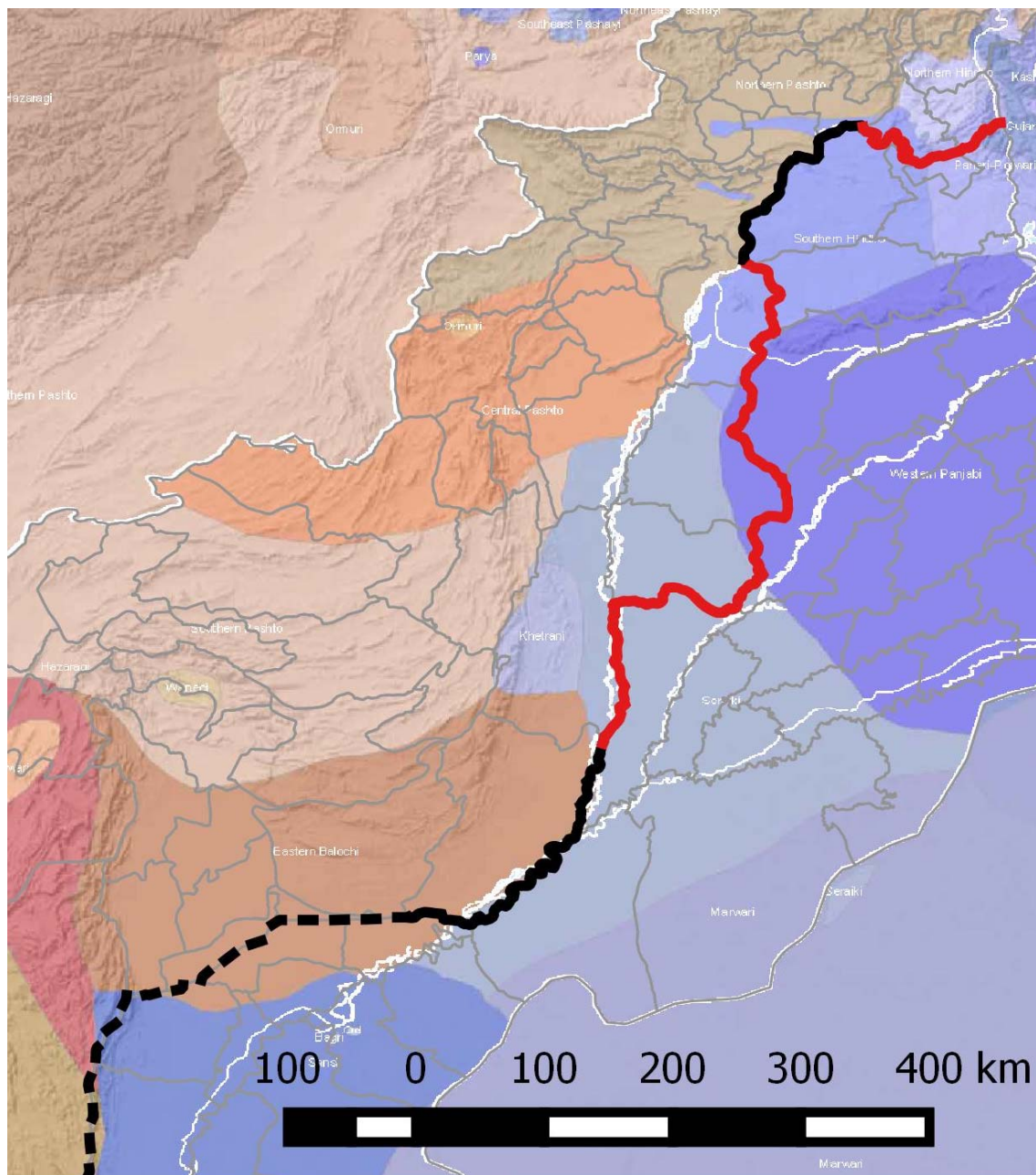
Using the 1998 Housing and Population Census and the 2008 Mouza Census aggregated at the village/city level, we estimate that the long-lasting effect of indirect rule decreased the literacy rate by about 5%. Indirect rule had also a negative effect on schooling attainment. We find that the indirectly ruled areas have inferior access to public goods as compared with the directly ruled areas. The villages on the indirect ruled side of the FCR boundary are on average 3 km further removed from a hospital and 1 km far from a road. The indirect rule has also had a negative effect on the availability of bricked streets and bricked drainage. We do, however, not find any significant effect on the provision of piped water and electricity. Our findings suggest that the FCR had a negative effect on the availability of those public goods which began to be provided in the colonial era and conclude that the divergence started in that time and persists today. We also show that the FCR did not harm economic prosperity measured by housing quality.

The FCR entitled the colonial authorities of each district to decide whether to refer a case to Jirga under the FCR or to colonial government courts under the Anglo-Indian Codes. The size and importance of tribes which applied Jirga and the geopolitical importance of the home district of these tribes could have influenced these decisions. Consequently, the ratio of cases dealt with under the FCR

⁶ We exclude this border because the FCR was not the only institutional difference between the two provinces and also because of a time-lag in the annexations of the two provinces. Sind and Baluchistan were annexed in 1843 and 1884, respectively.

⁷ One segment corresponds to the Punjabi-Pashtu border and one segment lies on the Punjabi-Baluchi border.

Figure 1: The FCR Boundary and Ethnic Distribution



Note: The map of the language distribution, shown in the background, is extracted from the maps of the Genetic Relationship of the World's Languages produced by Huffman (Huffman, 2015). The colors in the shades of red show Baluchi and Pashtu ethnic groups, and those colors in the shades of blue show Punjabi, Sindi, and Hindko ethnic groups. The gray lines are borders of Pakistani districts. The FCR boundary is plotted by dashed and solid black and solid red lines. The dashed black line is the Sind-Baluchistan border. The solid black lines are the ethnic borders. The solid red line is the portion of the boundary we exclusively focus on.

to the total number of civil and criminal cases could have varied across colonial districts that might have led to heterogeneity in the FCR treatment. We follow Lange (2009) and use this ratio to measure heterogeneity in indirect rule. When we add this continuous measure of indirect rule to our RDD specification, the binary measure of indirect rule loses statistical significance in most of the regressions, and our estimates suggest that areas where the indirect rule was applied to a lesser extent are more developed today.

Georeferencing individual-level survey data from the PEW Global Attitudes Project, we analyze the cultural legacy of the indirect rule as one of the mechanisms underlying its long-term effect. Our hypothesis is that direct rule allowed better legal enforcement and thereby increased the level of cooperation and trust and decreased corruption (Guiso, Sapienza, and Zingales, 2008). We show that the indirect rule has had a significantly negative effect on interpersonal trust and a positive effect on corruption.

In the next sections, we provide an overview of the related literature and in section 3 some historical background. We then discuss the data in section 4 and our identification strategy in section 5. In section 6, we present the results, and section 7 concludes.

2.2. Related Literature

On a broad scale, our study is related to the literature dealing with the role of historical institutions in shaping economic development. As a pioneer of this line of research, North (1990) argues that since institutional change is incremental, institutions are likely to have a persistent effect on economic outcomes and advocates quantitative research analyzing these effects of institutional change. Following North's suggestion, many scholars have conducted studies in the last two decades that link history to economic development, mostly on European expansion and colonization (see Nunn, 2009, and Acemoglu, et al., 2005, for a review). More specifically, our project contributes to the literature that examines the variation of colonial institutions across and within colonies. Acemoglu, et al. (2001, 2002) and La Porta et al. (1997, 1998) initiated this strand of research. The former focused on the variation in property rights institutions, the latter studied the differences in legal systems across colonies. Afterwards, some scholars have examined the spatial variation in colonial institutions within colonies. Banerjee and Iyer (2005) focus on variations in the British colonial land revenue systems and show that these variations have had a long-run effect on economic development. Dell (2010) finds that a forced labor system instituted by the Spanish government in Peru and Bolivia in 1573 has had a persistent influence on land tenure and public goods provision.

Our study is most closely related to a small literature that explores differences between the systems of indirect and direct rule across and within former British colonies. Indirect rule implies a collaboration between the colonial state and local intermediaries who controlled regional political institutions. In the indirect rule system, the colonial state thus delegated limited power to local elites.

Creating indices to measure the extent to which each former British colony was ruled through indirect rule, Lange (2009) shows that indirectly ruled colonies became significantly less developed countries than directly ruled colonies. He argues that the legal-administrative capacities of directly ruled states promote the ability of the state to provide infrastructure and public goods. In the context of Africa, Acemoglu, et al. (2014) explain how the practice of indirect rule resulted in insufficient postcolonial state capacity and, as a consequence, in a lack of contemporary development. Focusing on the experience of Sierra Leone, Acemoglu, et al. (2014) also find negative contemporary development effects of empowering local elites through indirect rule. In India, however, Iyer (2010) finds that areas directly ruled by the British have now less access to public goods. Using an Instrumental Variable method, she estimates the causal effect of indirect rule as compared to direct rule.

Our study also contributes to a growing literature on the long-run effects of pre-colonial institutions. Some recent studies use the quality of pre-colonial institutions to explain the observed variation in present-day development among former European colonies. These studies use anthropological data from Murdock's (1967) *Ethnographic Atlas* that contains categorical variables describing several institutional and cultural features of ethnic institutions. Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013) use the degree of political centralization of African ethnic groups as a determinant of post-colonial African development and show a positive correlation between political centralization and economic development. Angeles and Elizalde (2017) use a similar approach to study Latin American countries and find the same positive association. Fenske (2013) shows that in Africa pre-colonial polygamy, slavery, and land rights institutions are able to predict post-colonial land inheritance and polygamy institutions.

2.3. Historical and Institutional Background

A. The Frontier Crimes Regulation

In 1849, the British displaced the Sikh Empire, annexed the Punjab, and advanced up to the rugged mountainous regions inhabited by Pashtun tribes. In colonial records, the mountainous area is referred to as the "tribal areas." Expanding their territory by moving forward into the tribal areas, the British drew in 1895 an international boundary, the so-called "Durand Line", that separated Afghanistan from British India (Baha, 1978). To the east of the Durand Line, an internal boundary separated the tribal areas from the so-called "settled districts", also known as the north-west frontier districts of British India. While the tribal areas were politically controlled by the British, they were free from any interference in internal affairs⁸. In the settled districts, however, the British promoted land settlement, collected taxes, and provided public goods. To establish political control in addition to effective tax collection, the colonial state tried to establish a regime of law and order in the settled districts. This

⁸The tribal areas consisted of the semi-autonomous princely states and the autonomous tribal agencies.

endeavor was, however, only a limited success. This is why the colonial state decided to implement a region-specific legal system in cooperation with local elites.

In the mid-nineteenth century, the government of colonial India institutionalized a uniform judicial system known as the Anglo-Indian Codes that included codes of criminal law (the 1860 Indian Penal Code) and codes of criminal procedure (the 1861 Code of Criminal Procedure). Shortly after the enactment of these codes, to protect Europeans against attacks by natives of the north-western frontier regions, the colonial state passed the 1867 Punjab Murderous Outrages Act⁹. The British realized however quickly that the Anglo-Indian Codes were not adequate to establish public order in the frontier regions. The crime rate increased, and indigenous people, especially local elites, resisted the European legal system; they favored their customary law. The colonial state reacted by establishing the Punjab Frontier Regulation in 1872 which allowed the customary law to supplement to the European system¹⁰. By 1901, the Punjab Frontier Regulation was revised twice in 1887 and 1901, and finally renamed *Frontier Crimes Regulation (FCR)*¹¹.

In the north-western frontier regions, the ethnic method of conflict resolution and decision making, Jirga, was applied in accordance with an unwritten ethnic code of law called Pashtunwali. Jirga, or council of elders, was a gathering of all adult male members of a society to make decisions concerning the tribe's (village's, clan's, etc.) collective actions and to settle disputes. The FCR authorized the British rulers to send both civil and criminal cases to Jirga. However, the FCR-Jirga differed from the traditional form. The FCR Jirga consisted of three members of the local elites officially appointed by the colonial state (Nichols, 2013).¹² Although the FCR appeared to indicate colonial respect for ethnic traditions, the system was redesigned to gain the support of local allies. By strengthening the position of the local elites, the FCR-Jirga ensured that the elites remained loyal to the colonial state. Both the colonial state and the elites benefited from the FCR-Jirga. Nichols (2001) provides some examples how the elites exercised their political power using the FCR-Jirga, and Tripodi (2011) explains how the British managed to influence the society via this redesigned traditional institution.

B. The FCR boundary

The implementation of the FCR discretely changed at the administrative boundary of the subjected colonial districts and sub-districts, on one side the colonial district authorities could decide whether to deal with a case under the FCR, while on the other side all cases were subject to the Anglo-Indian Codes. The FCR boundary was established in two steps. The first 1872 version of the FCR was implemented

⁹An Act for the Suppression of Murderous Outrages in Certain Districts of the Punjab

¹⁰ Berry (1966) illustrates it as an "official acknowledgement of the inability of Western legal institutions to check the proliferation of crime on the Frontier."

¹¹It is noteworthy to mention that what we call as the FCR is different from what is known today as the 1901 FCR that is applied to the Federally Administered Tribal Areas of Pakistan. The tribal areas enjoyed in colonial era and enjoy today administrative and legal autonomy while the settled districts only had the legal autonomy provided by the FCR by 1962.

¹² According to the colonial records, the appointed members of the FCR Jirga were "from among the most prominent men" of the society (Gazetteer of the Muzaffargarh District and Leiah Tehsil, 1916).

in the settled districts and the tribal areas that were occupied by that time. The 1872 FCR boundary therefore followed the pre-existing administrative borders of the settled districts demarcated in the early 1850s. The colonial state established the administrative borders of these districts after annexing Punjab in 1849 to increase “administrative efficiency” (Schwartzberg, 1978; Mathur, 1973). The artificial drawing of the district borders took place almost two decades before enacting the first revision of the FCR, i.e. at the time when the colonizers had hardly any detailed knowledge of ethnic institutions and ethnic borders¹³.

The revised 1887 FCR changed the FCR boundary. The Baluchistan Province, annexed in 1884, and the Dera Ghazi Khan District of Punjab were included in the 1887 revision¹⁴. The 1887 FCR boundary had two parts (illustrated in figure 1). The first part (illustrated by solid black and red lines) passed through the 1849 province of Punjab, and the second part (illustrated by dashed black line) was the border between the two provinces of Sind and Baluchistan that were annexed in 1843 and 1884, respectively. Two parts of the border passing through the 1849 Punjab province correspond to ethnic divides (illustrated by solid black lines): the Punjabi-Pashtu and the Punjabi-Baluchi divide. We focus only on that portion of the boundary passing through the 1849 province of Punjab where the ethnic distribution is identical on both sides (illustrated by the solid red lines).

In 1901, proposed by the Viceroy Lord Curzon, some parts of the settled districts were separated from Punjab to form the new province, the North West Frontier Province (NWFP)¹⁵. The creation of the NWFP, however, made no change in the FCR boundary, i.e. those areas that were under the FCR until 1901 were also subject to the revised 1901 FCR. After Pakistan became independent in 1947, the FCR was left intact until 1952. In several steps in the 1950s and 1960s, the state of Pakistan abolished the FCR in most regions of the country except for small tribal areas along the Afghan border (Berry, 1966). Nevertheless, in some villages on both sides of the FCR boundary people still use Jirga as an informal community reconciliation system.

2.4. Data

We use colonial records, the *Punjab District Gazetteers* and the *NWF Province District Gazetteers*, to identify where the FCR applied. Volume A of each *District Gazetteer*, under the section *Civil and Criminal Justice*, details whether the FCR applies to the district or not¹⁶. The FCR boundary consisted

¹³ Although over one century annexation of Punjab its administrative borders changed many times, the borders of the so-called settled districts experienced very little changes. For example, the Hazara-Rawalpindi border was as drawn in 1950 or the Leiah border had negligible changes.

¹⁴ According to the *Baluchistan District Gazetteer Series*, the FCR applied to the entire province of Baluchistan.

¹⁵ The three settled districts of Peshawar, Kohat, and Hazara remained unchanged formed three districts of the NWFP. The two settled districts of Bannu and Dera Ismail Khan were split up into two parts: the sub-districts of Mianwali, Isa Khel, Bhakkar, and Leiah were left in Punjab, and the rest formed two districts of the NWFP.

¹⁶ Callen et al. (2015) create a dataset of when and where the FCR has been applied since 1901. The dataset is not available. However, Figure 1 of the paper that shows the FCR application over time confirms what we collected from the gazetteers.

of the borders separating the indirectly ruled districts from those that were directly ruled. To geo-locate the FCR boundary, we started out with a GIS map of the Pakistani administrative district borders created by Pakistan's Population Census Office in 2011. We then georeferenced colonial maps of those colonial districts that are (1) identified as indirectly ruled districts and (2) shared a border with directly ruled areas. Finally, we compared the FCR boundary extracted from the current GIS map with the georeferenced colonial maps to ensure that there has been no considerable change in the FCR boundary since the colonial era.

To estimate the persistent effect of the spatial variation in direct and indirect rule on economic development, our primary data sources are the 1998 Housing and Population Census (HPC) and the 2008 Mouza¹⁷ Census (MC) of Pakistan. We obtained access to the HPC at census tract level that coincides with the administrative boundaries of a village in rural areas and boroughs in urban areas. The 1998 HPC includes educational outcomes, access to piped water and electricity, and housing quality. The 2008 MC includes information on infrastructure and public goods at the village level. Using the GPS coordinates of Pakistani villages/cities produced by the National Geospatial-intelligence Agency, we geo-referenced the villages/cities of the 1998 HPC and the 2008 MC, and then computed the great-circle distance of each village/city to the FCR boundary.

We exclude Islamabad (the capital city of Pakistan), because Islamabad did not exist during the British era and was built in the 1960s. Table 1 reports descriptive statistics of the employed development outcomes of the villages/cities located within 15km on either side of the FCR boundary. The literacy rate¹⁸ is quite low (36%) and slightly lower (35%) for villages/cities that were under indirect rule. The 1998 HPC dataset that we managed to obtain only contains the number of illiterate people, literate people who are able to read and write but do not hold any school certificate, literate people who hold a school certificate below the Secondary School Certificate (SSC) (grade 10), and literate people holding a SSC or above. We, therefore, create an index of schooling attainments, i.e. a continuous variable ranging from 0 to 3. This index is simply a weighted mean of the proportion of each group where the weights 0, 1, 2, and 3 are assigned to illiterate, literate without school certificate, holding a certificate below the SSC, and holding a SSC or above.

The variables "access to electricity" and "access to piped water" are the shares of houses that have access to electricity and piped water. The mean of the "access to piped water" variable is about 11%, reflecting the very poor provision of piped water. The mean is even lower (about 9%) for villages/cities previously ruled directly. The mean of "access to electricity" (45%) is higher than access to piped water and is somewhat higher (48%) for villages/cities previously ruled directly. "Housing quality" is the share of brick houses. Housing quality is on average about 42% and very similar on both sides of the FCR boundary. "Distance to hospital" and "distance to road" measure the distances between

¹⁷ Mouza means village.

¹⁸ Proportion of people in a village/city above age 10 who could read and write in any language.

Table 1. Descriptive Statistics of the Border Sample

	Distance < 15 km to the FCR Boundary											
	Panel A: Both sides				Panel B: Indirect Rule side				Panel C: Direct Rule side			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
A. Development Outcomes, by data source												
2008 Mouza Census												
Distance to the Closest Hospital (km)	9.378	10.75	0	85	11.21	13.05	0	85	8.062	8.491	0	70
Distance to the Closest Road (km)	1.333	3.795	0	60	1.940	4.079	0	40	0.898	3.516	0	60
Bricked Street	1.815	0.875	1	4	1.625	0.756	1	4	1.951	0.928	1	4
Bricked Drainage	1.674	0.901	1	4	1.441	0.759	1	4	1.841	0.957	1	4
Obs.	1,308				547				761			
1998 Housing and Population Census												
Literacy rate	0.359	0.177	0	1	0.350	0.170	0	0.891	0.366	0.183	0	1
Schooling Attainment	0.585	0.321	0	2.5	0.568	0.307	0	1.734	0.599	0.332	0	2.5
Housing Quality	0.422	0.338	0	1	0.428	0.353	0	1	0.417	0.325	0	1
Access to piped water	0.108	0.218	0	1	0.129	0.242	0	1	0.0891	0.192	0	1
Access to electricity	0.452	0.375	0	1	0.423	0.377	0	1	0.477	0.372	0	1
Obs.	1,647				762				885			
B. Geographical Characteristics												
Altitude	383.7	428.4	106.1	2,396	443.0	486.1	108.6	2,396	332.7	364.2	106.1	1,859
Ruggedness	13.87	27.60	0.706	187.5	17.87	32.44	0.706	187.5	10.43	22.08	0.794	102.2
Distance to the Closet River in km	19.83	13.69	0.093	54.76	21.40	14.19	0.093	50.27	18.48	13.10	0.234	54.76

Note: The sample is restricted to the villages within 15km distance to the border. The unit of observation is village. “Distance to hospital” and “distance to road” measure the distance between a village and the closest hospital and road in kilometers. “Bricked street” and “bricked drainage” are categorical variables ranging from 1 to 4; 1: none, 2: some, 3: almost all, 4: all. The variable “bricked street” measuring the provision of bricked streets and bricked drainage system in the village. “Schooling attainment” is a continuous variable ranging from 0 to 3. “Relative gender gap in schooling attainment” is computed as female schooling attainments over male schooling attainments. “Access to electricity” and “access to piped water” are the ratio of houses that have access to electricity and piped water to the total number of houses. “Housing quality” is the ratio of brick houses to the total number of houses. Altitude is the mean of altitudes in each 1x1 km grid cells coming from the SRTM. Ruggedness is the standard deviation of altitudes in each 1x1 km grid cells.

a village and the closest hospital and road in kilometers. These two variables are zero when there is a hospital in the village or when a road runs through the village. “Bricked street” and “bricked drainage” are categorical variables ranging from 1 to 4; 1: none, 2: some, 3: almost all, 4: all. The variable “bricked street” measures the provision of bricked streets in the village; the mean of 1.81 shows that fewer than “some of the streets” are bricked. The “bricked drainage” variable indicates to what extent the houses in a village are connected to a bricked drainage system.

Geographic characteristics of the villages/cities are inferred from 3-arc-second-resolution (about one hundred m²) elevation data produced by NASA’s Shuttle Radar Topography Mission (SRTM). We divide the region into 1x1 km grid cells, and calculate the mean altitude and terrain ruggedness (standard deviation of altitudes) within each grid cell. Finally, linking the GPS coordinates of the villages/cities with the closest centroid of the grid cells, we obtain the altitude and ruggedness of each village/city. We also compute the great-circle distance of each village/city to a river that is reported in Schwartzberg (1978) as navigable in the colonial era. Since rivers were important traditional means for transportation and an important input for agriculture, we also control for the distance to the rivers.

2.5. Empirical Strategy

To estimate the persistent effect of the FCR indirect rule on development outcomes, we exploit the discontinuous change in exposure to the indirect rule across the FCR boundary. We compare the outcomes of villages/cities which were subject to the indirect rule with the outcomes of directly ruled villages/cities by using a Regression Discontinuity Design (RDD) with the following baseline specification:

$$y_{ib} = \alpha + \beta \text{IndirectRule}_i + \gamma f(d_i) + \delta \text{IndirectRule}_i \cdot f(d_i) + X_i' \theta + S_b + \varepsilon_{ib} \quad (1)$$

where y_{ib} is the respective outcome of village/city i located in a neighborhood of segment b of the FCR boundary, and IndirectRule is a dummy variable equal to 1 if the village/city i was exposed to indirect ruling under the FCR and equal to 0, otherwise. S_b is the set of fixed effects relating to 40km segments of the 781 km long FCR boundary, and $f(d_i)$ is a polynomial of the great-circle distance d_i of village/city i to the boundary. X_i' , finally, is a vector of covariates, allowing us to control for observable factors that vary across villages/cities. These covariates (geographical factors (altitude and terrain ruggedness) and distances to rivers) are exogenous to the FCR boundary. We also control for the area of the village/city (in acres) since area may influence access to public goods.

In the baseline specification, we make sure that we compare observations in close geographic proximity by restricting the sample to the villages/cities located within 15km distance of each side of the boundary. To the best of our knowledge there is no widely accepted optimal bandwidth for an RDD

with several outcomes.¹⁹ We also use a local linear RD polynomial (Gelman and Imbens, 2014) for the baseline specification, and then report the robustness of our estimates by reporting estimates based on a wide range of bandwidths and on a quadratic RD polynomial. Figure 2 illustrates the outcomes of interest and a local linear polynomial fitted for observations within a 15km distance.

Lee and Card (2008) argue that if forcing variable is discrete, then standard errors should be clustered by groups. In our case, the forcing variable (distance to the FCR boundary) is continuous. However, in our sample, distance to the FCR boundary is almost equal within a higher administrative level than the village level, the so-called union council. We, therefore, cluster standard errors at union council level. Following Lee and Lemieux (2010), we also use robust standard errors for a robustness check and find no change in the results.

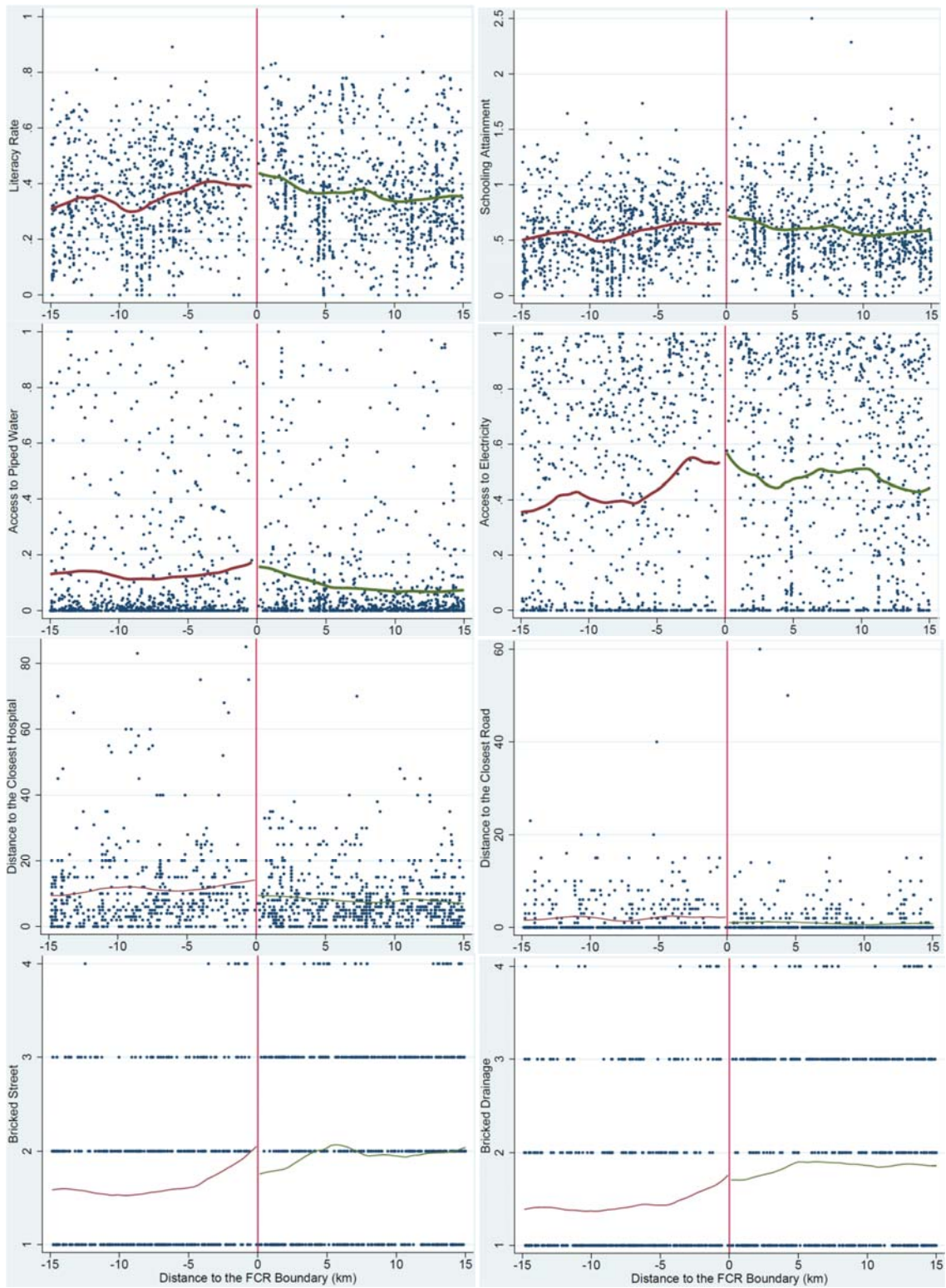
Following Becker et al. (2016), we employ a geographic matching algorithm. We first split the boundary into 20km segments and use $v_{IndirectRule,S}$ to denote a village/city v . *IndirectRule* indicates whether the village/city was under the FCR or not, and S indicates the closest 20km segment. Then, we exclude villages/cities denoted by $v_{1,S}$ ($v_{0,S}$) from the sample if there is no village/city within 15km distance of the boundary denoted by $v_{0,S}$ ($v_{1,S}$). This procedure ensures that our sample includes observations only if reasonably close-by observations exist on the other side of the border. This means that the FCR boundary that we focus on is effectively shorter than 781 km.

To identify the causal effect of the institutional variation, the key identifying assumption of the RDD is that the border was locally drawn quasi-randomly which would imply that the boundary does not follow any preexisting discontinuous difference in socio-economic and geographic factors correlated with contemporary outcomes. If this assumption is correct, all potentially relevant factors besides treatment vary smoothly at the border. To test the plausibility of this assumption, we estimate equation (1) using altitude, terrain ruggedness, and distance to the next river as outcome variables. The identification assumption requires that β is not statistically significant when we use these three exogenous geographic characteristics instead of the outcomes of interest. In table 2, we report the estimates of β for mean altitude, ruggedness, and distance to the next river, using local linear and quadratic RD polynomials. We find no jump in any of the three outcomes except for terrain ruggedness in the local linear specification which is statistically significant but relatively small.

Our dependent variables are of two types: continuous (educational outcomes, access to piped water and electricity, housing quality, and distances to the next hospital and road) and categorical variables (bricked streets, and bricked drainage). To estimate equation (1), we estimate OLS when the outcome variable is continuous and ordered logit regression when it is categorical.

¹⁹ Imbens and Kalyanaraman (2012) (IK) propose a method for choosing optimal bandwidth when employing one outcome.

Figure 2: Development Outcomes, Local Linear Regressions at the FCR Boundary



Note: Fitted line is a non-parametric local linear polynomial for either side of the FCR boundary.

Table 2. Falsification Tests

	Linear RDD	Quadratic RDD	Linear RDD	Quadratic RDD	Linear RDD	Quadratic RDD
	Altitude		Ruggedness		Distance to the Closet River	
Indirect Rule	48.113 (38.1190)	32.346 (47.6009)	9.003** (4.2639)	9.213 (6.1384)	0.580 (2.2474)	2.753 (2.8127)
Distance	-2.373 (2.1083)	-4.878 (9.4016)	-0.123 (0.1506)	0.195 (0.6751)	0.045 (0.1807)	-0.300 (0.6333)
Distance ²		0.158 (0.5751)		-0.020 (0.0427)		0.022 (0.0407)
Indirect Rule*Distance	2.811 (3.4257)	2.448 (13.9660)	0.697* (0.4021)	0.178 (1.4541)	-0.107 (0.2562)	1.218 (0.8468)
Indirect Rule*Distance ²		-0.332 (0.8384)		0.008 (0.0747)		0.038 (0.0513)
Observations	1,647	1,647	1,647	1,647	1,647	1,647
R-squared	0.877	0.877	0.720	0.720	0.727	0.729
Segment FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The unit of observation is village. Altitude is the mean of altitudes in each 1x1 km grid cells coming from the SRTM. Ruggedness is the standard deviation of altitudes in each 1x1 km grid cells. Distance is the distance to the FCR boundary in km. The unit of analysis is the village/city. Standard errors cluster at a higher administrative level than the village level, the so-called union council are reported in parentheses. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

2.6. Results

A. Development Outcomes

In this section we report our results on the effect of indirect rule on economic development and public goods provision. Table 3 reports the estimates of the baseline RDD specifications. In all columns, *indirect rule* refers to the discontinuous jump at the FCR boundary, i.e. the respective entry reports the estimate of β in equation (1). All models control for mean altitude, ruggedness, and distance to the next river. Dependent variables in columns 1 to 5 are from the 1998 Housing and Population Census and in columns 6 to 9 are from the 2008 Mouza Census.

The estimates reported in columns 1 and 2 in Table 3 show that past exposure to indirect rule has a negative effect on today's educational outcomes: it lowers today's literacy rate by around 5% and today's schooling attainments by around 0.085. The estimates are not only statistically significant but also economically sizable considering the average of literacy rate (36%) and schooling attainments (0.58) and also their standard deviations, which implies that direct rule had a positive effect on human capital. The estimates reported in columns 3 to 5 are statistically insignificant, indicating that indirect rule had no significant effect on access to piped water and electricity and also quality of houses. Both piped water and electricity began to be provided after independence in 1947, suggesting that indirect rule did not have any detrimental effects on the provision of more modern types of infrastructure.

The estimates reported in columns 6 to 9 show that villages/cities exposed to indirect rule have today inferior access to infrastructure and public goods as measured by the 2008 Mouza Census. The villages under the FCR indirect rule are on average 3km further removed from a hospital and 1km from a road. These estimates are economically large considering the average distance to hospitals of 9km and to roads of 1km. The indirect rule has also had a statistically significant and negative effect on the availability of bricked streets and bricked drainage. The British colonial state made a massive investment in constructing roads and in building hospitals and sanitary systems. The results, therefore, suggest that colonial investment policies can explain the results.

In total, these results suggest that indirect rule only retarded the provision of those public goods which began to be provided in the colonial era. On both sides of the FCR boundary, the colonial state determined where and how much to invest on education, health, transportation, agriculture, and communication. While extractive institutions, i.e. taxation, were similar on the both sides, public investment were arguably made unevenly. We conclude that the observed divergence in development started in the colonial period by way of an uneven investment in publicly provided goods on the two sides of the FCR boundary, and this divergence persists until today.

B. Robustness checks

To check the robustness of our findings, we first test how the choice of bandwidth changes the estimates. Each panel in Figure 3 plots point estimates of β in equation (1) for bandwidths between 10 to 60 kilometers, where the x-axis measures the bandwidth and the y-axis reports the point estimates. The point estimate confidence intervals are indicated by the blue lines. The thickest lines indicate the 90% and the thinnest lines the 99% confidence intervals. Following Imbens and Kalyanaraman (2012) (IK), we compute optimal bandwidth for each outcome. To compute the IK optimal bandwidth, we restricted the sample to observations located within those districts that are next to the FCR boundary.²⁰ The red vertical line in each panel depicts the IK optimal bandwidth. In Table 4, we report the estimates of the discontinuous jump (the estimates of β in equation 1) when using the linear RDD polynomial for the observations falling within a 30, 45, and 60 km band along the boundary and when using the quadratic RD polynomial specification. Widening the bandwidth, two patterns emerge from Figure 3 and Table 4. First, the point estimates for those dependent variables that represent local development, i.e. educational outcomes and bricked street and drainage, remain negative and statistically significant, although the magnitudes decrease. The second is that for larger bandwidths the effect of indirect rule on the distance to the next hospital and road changes in favor of areas ruled indirectly. These findings can be explained by the construction of transportation infrastructure, i.e. railways and roads, in the colonial era. In areas previously ruled indirectly, railways and roads constructed by the colonial state passed through areas that were far from the FCR boundary, and therefore, the further the villages/cities are from the FCR

²⁰ The IK optimal bandwidth is sensitive to observations included in sample.

Table 3: development outcomes, baseline specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Literacy Rate	Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity	Distance to the Closest Hospital	Distance to the Closest Road	Bricked Street	Bricked Drainage
Indirect Rule	-0.047** (0.0211)	-0.085** (0.0385)	-0.035 (0.0282)	-0.033 (0.0338)	-0.056 (0.0432)	3.134* (1.6193)	1.098** (0.4540)	-1.298*** (0.3221)	-1.434*** (0.3625)
Distance	0.0004 (0.0016)	0.001 (0.0029)	0.0004 (0.0022)	-0.003 (0.0020)	0.0004 (0.0031)	-0.177** (0.0780)	-0.055* (0.0290)	0.043** (0.0192)	0.029 (0.0196)
Indirect Rule*Distance	-0.001 (0.0026)	-0.003 (0.0049)	-0.002 (0.0032)	-0.0004 (0.0037)	0.001 (0.0053)	0.418** (0.1637)	0.120** (0.0564)	-0.135*** (0.0344)	-0.094** (0.0379)
Observations	1,647	1,647	1,647	1,647	1,647	1,308	1,308	1,308	1,308
R-squared	0.502	0.481	0.666	0.233	0.404	0.108	0.098		
Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

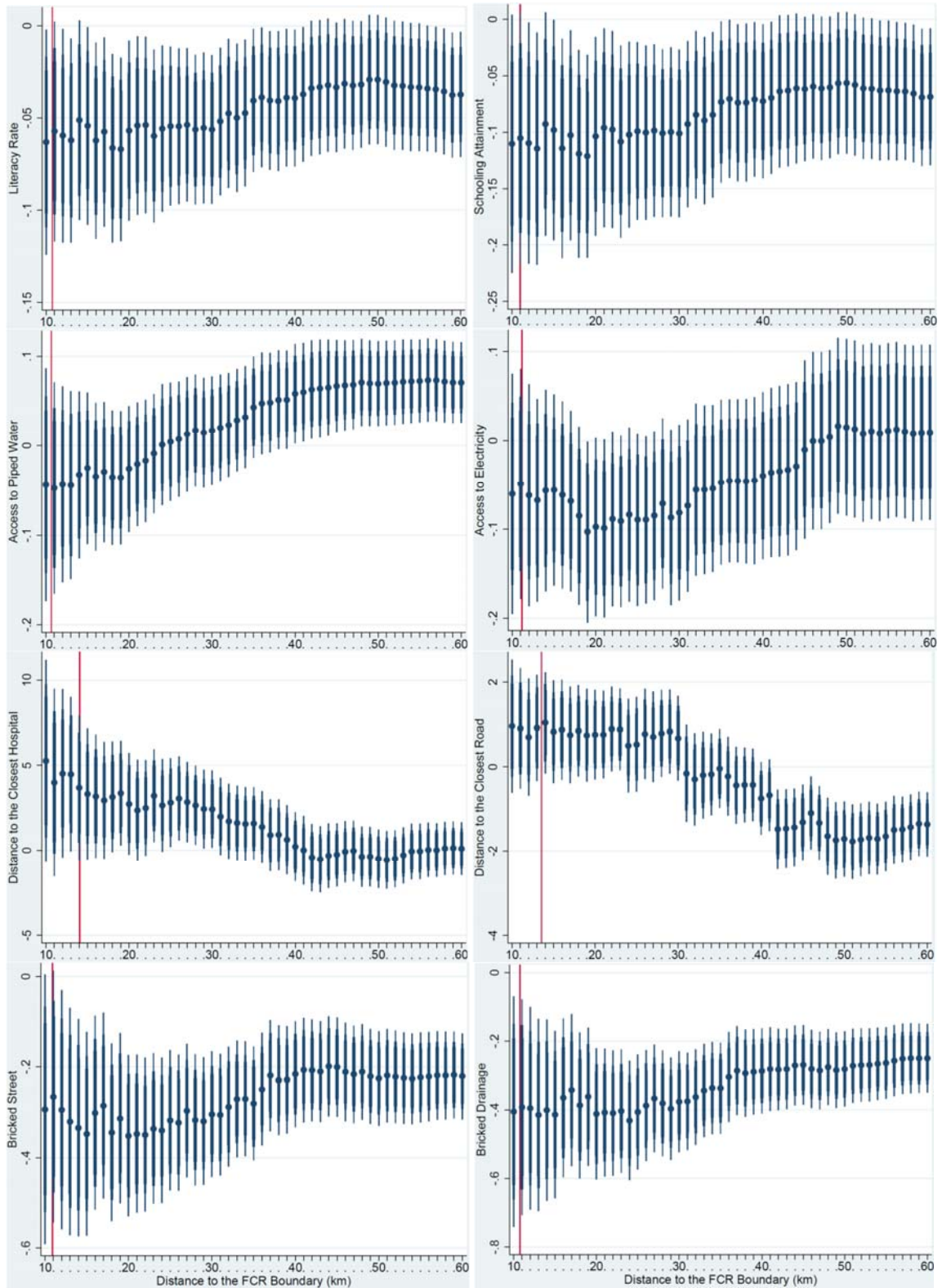
Note: The unit of analysis is village/city. Sample contains villages/cities located within 15km of the FCR boundary. All columns include geographical controls for altitude and ruggedness, distance to river, area of the village/city, and 40km segment fixed effects. Standard errors cluster at a higher administrative level than the village level, the so-called union council are reported in parentheses. Coefficients and standard errors are from OLS estimation in columns 1 to 7 and from ordered logit estimation in columns 8 and 9. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

Table 4: Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Literacy Rate	Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity	Distance to the Closest Hospital	Distance to the Closest Road	Bricked Street	Bricked Drainage
	<i>Local Quadratic Regression, 15km</i>								
Indirect Rule	-0.084*** (0.0305)	-0.151** (0.0585)	0.077* (0.0455)	-0.091 (0.0700)	-0.082 (0.0728)	5.828* (2.9901)	0.994 (0.7527)	-0.344 (0.4849)	-0.885 (0.5949)
Observations	1,647	1,647	1,647	1,647	1,647	1,308	1,308	1,308	1,308
R-squared	0.504	0.484	0.669	0.236	0.404	0.110	0.098		
	<i>Local Linear Regression, 30km</i>								
Indirect Rule	-0.056*** (0.0156)	-0.101*** (0.0287)	-0.0004 (0.0221)	0.017 (0.0235)	-0.081** (0.0382)	2.422*** (0.8784)	0.674* (0.3861)	-0.897*** (0.1673)	-1.281*** (0.1916)
Observations	3,738	3,738	3,738	3,738	3,738	3,682	3,682	3,682	3,682
R-squared	0.496	0.476	0.635	0.214	0.334	0.057	0.045		
	<i>Local Linear Regression, 45km</i>								
Indirect Rule	-0.033** (0.0136)	-0.062** (0.0248)	0.016 (0.0198)	0.067*** (0.0198)	-0.010 (0.0388)	-0.257 (0.7217)	-1.316*** (0.3479)	-0.548*** (0.1276)	-0.865*** (0.1418)
Observations	5,489	5,489	5,489	5,489	5,489	5,427	5,427	5,427	5,427
R-squared	0.500	0.482	0.587	0.148	0.306	0.084	0.148		
	<i>Local Linear Regression, 60km</i>								
Indirect Rule	-0.037*** (0.0131)	-0.069*** (0.0235)	-0.005 (0.0195)	0.071*** (0.0174)	0.009 (0.0382)	0.100 (0.6041)	-1.366*** (0.2969)	-0.560*** (0.1117)	-0.739*** (0.1227)
Observations	6,947	6,947	6,947	6,947	6,947	6,908	6,908	6,908	6,908
R-squared	0.508	0.489	0.558	0.125	0.303	0.115	0.271		

Note: The unit of analysis is village/city. All specifications include geographical controls for altitude and ruggedness, distance to river, area of the village/city, and 40km segment fixed effects. Standard errors cluster at a higher administrative level than the village level, the so-called union council are reported in parentheses. Coefficients and standard errors are from OLS estimation in columns 1 to 7 and from ordered logit estimation in columns 8 and 9. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

Figure 3: Point Estimates for Choosing Bandwidth between 10 to 60 Kilometers



Note: Each panel plots point estimates of β in equation (1) for choosing bandwidth between 10 to 60km. Lines stemming from the point estimates show confidence intervals. The red vertical line in each panel depicts the Imbens and Kalyanaraman (2012) optimal bandwidth.

boundary, the closer they are to the railroads. The roads and railways promoted growth and development, and those villages/cities that had better access to them grew faster. When we enlarge the bandwidth, villages/cities that have had better access to the roads and railways are included in the sample. Our finding is thus in line with the argument that the British colonial investment had beneficial long-run effect.

C. *An index measuring the extent of the FCR*

In the sharp RDD, used so far, the probability of the FCR treatment jumps from 0 in the areas ruled directly to 1 in the indirectly ruled areas. It, therefore, does not capture heterogeneity in the implementation of the FCR across the indirectly ruled areas. The FCR entitled the colonial authorities of each district to decide whether to refer a case to Jirga under the FCR or to courts on colonial government under the Anglo-Indian Codes. Their decisions depended on the size and importance of tribes which applied Jirga and the geopolitical importance of the home district of these tribes. We follow Lange (2009) and use the percentage of cases allocated to Jirga courts to measure heterogeneity in the implementation of indirect rule. In volume B of the *Punjab District Gazetteers*, the number of cases referred to Jirga is documented. We computed the share of civil and criminal cases dealt with under the FCR for each district that was ruled indirectly in the period of 1904 to 1911, the only period when data is available for all districts.

We add this measure to our RD specification and report the estimates in Table 5. When including “*heterogeneity in indirect rule*”, the binary *indirect rule* variable does not have a statistically significant effect on educational outcomes and distance to hospitals and roads, implying that the index measuring heterogeneity in the implementation of indirect rule captures the effect of the binary variable. The point estimates of the index are remarkably larger but not statistically significant in all models. The findings suggest that areas where the indirect rule was applied to a lesser extent are more developed today, and thereby supports our main result.

2.7. Cultural legacy of the FCR

If the observed divergence in development commenced in the colonial era, why does it persist today, almost 50 years after abolishing the FCR? Our hypothesis is that the FCR influenced social norms, specifically interpersonal trust and corruption. The direct rule was better suited for establishing and sustaining the rule of law than the indirect rule that empowered local elites to exercise power with the help of an informal method of serving justice. Guiso, Sapienza, and Zingales (2008) argue that the rule of law has a permanent and positive effect on cooperation and trust. This gives rise to the hypothesis that in districts that were formerly subject to direct rule, the level of cooperation and trust is higher and corruption lower than in districts with a legacy of indirect rule. We thus suggest that the mechanisms underlying the long-term development effect of indirect rule is to be found in the cultural legacy of indirect ruling. We test this hypothesis by using individual-level survey data from the PEW Global

Table 5: A measure of heterogeneity in the FCR indirect rule

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Literacy Rate	Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity	Distance to the Closest Hospital	Distance to the Closest Road	Bricked Street	Bricked Drainage
Indirect Rule	-0.031 (0.0234)	-0.056 (0.0420)	-0.021 (0.0300)	-0.044 (0.0357)	-0.034 (0.0476)	2.715 (1.7809)	0.536 (0.4934)	-1.500*** (0.3647)	-1.442*** (0.4342)
Heterogeneity in Indirect Rule	-0.298 (0.1901)	-0.542* (0.3275)	-0.263 (0.2868)	0.194 (0.2012)	-0.400 (0.5156)	8.823 (13.5126)	11.832** (5.4158)	3.891 (3.1124)	0.143 (3.9954)
Observations	1,647	1,647	1,647	1,647	1,647	1,308	1,308	1,308	1,308
R-squared	0.503	0.483	0.667	0.234	0.404	0.108	0.101		

Note: The unit of analysis is the village/city. Sample contains villages/cities located within 15km of the FCR boundary. All columns includes geographical controls for altitude and ruggedness, distance to river, area of the village/city, and 40km segment fixed effects. Standard errors cluster at a higher administrative level than the village level, the so-called union council are reported in parentheses. Coefficients and standard errors are from OLS estimation in columns 1 to 7 and from ordered logit estimation in columns 8 and 9. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

Attitudes Project. Two waves of the survey, 2007 and 2009, include the names of the cities where the respondents live. This allows us to georeference these cities and to compute their great-circle distance to the FCR boundary.

The survey includes one question about interpersonal trust. It asks respondents whether “most people in this society are trustworthy.” Answer categories are: 1 = completely agree; 2 = mostly agree; 3 = mostly disagree; 4 = completely disagree. The survey also includes a question on bribery. Respondents are asked the question: “in the past year, how often, if ever, have you had to do a favor, give a gift or pay a bribe to a government official in order to get services or a document that the government is supposed to provide?” Answer categories are: 1 = very often; 2 = somewhat often; 3 = not too often; 4 = not at all. Again, we estimate the equation (1), now with these two cultural indicators as dependent variables, using ordered logit.

Table 6 reports our findings for four bandwidths: cities located within 30, 45, and 60km distance to the FCR boundary. We control for geographical characteristics of cities and some individual characteristics: gender, age, and religion. The significantly positive coefficients of *IndirectRule* in the trust columns indicate a significantly negative effect of indirect rule on interpersonal trust. For bribery the results are similar: the significantly negative coefficients indicate a positive effect of indirect rule on bribery. Although magnitudes change with changes in bandwidth, the signs and statistical significance are stable, except for the estimates of bribery when we widen the bandwidth to 60km. Table 7 reports the marginal effects. This allows us to discuss the size of the effects estimated by the ordered logit model. Table 7 shows the marginal effects of past exposure to indirect rule category by category and also the average absolute marginal effects. For example, for the bandwidth that includes cities located within 45km distance to the FCR boundary, past exposure to indirect rule decreases the probability of answering “completely agree” that “most people in this society are trustworthy” by about 27% percentage points and increases complete disagreement by 33 percentage points. Holding the other variables constant at their means, past exposure to indirect rule, again for the 45km bandwidth, increases the probability of moving to a higher category in trust, i.e. going towards disagreement with trustworthiness, by 20%.

2.8. Conclusion

Iyer (2010), comparing contemporary outcomes of areas ruled directly by the British colonial state of India with those ruled indirectly, finds that the directly ruled areas are less developed today. Focusing only on India (excluding present-day Pakistan and Bangladesh), in her study directly ruled areas are those that were under the administration of the British Crown, and indirectly ruled areas are those that were subject to administrative and legal autonomy. She argues that the British colonial state did not invest in physical and human capital in the directly ruled areas as much as local states did in the indirectly ruled areas.

Table 6: Cultural legacy of the FCR indirect rule

	Trust			Bribery		
	Distance < 30	Distance < 45	Distance < 60	Distance < 30	Distance < 45	Distance < 60
Indirect Rule	2.003*** (0.4362)	1.737*** (0.3013)	1.829*** (0.2599)	-1.547* (0.8604)	-1.816** (0.7344)	-0.685 (1.2061)
Distance	0.047** (0.0200)	0.028** (0.0120)	0.029*** (0.0075)	-0.000 (0.0688)	-0.043*** (0.0165)	-0.028 (0.0218)
Indirect Rule*Distance	-0.063* (0.0339)	-0.031* (0.0178)	0.001 (0.0086)	-0.073 (0.0899)	-0.006 (0.0290)	0.011 (0.0440)
Female	0.914 (0.5696)	0.755 (0.5167)	1.016** (0.4778)	0.717 (0.7062)	0.865 (0.5674)	0.961* (0.5240)
Age	0.005 (0.0055)	0.001 (0.0037)	0.000 (0.0029)	-0.002 (0.0136)	-0.002 (0.0109)	-0.001 (0.0104)
Muslim	-14.289*** (0.9754)	-13.108*** (0.9566)	-13.875*** (0.9529)	-14.846*** (0.7258)	0.287 (1.9294)	0.332 (1.6424)
Observations	157	220	268	255	400	472
Number of Cities	10	14	17	17	25	30

Note: The unit of analysis is individual. All columns includes geographical controls for altitude and ruggedness, and distance to the closest river. Standard errors are clustered at each city. Coefficients and standard errors are from ordered logit estimation. Trust is answer to the question “Please tell me whether most people in this society are trustworthy.” Answer categories are: 1 = completely agree; 2 = mostly agree; 3 = mostly disagree; 4 = completely disagree. Bribery is answer to the question “In the past year, how often, if ever, have you had to do a favor, give a gift or pay a bribe to a government official in order to get services or a document that the government is supposed to provide?” Answer categories are: 1 = very often; 2 = somewhat often; 3 = not too often; 4 = not at all. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

Table 7: Marginal effects of the FCR indirect rule of social norm

	Trust				Bribery		
	Distance < 30	Distance < 45	Distance < 60		Distance < 30	Distance < 45	Distance < 60
Completely agree	-0.345	-0.273	-0.331	Very often	0.100	0.143	0.046
Mostly agree	-0.104	-0.131	-0.078	Somewhat often	0.069	0.100	0.036
Mostly disagree	0.137	0.078	0.085	Not too often	0.048	0.050	0.016
Completely disagree	0.313	0.327	0.324	Not at all	-0.216	-0.29	-0.098
Average absolute marginal effect	0.225	0.202	0.204	Average absolute marginal effect	0.108	0.146	0.049

Note: Marginal effects of past exposure to indirect rule in a model specification that includes all control variables shown in Table 6, holding other variables constant at their mean. For instance, the upper-left entry of -0.345 indicates that being exposed to indirect rule decreases the probability of answering “complete agree” by 34.5 percentage points. The last row displays the average absolute percentage change across categories. It indicates how much being exposed to indirect rule changes the probability of moving to a higher category, on average across four categories. Trust is answer to the question “Please tell me whether most people in this society are trustworthy.” Answer categories are: 1 = completely agree; 2 = mostly agree; 3 = mostly disagree; 4 = completely disagree. Bribery is answer to the question “In the past year, how often, if ever, have you had to do a favor, give a gift or pay a bribe to a government official in order to get services or a document that the government is supposed to provide?” Answer categories are: 1 = very often; 2 = somewhat often; 3 = not too often; 4 = not at all.

We focus on a special case of indirect rule in the northwest of British India (present-day Pakistan) where both directly and indirectly ruled areas were administered by the colonial state but under two different legal codes. Directly ruled areas were under British law while indirectly ruled areas could apply their ethnic method of justice. Using a regression discontinuity design, we show that past exposure to indirect rule has had a negative effect on development and also on the provision of those public goods that began to be provided in the colonial era. We argue that the British colonial state did not invest in physical and human capital in the indirectly ruled areas as much as it did in the directly ruled areas. Our results are in line with Lange (2009), who suggests that the legal-administrative capacity of directly ruled states promote the ability of the state to provide public goods. We also propose a channel of persistence: the indirect rule decreased interpersonal trust and increased corruption.

Reference

- Acemoglu, D., Chaves, I. N., Osafo-Kwaako, P., and Robinson, J. A. (2014). “*Indirect rule and state weakness in Africa: Sierra Leone in comparative perspective.*” No. w20092. National Bureau of Economic Research.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001): “*The Colonial Origins of Comparative Development: An Empirical Investigation.*” *American Economic Review*, 91, 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). “*Reversal of fortune: Geography and institutions in the making of the modern world income distribution.*” *The Quarterly journal of economics*, 117(4), 1231-1294.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2005): “*Institutions as a Fundamental Cause of Long-Run Growth,*” in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf. Amsterdam, The Netherlands: Elsevier North-Holland, 109–139.
- Acemoglu, D., T. Reed, and J. A. Robinson (2014). “*Chiefs: Economic development and elite control of civil society in Sierra Leone.*” *Journal of Political Economy*, 122(2), 319-368.
- Alsan, M. (2015), “*The effect of the tsetse fly on African development*”, *American Economic Review* 105, 382-410.
- Angeles, L. and A. Elizalde (2017). “*Pre-colonial institutions and socioeconomic development: The case of Latin America.*” *Journal of Development Economics* 124: 22-40.
- Baha, L. (1978). “*N.-WFP Administration under British Rule, 1901-1919.*” National Commission on Historical and Cultural Research.
- Bandyopadhyay, S., and E. Green (2016). “*Precolonial Political Centralization and Contemporary Development in Uganda.*” *Economic Development and Cultural Change*, 64(3), 471-508.

- Banerjee, A., and L. Iyer, (2005). “*History, institutions, and economic performance: the legacy of colonial land tenure systems in India.*” *The American Economic Review*, 95(4), 1190-1213.
- Becker, S. O., K. Boeckh, C. Hainz, and L. Woessmann, (2016). “*The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy.*” *The Economic Journal*, 126(590), 40-74.
- Berry, W. (1966). “*Aspects of the Frontier Crimes Regulation in Pakistan.*” (No. 3) Duke University, Program in Comparative Studies on Southern Asia.
- Caroe, O. (1958). “*The Pathans: 550 bc-ad 1957.*” MacMillan; St Martin's Press.
- Callen, M., Gulzar, S., Rezaee, A., and Shapiro, J. (2017). “*Choosing Ungoverned Space: Pakistan's Frontier Crimes Regulation*”, Working Paper.
- Davies, C. C. (1932). “*The problem of the North-West frontier, 1890-1908.*” Cambridge University Press.
- Dell, M. (2010). “*The Persistent Effects of Peru's Mining Mita.*” *Econometrica*, 78(6), 1863-1903.
- Dell, M., N. Lane, and P. Querubin, (2017). “*The Historical State, Local Collective Action, and Economic Development in Vietnam.*” (No. w23208) National Bureau of Economic Research.
- Fenske, James. 2013. “*Does Land Abundance Explain African Institutions?*” *Economic Journal* 123(573): 1363–90.
- Gelman, A., and Imbens, G. (2014). “*Why high-order polynomials should not be used in regression discontinuity designs.*” (No. w20405) National Bureau of Economic Research.
- Gennaioli, N. and I. Rainer. 2007. “*The Modern Impact of Precolonial Centralization in Africa.*” *Journal of Economic Growth* 12 (3): 185–234.
- Guiso, L., P. Sapienza, L. Zingales (2008). “*Social Capital as Good Culture.*” *Journal of the European Economic Association*; 6 (2-3): 295-320.
- Huffman, S. (2015). “*Mapping the genetic relationships of the world's languages.*” Retrieved from http://www.worldgeodatasets.com/files/8113/0193/1121/Mapping_Genetic_Relationships_World_Languages.pdf.
- Imbens, G., and K. Kalyanaraman (2012) “*Optimal Bandwidth Choice for the Regression Discontinuity Estimator.*” *Review of Economic Studies* 79 (3): 933-59.
- Iyer, L. (2010). “*Direct versus indirect colonial rule in India: Long-term consequences.*” *The Review of Economics and Statistics*, 92(4), 693-713.

- La Porta R, Lopez-de-Silanes F, Shleifer A, and Vishny R. (1997). “*Legal determinants of external finance.*” *Journal of Finance*. 52:1131–50.
- La Porta R, Lopez-de-Silanes F, Shleifer A, and Vishny R. (1998). “*Law and finance.*” *Journal of Political Economy*. 106:1113–55
- Lange, M. (2009). “*Lineages of despotism and development: British colonialism and state power.*” University of Chicago Press.
- Lee, D. S., and Card, D. (2008). “*Regression discontinuity inference with specification error.*” *Journal of Econometrics*, 142(2), 655-674.
- Lee, D. S., and Lemieux, T. (2010). “*Regression discontinuity designs in economics.*” *Journal of economic literature*, 48(2), 281-355.
- Michalopoulos, S. and E. Papaioannou. 2013. “*Pre-colonial Ethnic Institutions and Contemporary African Development.*” *Econometrica* 81 (1): 113–52.
- Michalopoulos, S. and E. Papaioannou. 2014. “*National Institutions and Subnational Development in Africa.*” *Quarterly Journal of Economics* 129 (1): 151–213.
- Mathur, Y. B. (1973). “*British administration of Punjab, 1849-75.*” Delhi: Surjeet Book Depot.
- Nichols, R. (2001). “*Settling the frontier: Land, law and society in the Peshawar Valley, 1500—1900.*” Oxford University Press.
- Nichols, R. (2013). “*The Frontier Crimes Regulation: A History in Documents.*” Oxford University Press.
- North, D. C. (1990). “*Institutions, institutional change and economic performance.*” Cambridge university press.
- Nunn, N. (2009). “*The Importance of History for Economic Development.*” *Annual Review of Economics* 1: 65-92.
- Schwartzberg, J. E. (1978). “*A Historical Atlas of South Asia.*” University of Chicago Press.

Chapter 3

Was Railroad Expansion a Success Story of the British Colonial Rule in India? A Long-Term Perspective

3.1. Introduction

A large and growing literature in economics and political science has emphasized the detrimental effect of colonial extractive institutions. In some cases, however, infrastructure built by colonial rulers to extract economic surplus may have had a positive effect on local development. Dell and Olken (2017) and Jedwab and Moradi (2016), for example, show positive economic effect of the colonial Dutch Cultivation System in Java and the colonial railroads in Africa, respectively. We investigate the economic effect of railroads constructed by the colonial British state in India. This railroad network is generally acknowledged to be a success story of the British colonial rule in India.

The construction of British India's railroad network began in 1853 and expanded so fast that, by the 1900s, it became the fourth largest railroad network of the world.²¹ Prior to the development of railroads, most of the internal transportation was slow, unreliable, and costly, and rates of spoilage were high (Hurd, 1983). Donaldson (2010), estimating the short-term effect of railroads, shows that railroads greatly reduced costs of trading and increased real income. Andrabi and Kuehlwein (2010) find that, shortly after introduction of railroads, price dispersion in annual retail wheat and rice prices in Indian districts was reduced. Bogart and Chaudhary (2013) show that the railroad led to rapid economic growth at the time. We contribute to this research by estimating the long-term causal effect of British India's railroads on local development, focusing on the northwest of the colony (present-day Pakistan).

²¹ After the United States, Russia, and Germany.

At the early stages, railroad expansion was targeted at connecting major ports to the interior for commercial purposes.²² However, since the 1870s, military and strategic concerns dominated the commercial objectives (Bogart and Chaudhary, 2015). In particular, in the northwest of the colony railroads were expanded to facilitate movements of troops to the frontier and to form a network among important cantonments (Andrabi and Kuehlwein, 2010). End points of each line were either pre-existing important cities (e.g. ports, big cities, or salt mines) or new cantonments that grew eventually into big municipalities. Track placement and cities that received railway stations were also influenced by strategic concerns (Hurd, 1983). This implies that the track placement was not chosen randomly. Villages and cities that were close to the strategic locations therefore enjoyed better access to railroads. To address the endogeneity issue, we use an instrumental variable (IV) method that has been proposed in empirical literature on economic effect of transportation infrastructure (Michaels, 2008; Atack, et al., 2010; Banerjee, et al., 2012; Faber, 2014; Hornung, 2015).

The instrumental variable is based on a hypothetical network of straight-lines corresponding to tracks that would have been built if only the two end points of the railway were chosen strategically. To construct the IV, we geolocate the railroad network of British India, obtain years of opening for all line sections from colonial records, and identify as “segments” a set of connected sections that opened for traffic in one year. We then construct the hypothetical network of straight-lines between endpoints and junctions of segments (nodes). Our primary data source of economic outcomes is the 1998 Housing and Population Census (HPC) at village level. We geo-reference villages and compute their distance to the railroad network as actual railroad access and to the hypothetical network of straight lines as an IV for actual railroad access.

The identifying assumption is that the distance to the straight-lines affects contemporary outcomes only through actual railroad access, conditional on geographical factors, distance to the closest river, and distance to the closest node. This assumption would be violated if the distance to the straight-lines were correlated with pre-existing factors that are also correlated with contemporary outcomes. In a placebo falsification test, we show that altitude and the distance to the closest river, i.e. two factors influencing development, are not correlated with the IV.

In both the OLS and the IV specification, we find a statistically and economically significant effect of railroad access. Using the IV specification, we estimate that the long-lasting effect of railroad access increased the literacy rate by about 2%. Railroad access has also a positive effect on schooling attainment. Furthermore, we find a positive effect of railroad access on living standards: access to piped water, access to electricity, and housing quality. Our IV point estimates are almost twice as large as the OLS estimates. This is in line with the argument that train stations were chosen strategically from the set of pre-existing developed cities.

²² There were also some very short branches connecting salt mines to the main railroad network.

In the next sections, we provide an overview of the related literature and in section 3 some historical background. We then discuss the data in section 4. We present our identification strategy and the results in section 5. Section 6 concludes.

3.2. Related Literature

On a broad scale, our study is related to the literature on the long-term effect of colonial rules. In the past two decades, many economists have studied the effect of colonial institutions on economic development (see Nunn, 2009, and Acemoglu, et al., 2005, for a review). Our project contributes to these studies, more specifically to those studies that examine the effect of colonial investment in infrastructure and public goods. Colonial powers invested in infrastructure to maximize their extractive returns from exploitable resources. A variety of studies suggests that infrastructure constructed by colonizers had also positive effects on economic development. Dell and Olken (2017) explore the colonial experience of Java and estimate the long-term effect of the Dutch Cultivation System. The colonial state introduced a system of sugar production and processing to extract revenue. Dell and Olken show that areas that were closer to sugar factories are more developed today. Jedwab and Moradi (2016) study the impact of railroad expansion in Africa and find that railroads had short- and long-term positive effects on economic development. Jedwab et al. (2015), focusing on railroad expansion in colonial Kenya, document an immediate population growth in cities close to the railroads.

Our study also contribute to the literature on transportation infrastructure. Several studies investigate the impact of railroad expansion in 19th-century America and find an improvement in agricultural productivity (Atack and Margo, 2009), a positive effect on population density and urbanization (Atack et al., 2010), and a rise in agricultural land values (Donaldson and Hornbeck, 2016). Many studies estimate the impact of the U.S. Interstate Highway System constructed in the second half of 20th century. These studies find a decline in urban population (Baum-Snow, 2007), an increase in trade-related activities and in demand for skilled workers (Michaels, 2008), and a growth in employment (Duranton and Turner, 2012). Focusing on China, Banerjee et al. (2012) find a long-term positive effect of access to railroads on per capita GDP, Faber (2014) finds a negative effect of the National Trunk Highway System on local GDP growth in connected regions relative to non-connected regions, and Baum-Snow et al. (2017) find that highway and railroad networks led to a reduction in the cities' industrial GDP. Finally, Hornung (2015) estimates a positive effect of railroad access in 19th-century Prussia on urban population growth.

Focusing on colonial India, our project is, of course, most closely related to the literature on the economic effect of colonial railroad expansion in the sub-continent. Hurd (1983) documents various positive and negative short-term effects of railroad expansion on different sectors of the Indian economy and speculates on mechanisms underlying these effects, some of which have been more closely investigated in the meantime: Donaldson (2010) finds a decline in trade costs, a reduction in the price

gap between and within the regions, an increase in trade flows, and an increase in real income levels. Bogart and Chaudhary (2013) find an increase in total factor productivity.

3.3. Historical background

A. *Motives for Construction and Expansion*

Railroad construction in British India began in 1853 by connecting the major ports to the interior. Construction and expansion progressed rapidly, and by the 1900s the railway network became the fourth largest in the world. The British government that determined the placement of tracks had three motives for building railroads: commercial, military, and humanitarian motives. The very first railways that connected main ports to the interior and major inland cities with one another, were constructed “according to the extent of political and commercial advantages which they are calculated to afford”, as advised by Lord Dalhousie. Beginning in the 1870s, military motives became increasingly important (MacPherson, 1955). The colonial government had military concerns about controlling the population and defending the frontier.²³ Another consideration was to lower the risk of famine, by being able to transport food to regions suffering from famines, the so-called humanitarian railways. The military and humanitarian lines were built in low-profit areas with light civilian traffic. Hurd (1983) shows that unprofitable lines²⁴ formed 70% of the total length of the railroad network in 1900. Not only were the two endpoints of the railways chosen strategically but also the railway routes. For strategic reason, lines “normally avoided the central business districts and passed through the outskirts” (Hurd, 1983).

The railroad system began to operate in the North West of the British India (present-day Pakistan) in 1861 by opening a short line connecting Karachi to Kotri. The first lines were constructed for political and commercial reasons. They connected the main port of Karachi to the major cities of Multan, Lahore, and Rawalpindi in the fertile province of Punjab. The next lines, however, were mostly built to facilitate movement of troops to the frontier.²⁵ The British believed that a Russian attack on India might be launched from the North West frontier with Afghanistan. Strategic requirements therefore played an important role in determining the location of lines (MacPherson, 1955). Even after ensuring peace on the Afghan border, the British government continued to build new railways in that region because of the uprisings among the Frontier Tribes.²⁶

²³ Especially after the 1857 mutiny and the threat posed by Russian railways.

²⁴ Those lines that earned less than 5 per cent and received subsidies under the so-called system of “the guarantee.”

²⁵ Some lines were exclusively built for military purposes such as those lines passing through Peshawar and Khyber Pass or Quetta and Bolan Pass. Several other lines were officially proposed as commercial, however, had significant military purposes as well, such as those to Dera Ismael Khan, Attock, and Peshawar which had important cantonments (Administration reports on Railways in India, 1907, P. 171).

²⁶ Administration Report on the Railways in India for 1919 -1920, P. 8. As an example, in 1895 the North Western railway “was fully occupied in the movement and concentration of troops, animals and stores in connection with the late Chitral expedition” (Administration Report on the Railways in India for 1895-96, Part I, P.6.).

B. Effects of the railroads

The railroad expansion had an instant impact on all sectors of the Indian economy. Connecting the fertile plains to the main ports, the railroads connected the Indian market with the world market, reduced shipping costs, made agricultural commodities competitive internationally, boosted the export of agricultural products²⁷, increased the import of manufactured products, created new economic activities, and promoted internal trade (Tiwari, 1921). All these effects caused agriculture to become more commercialized which brought about higher land prices, rents, and tax revenues. It is not surprising that these developments induced far-reaching socio-economic changes to the rural areas (Hurd, 1983).

The colonial records document some of these effects of the introduction of railroads. In the North West of British India, connecting irrigated tracts in the provinces of Sind and Punjab to Karachi increased the export of wheat and oilseeds considerably²⁸ and decreased price differences across regions²⁹. Cities located along the railways experienced rapid population growth³⁰ and became politically more important³¹. Railroads were the engine of economic prosperity³², raised the standard of living³³, and created new markets³⁴. Railroad expansion also changed the religious composition of the population. Hindu traders moved to regions where no Hindu had lived before. Christian missionaries had better access to areas where they did not travel before.³⁵ Railroads also had an influence on culture and education, especially in the tribal areas.³⁶

²⁷ For example, India exported no wheat before railway era, while it was supplying 23 per cent of Britain's imports of wheat by 1886 (Hurd, 1983).

²⁸ Imperial Gazetteer V.15, P. 15.

²⁹ For example, the experience of Quetta after opening of the Nushki railway at 1905 (Gazetteer of Bolan and Chagai District, V. A 1907, P. 119).

³⁰ For instance, Katha that was "no towns of importance" saw "a rapid increase of population in the country lying along the railway" (Imperial Gazetteer V.15, P. 157).

³¹ For instance, Campbellpur (present-day Attock) became the head-quarters of a district after the Mari-Attock Railway had its terminus at Campbellpur (Gazetteer of Attock District, 1907, P. 202).

³² For instance, opening of Wazirabad-Khanewal railway in 1900 and Jech Doab railway in 1906 (Gazetteer of Jhang district, V. A 1929, P. 125).

³³ The experience of Bahawalpur as an example (Gazetteer of Bahawalpur District, V. A 1904, P. 275).

³⁴ For example, in Mianwali sub-district where there were no large markets, after opening of Kalabagh-Bannu railway, every railway station became itself a small center of export and a more important city (Gazetteer of Mianwali district, V. A 1915, P. 141).

³⁵ For example, According to Gazetteer of Bahawalpur district, they "commonly travel long distances on pilgrimages to the places sacred to them, while Christianity is now more frequently preached in the state by missionaries from Sindh and the Punjab (Gazetteer of Bahawalpur district, V. A 1904, P. 275).

³⁶ For instance, by opening the Quetta-Nushki railway, It had an "educating and civilizing result" which was "doubtless producing its effects among the tribesmen." (Gazetteer of Bolan and Chagai district, V. A 1907, P. 66), or, by opening Sind-Pishin railway, "the impression created on the wild tribes of the frontier" was immense and "their civilizing influence felt ... far beyond the political frontier, for every year many thousands of trans-border Afghans travel to India by their means to find remunerative employment." (Imperial Gazetteer of India Provincial Series Baluchistan, 1908, P. 53)

3.4. Data

To geolocate the British Raj railroad network in the North West of British India,³⁷ we started out with a GIS database of the Digital Chart of the World that contains the contemporary railroad network and then used colonial records and maps³⁸ to georeference abandoned railroads. The railroad network that was constructed in the colonial era (illustrated by black lines in Figure 1) makes up 96 percent of the total railroad length of present-day Pakistan.³⁹ We focus exclusively on the colonial railroad network.

To construct a hypothetical network as an instrument for the actual railroad network, we first obtained the years of opening for all line sections from the colonial records. Then, we define a segment as a set of connected sections that opened for traffic in one year and identify endpoint cities and junctions of segments as nodes.⁴⁰ We finally draw a straight line between every two connected nodes (see Figure 2). The network of straight lines corresponds to a hypothetical railroad track system that would have been constructed if the sole objective had been to connect the nodes.

Our primary data source of contemporary economic outcomes is the 1998 Housing and Population Census (HPC) at census tract level that coincides with the administrative boundaries of the rural villages. Using the GPS coordinates of Pakistani villages produced by the National Geospatial-intelligence Agency, we geo-referenced villages and then computed the great-circle distance of each village to (i) the actual railroad network, (ii) the hypothetical network of straight lines as an instrument for actual railroad access, and (iii) rivers that are reported in Schwartzberg (1978) as navigable in the colonial era.

The 1998 HPC includes educational outcomes, access to piped water and electricity, and housing quality. The dataset that we managed to obtain only contains the number of literate people⁴¹, literate people who are able to read and write but do not hold any school certificate, literate people who hold a school certificate below the Secondary School Certificate (SSC) (grade 10), and literate people holding an SSC or above. Based on this information, we created an index of schooling attainments that is a continuous variable ranging from 0 to 3. This index is simply a weighted mean of the proportion of each group, where the weights 0, 1, 2, and 3 are assigned to being illiterate, being literate without a school certificate, holding a certificate below the SSC, and holding an SSC or above. We also compute the relative gender gap in schooling attainments as the ratio of female schooling attainments and male schooling attainments. “Access to electricity” and “access to piped water” are the ratio of houses that have access to electricity and piped water and the total number of houses. Finally, “housing quality” is the ratio of brick houses and the total number of houses.

³⁷ Donaldson (2017) constructs a similar database, but the shapefiles are not available.

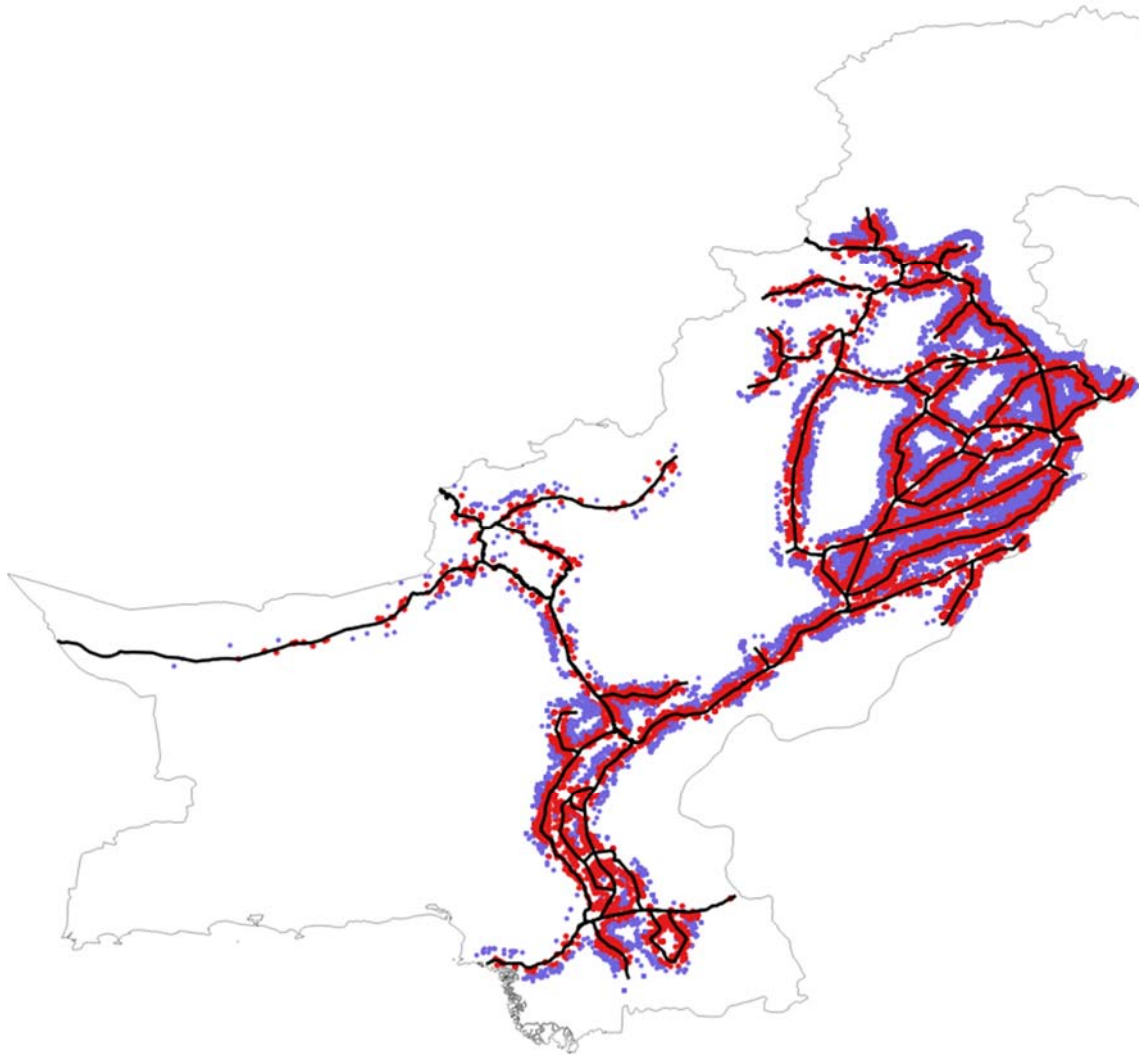
³⁸ Annual Administration Report on the Railways in India and maps of district and imperial gazetteers.

³⁹ Only two segments were constructed after independence.

⁴⁰ The opening years and the length of the segments that are labeled by their two endpoints are provided in supplementary materials.

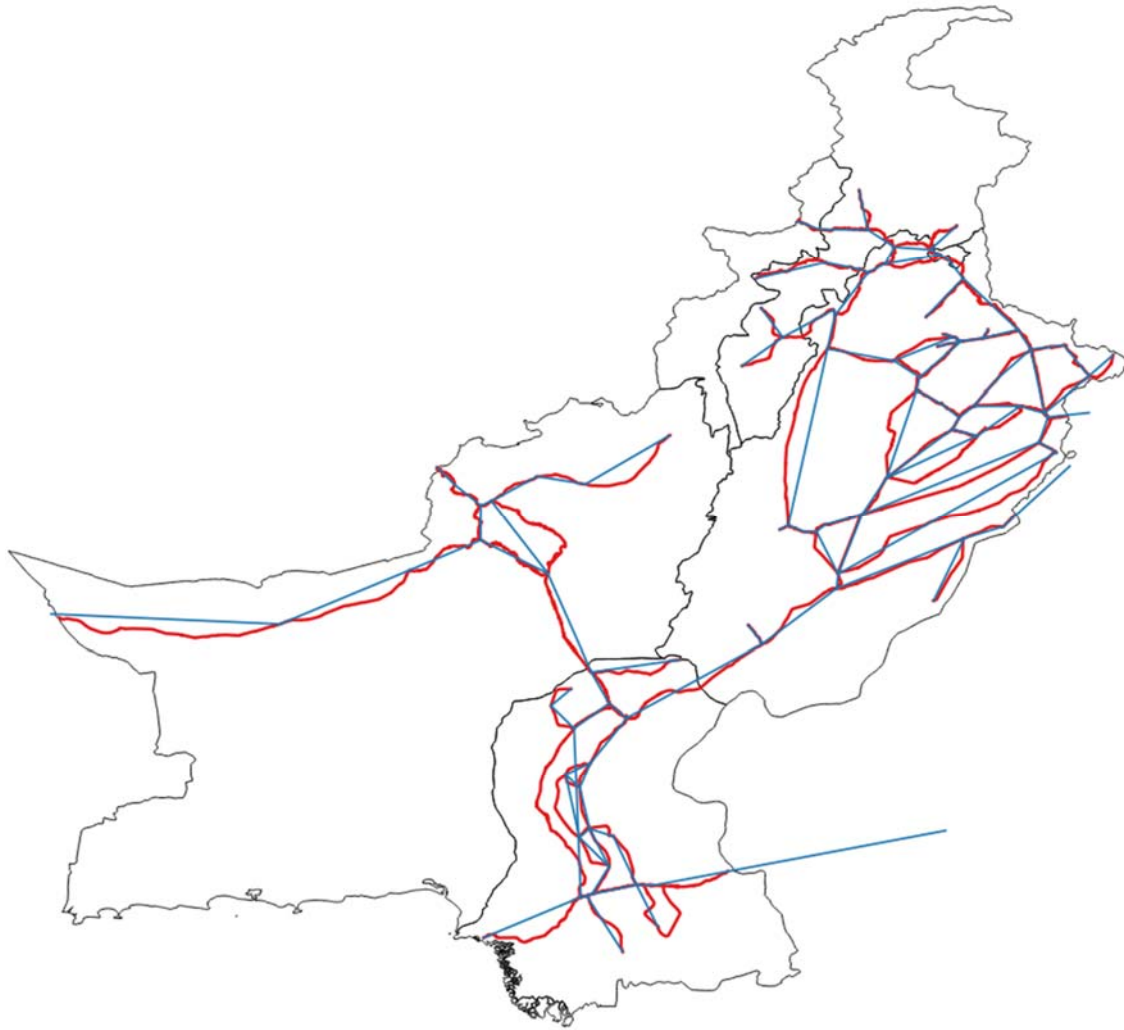
⁴¹ People in a village above age 10 who could read and write in any language.

Figure 1: Railroads Constructed in the Colonial Era



Note: The black lines illustrate railways constructed up to independence in 1947, consisting of 96 percent of the total railroad system of present-day Pakistan. The dots represent villages geo-referenced by using the GPS coordinates of Pakistani villages produced by the National Geospatial-intelligence Agency. The red dots are those villages within a 10km distance (treatment group) and the blue dots are those within 10km to 20km distance to a railroad (control group).

Figure 2: Hypothetical Railroad Track System



Note: A hypothetical railroad track system, consisting of straight lines between every two connected nodes that are defined as either two endpoints or junctions of railway segments, is illustrated by the blue lines. Railways constructed up to independence in 1947 are illustrated by the red lines. The map shown in the background is an administrative map of Pakistani provinces.

Geographic characteristics of the villages are inferred from 3-arc-second-resolution (about one hundred m²) elevation data produced by NASA’s Shuttle Radar Topography Mission (SRTM). We divide the region into 1x1 km grid cells, and calculate the mean altitude and terrain ruggedness (standard deviation of altitudes) within each grid cell. Finally, linking the GPS coordinates of the villages with the closest centroid of the grid cells, we obtain the altitude and ruggedness of each village.

We restrict our sample to villages located within 20km of the railroads. To address the endogeneity issue deriving from short distance to nodes, we also exclude those villages are located within a 10km radius of a node. In Figure 1, we depict the sample: red dots are those villages within a 10km distance (treatment group) and blue dots are those within 10km to 20km distance to a railroad (control group). Table 1 (Panel A) reports descriptive statistics of the employed development outcomes of our sample. Panels B and C in Table 1 report summary statistics of treatment and control groups. The literacy rate is quite low (35%) in the entire sample and slightly higher for the villages which are closer to the railroads. Schooling attainment also decreases with distance to the railroads. Only 10% of houses have “access to piped water” on average, reflecting the very poor provision of piped water. The mean is even lower for villages far from the railroads. The mean of “access to electricity” (59%) is somewhat higher (61%) for villages closer to the railroads. 43% of houses are brick houses.

3.5. Empirical Strategy and Estimation Results

A. Ordinary Least Squares (OLS)

In order to estimate the effect of historic railroad access on contemporary economic outcomes, we start from an OLS estimation. The respective outcome of village i in province p , y_{ip} , is a function of railroad access, D_i , and a vector of covariates, X_i ,

$$y_{ip} = \alpha + \beta D_i + \gamma X_i + \mu_p + \varepsilon_{ip}. \quad (1)$$

We estimate equation (1) with two alternatives for railroad access: a dummy variable equal to 1 if a village is within 10km distance to railroads and a continuous variable indicating distance to railroads. We also add province fixed effect, μ_p , because of differences in colonial policies imposed on provinces⁴². We control for geographical factors, altitude and terrain ruggedness, and distance to the closest river because rivers were important traditional means for transportation and an important input for agriculture. We also control for distance to the closest nodes, mostly were important pre-existing cities.

⁴² The North West of British India was composed of four provinces: Punjab, Sindh, Baluchistan, and the North West Frontier Province. After independence in 1947, number and boundaries of the provinces remained unchanged except for the Islamabad Capital Territory that was separated from Punjab. We consider Islamabad as a part of Punjab.

Table 1: Descriptive Statistics of the Sample

	Panel A				Panel B				Panel C			
	within 20km distance to railroads				within 10km distance to railroads				within 10km to 20km distance to railroads			
	Mean	SD	MIN	MAX	Mean	SD	MIN	MAX	Mean	SD	MIN	MAX
<i>Dependent Variables</i>												
Literacy Rate	0.347	0.175	0	1	0.351	0.175	0	1	0.342	0.174	0	1
Schooling Attainment	0.577	0.312	0	3	0.586	0.316	0	3	0.565	0.306	0	3
Relative Gender Gap in Schooling Attainment	0.321	0.223	0	3.6	0.327	0.221	0	3.6	0.312	0.226	0	2.5
Housing Quality	0.434	0.332	0	1	0.433	0.329	0	1	0.434	0.335	0	1
Access to Piped Water	0.099	0.187	0	1	0.108	0.194	0	1	0.0875	0.176	0	1
Access to Electricity	0.594	0.341	0	1	0.608	0.332	0	1	0.574	0.353	0	1
<i>Control Variables</i>												
Altitude	236.6	296.5	3.18	3222	222.3	267.4	4.23	2716	255.5	330.0	3.18	3222
Ruggedness	3.308	8.975	0	135.3	2.531	5.354	0	111.4	4.331	12.13	0	135.3
Distance to the Closet River in km	27.28	28.52	0.03	495.9	27.21	28.59	0.05	425.6	27.38	28.44	0.03	495.9
Distance to the Closet Node in km	31.11	18.91	10.01	136.3	30.21	19.06	10.01	122.7	32.30	18.65	10.01	136.3
Obs.	23169				13177				9992			

Note: The unit of observation is village. “Schooling attainment” is a continuous variable ranging from 0 to 3. “Relative gender gap in schooling attainment” is computed as female schooling attainments over male schooling attainments. “Access to electricity” and “access to piped water” are the ratio of houses that have access to electricity and piped water to the total number of houses. “Housing quality” is the ratio of brick houses to the total number of houses. Altitude is the mean of altitudes in each 1x1 km grid cells coming from the SRTM. Ruggedness is the standard deviation of altitudes in each 1x1 km grid cells.

Table 2 reports the estimates of equation 1. In all columns, D_i denotes *railroad access*, i.e. the respective entry reports the estimate of β in equation 1, and standard errors are clustered at a higher administrative level than the village level, the so-called union council. Panel A reports the estimation results of OLS specification when railroad access is measured by a dummy variable. All estimates are positive, implying that the mean of the respective outcome is higher for villages closer to the railroads. The average literacy rate in villages that have better access to the railroads is 1% higher than villages that are further away. Taking into account the low average literacy rate in our sample (35%), 1% is economically significant. Access to the railroads slightly increases schooling attainment and the relative gender gap in schooling attainment. The statistically significant and positive effect of railroad access may be explained by non-agricultural opportunities that railroad expansion opened up. Access to railroads also improved living standards as measured by an increase in access to piped water and electricity (2% and 3%, respectively).

Using the linear measure of distance to railroad as a measure of railroad access, Panel B reports similar results. Distance to the railroads has a negative effect on the economic outcomes. Point estimates are very small because they represent changes in the outcomes by going 1km further away from the railroads. For example, 1km increase in distance to the railroads decreases literacy rate by 0.1%, access to piped water by 2%, and access to electricity by 3%. These results are equal to the results of the models that use a binary variable for 10km distance to the railroads.

B. Instrumental Variables (IV)

The main shortcoming of OLS estimates is, of course, that railroad access may be endogenous to economic development: railroad access may be correlated with pre-existing factors that are also correlated with contemporary outcomes. To address this issue, we use an instrumental variable (IV) that has been well established in the literature on the effect transportation infrastructure (started with Michaels (2008) and used by Atack, et al (2010), Banerjee, et al. (2012), Faber (2014), and Hornung (2015)). Actual railroad access is instrumented by access to a hypothetical network of straight lines described in the last section.

Table 3 reports first stage results of the IV strategy in Panel A. The instrument is statistically significant for both the dummy variable and the continuous measure of railroad access, and the first stage F-statistics are high, confirming the power of the instrument. The exclusion restriction would be violated if access to the hypothetical network of straight-lines were correlated with pre-existing factors that may have had influence on contemporary outcomes. Because of the lack of pre-existing data, we estimate as placebo falsification test the correlation between straight-line access and geographic factors. Neither altitude nor distance to river is correlated with the IV. The correlation between terrain ruggedness and the IV is statistically significant, but small considering the mean and standard deviation of terrain ruggedness. The overall result suggests the importance of controlling for the geographical factors and distance to the closest river.

Table 2: OLS specification

	Panel A					
	Literacy Rate	Schooling Attainment	Relative Gender Gap in Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity
Railroad Access, dummy for 10km	0.010** (0.005)	0.023** (0.009)	0.011* (0.006)	0.001 (0.009)	0.023*** (0.004)	0.028*** (0.010)
R-squared	0.308	0.284	0.218	0.383	0.029	0.179
	Panel B					
	Literacy Rate	Schooling Attainment	Relative Gender Gap in Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity
Railroad Access, distance	-0.001** (0.0005)	-0.003*** (0.0009)	-0.002*** (0.0006)	-0.001 (0.0009)	-0.002*** (0.0003)	-0.003*** (0.001)
R-squared	0.309	0.285	0.218	0.383	0.031	0.180
Observations	23164	23164	22791	23164	23164	23164

Note: The unit of analysis is village. Sample contains villages located within 20km of the railroad network. All columns include geographical controls for altitude and ruggedness, distance to the closest river, distance to the closest node, and a set of province fixed effects. Standard error reported in the parentheses are clustered at a higher administrative level than the village level, the so-called union council. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

Table 3: First Stage and Falsification Tests

	Panel A: First Stage		Panel B: Falsification Tests		
	Actual Railroad Access		Altitude	Ruggedness	Distance to River
	dummy for 10km	distance			
Straight-line Access, dummy for 10km	0.420*** (0.0179)	-4.934*** (0.2012)	1.039 (9.1671)	-0.742*** (0.2721)	0.827 (0.9748)
Natural Logarithm of Altitude	0.001 (0.0179)	-0.180 (0.2312)			
Natural Logarithm of Ruggedness	-0.019 (0.0126)	0.333** (0.1499)			
Natural Logarithm of Distance to the Closest River	-0.010 (0.0085)	0.105 (0.0927)			
Natural Logarithm of Distance to the Closest Node	-0.024 (0.0201)	0.527** (0.2375)			
Observations	23,164	23,164	23,169	23,169	23,169
R-squared	0.192	0.218	0.453	0.195	0.342
First stage F-Stat	90.73	93.71			

Note: The unit of analysis is village. Sample contains villages located within 20km of the railroad network. All columns include a set of province fixed effects. Standard error reported in the parentheses are clustered at a higher administrative level than the village level, the so-called union council. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

In Table 4, we report the second stage results and the reduced form estimates. The IV estimates (Panel A and B) correspond to the reduced form relationship (Panel C) between distance to the hypothetical network and all development outcomes. The IV estimates remain statistically significant. Even housing quality that was statistically insignificant using OLS specification becomes significant when we use the IV. However, the point estimates are higher, almost twice as high as the OLS estimates. Railroad access increases literacy rate by 2 percentage points, the ratio of houses that have access to piped water by 4 percentage points, and the ratio of houses that have access to electricity by 5 percentage points. This result suggests that track placement was not chosen randomly and that distance to actual railroads was correlated with pre-existing development.

3.6. Conclusion

Hurd (1983) documents short-term economic effects of railroad expansion in British India and proposes some hypotheses. Donaldson (2010), Andrabi and Kuehlwein (2010), and Bogart and Chaudhary (2013) investigate some of the effects anticipated by Hurd (1983). We estimate the long-term effects of railroad expansion on local development. Since track placement was not chosen randomly, we use an instrumental variable approach to address the pertinent endogeneity issue. Both OLS and instrumental variable findings show that railroad expansion had a statistically and economically significant long-term effect. Literacy rate and schooling attainment in villages that are closer to the old railroad system are higher than in villages that are further away. Access to railroads had also a positive effect on living standards as measured by access to piped water, access to electricity, and housing quality. Hurd (1983) also argue that railroad expansion increased inequality in land ownership. This hypothesis can be empirically tested. We leave it for future research.

Reference

- Acemoglu, D., S. Johnson, and J. A. Robinson (2005): “*Institutions as a Fundamental Cause of Long-Run Growth*,” in Handbook of Economic Growth, ed. by P. Aghion and S. N. Durlauf. Amsterdam, The Netherlands: Elsevier North-Holland, 109–139.
- Andrabi, T., and Kuehlwein, M. (2010). “*Railways and price convergence in British India*.” The Journal of Economic History, 70(02), 351-377.
- Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010). “*Did railroads induce or follow economic growth?*” Social Science History 34, no. 02: 171-197.
- Atack, J., and Margo, R. A. (2009). “*Agricultural Improvements and Access to Rail Transportation: The American Midwest as a Test Case, 1850-1860*.” (No. w15520) National Bureau of Economic Research.

Table 4: IV specification and Reduced Form

Panel A: IV specification, Second Stage						
	Literacy Rate	Schooling Attainment	Relative Gender Gap in Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity
Railroad Access, dummy for 10km	0.022* (0.0124)	0.043** (0.0218)	0.026* (0.0146)	0.043** (0.0213)	0.043*** (0.0084)	0.051** (0.0248)
R-squared	0.307	0.283	0.216	0.379	0.027	0.178
Panel B: IV specification, Second Stage						
	Literacy Rate	Schooling Attainment	Relative Gender Gap in Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity
Railroad Access, distance	-0.002* (0.0011)	-0.004** (0.0019)	-0.002* (0.0012)	-0.004** (0.0018)	-0.004*** (0.0007)	-0.004** (0.0021)
R-squared	0.308	0.285	0.218	0.380	0.029	0.180
Panel C: Reduced Form						
	Literacy Rate	Schooling Attainment	Relative Gender Gap in Schooling Attainment	Housing Quality	Access to Piped Water	Access to Electricity
Straight-line Access, dummy for 10km	0.009* (0.0052)	0.018** (0.0091)	0.011* (0.0061)	0.018** (0.0089)	0.018*** (0.0035)	0.021** (0.0104)
R-squared	0.308	0.284	0.218	0.384	0.028	0.179
Observations	23,164	23,164	22,791	23,164	23,164	23,164

Note: The unit of analysis is village. Sample contains villages located within 20km of the railroad network. All columns include geographical controls for altitude and ruggedness, distance to the closest river, distance to the closest node, and a set of province fixed effects. Standard error reported in the parentheses are clustered at a higher administrative level than the village level, the so-called union council. Coefficients significantly different from zero are denoted by: *** 1%, ** 5%, * 10%.

- Banerjee, A., Duflo, E., and Qian, N. (2012). “*On the road: Access to transportation infrastructure and economic growth in China.*” NBER (No. w17897).
- Baum-Snow, N. (2007). “*Did highways cause suburbanization?*” *The Quarterly Journal of Economics*, 122(2), 775-805.
- Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). “*Roads, railroads and decentralization of Chinese cities.*” *Review of Economics and Statistics*, 99(3): 435–448.
- Bogart, D., and Chaudhary, L. (2013). “*Engines of growth: the productivity advance of Indian railways, 1874–1912.*” *The Journal of Economic History*, 73(2), 339-370.
- Bogart, D and Chaudhary, L, (2015). “*Railways in Colonial India: An Economic Achievement?*” In Chaudhary, L, Gupta, B, Roy, T and Swamy, A (Eds), *A New Economic History of India*. Routledge, London, 140–160.
- Dell, M., and Olken, B. A. (2017). “*The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java.*” (No. w24009) National Bureau of Economic Research.
- Donaldson, D. (2017). “*Railroads of the Raj: Estimating the impact of transportation infrastructure.*” *American Economic Review*, forthcoming.
- Donaldson, D., and Hornbeck, R. (2016). “*Railroads and American economic growth: A “market access” approach.*” *The Quarterly Journal of Economics*, 131(2), 799-858.
- Duranton, G., and Turner, M. A. (2012). “*Urban growth and transportation.*” *Review of Economic Studies*, 79(4), 1407-1440.
- Faber, B. (2014). “*Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System.*” *The Review of Economic Studies*, 81, 1046–1070.
- Hornung, E. (2015). “*Railroads and growth in Prussia.*” *Journal of the European Economic Association*, 13(4), 699-736.
- Hurd, J. (1983), “*Chapter VIII.2: Railways,*” in *Cambridge Economic History of India*, volume 2, ed. by D. Kumar. Cambridge: Cambridge University Press.
- Jedwab, R., Kerby, E., and Moradi, A. (2015). “*History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya.*” *The Economic Journal*, 127 (August), 1467–1494.
- Jedwab, R., and Moradi, A. (2016). “*The permanent effects of transportation revolutions in poor countries: evidence from Africa.*” *Review of Economics and Statistics*, 98(2), 268-284.

- MacPherson, W. J. (1955). "*Investment in Indian railways, 1845–1875.*" *The Economic History Review*, 8(2), 177-186.
- Michaels, G. (2008). "*The effect of trade on the demand for skill: Evidence from the interstate highway system.*" *The Review of Economics and Statistics*, 90(4), 683-701.
- Nunn, N. (2009). "*The Importance of History for Economic Development.*" *Annual Review of Economics* 1: 65-92.
- Tiwari, P. (1921) "*Indian railways: Their historical, economical and administrative aspects.*"
Shiwahare.

Bibliography

- Acemoglu, D., Chaves, I. N., Osafo-Kwaako, P., and Robinson, J. A. (2014). “*Indirect rule and state weakness in Africa: Sierra Leone in comparative perspective.*” No. w20092. National Bureau of Economic Research.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2001): “*The Colonial Origins of Comparative Development: An Empirical Investigation.*” *American Economic Review*, 91, 1369–1401.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2002). “*Reversal of fortune: Geography and institutions in the making of the modern world income distribution.*” *The Quarterly journal of economics*, 117(4), 1231-1294.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2005): “*Institutions as a Fundamental Cause of Long-Run Growth,*” in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf. Amsterdam, The Netherlands: Elsevier North-Holland, 109–139.
- Acemoglu, D., T. Reed, and J. A. Robinson (2014). “*Chiefs: Economic development and elite control of civil society in Sierra Leone.*” *Journal of Political Economy*, 122(2), 319-368.
- Alsan, M. (2015), “*The effect of the tsetse fly on African development*”, *American Economic Review* 105, 382-410.
- Andrabi, T., and Kuehlwein, M. (2010). “*Railways and price convergence in British India.*” *The Journal of Economic History*, 70(02), 351-377.
- Angeles, L. and A. Elizalde (2017). “*Pre-colonial institutions and socioeconomic development: The case of Latin America.*” *Journal of Development Economics* 124: 22-40.
- Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010). “*Did railroads induce or follow economic growth?*” *Social Science History* 34, no. 02: 171-197.
- Atack, J., and Margo, R. A. (2009). “*Agricultural Improvements and Access to Rail Transportation: The American Midwest as a Test Case, 1850-1860.*” (No. w15520) National Bureau of Economic Research.
- Baha, L. (1978). “*N.-WFP Administration under British Rule, 1901-1919.*” National Commission on Historical and Cultural Research.
- Bandyopadhyay, S., and E. Green (2016). “*Precolonial Political Centralization and Contemporary Development in Uganda.*” *Economic Development and Cultural Change*, 64(3), 471-508.

- Banerjee, A., Duflo, E., and Qian, N. (2012). “*On the road: Access to transportation infrastructure and economic growth in China.*” NBER (No. w17897).
- Banerjee, A., and L. Iyer, (2005). “*History, institutions, and economic performance: the legacy of colonial land tenure systems in India.*” *The American Economic Review*, 95(4), 1190-1213.
- Baum-Snow, N. (2007). “*Did highways cause suburbanization?*” *The Quarterly Journal of Economics*, 122(2), 775-805.
- Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). “*Roads, railroads and decentralization of Chinese cities.*” *Review of Economics and Statistics*, 99(3): 435–448.
- Becker, S. O., K. Boeckh, C. Hainz, and L. Woessmann, (2016). “*The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy.*” *The Economic Journal*, 126(590), 40-74.
- Berry, W. (1966). “*Aspects of the Frontier Crimes Regulation in Pakistan.*” (No. 3) Duke University, Program in Comparative Studies on Southern Asia.
- Bogart, D., and Chaudhary, L. (2013). “*Engines of growth: the productivity advance of Indian railways, 1874–1912.*” *The Journal of Economic History*, 73(2), 339-370.
- Bogart, D and Chaudhary, L, (2015). “*Railways in Colonial India: An Economic Achievement?*” In Chaudhary, L, Gupta, B, Roy, T and Swamy, A (Eds), *A New Economic History of India*. Routledge, London, 140–160.
- Caroe, O. (1958). “*The Pathans: 550 bc-ad 1957.*” MacMillan; St Martin's Press.
- Callen, M., Gulzar, S., Rezaee, A., and Shapiro, J. (2017). “*Choosing Ungoverned Space: Pakistan's Frontier Crimes Regulation*”, Working Paper.
- Contini, D. (2013). “*Immigrant background peer effects in Italian schools.*” *Social Science Research*, 42(4), 1122-1142.
- Davies, C. C. (1932). “*The problem of the North-West frontier, 1890-1908.*” Cambridge University Press.
- Dell, M. (2010). “*The Persistent Effects of Peru's Mining Mita.*” *Econometrica*, 78(6), 1863-1903.
- Dell, M., N. Lane, and P. Querubin, (2017). “*The Historical State, Local Collective Action, and Economic Development in Vietnam.*” (No. w23208) National Bureau of Economic Research.
- Dell, M., and Olken, B. A. (2017). “*The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java.*” (No. w24009) National Bureau of Economic Research.

- Donaldson, D. (2017). "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, forthcoming.
- Donaldson, D., and Hornbeck, R. (2016). "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics*, 131(2), 799-858.
- Duranton, G., and Turner, M. A. (2012). "Urban growth and transportation." *Review of Economic Studies*, 79(4), 1407-1440.
- Entorf, H., and Minoiu, N. (2005). "What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration." *German Economic Review*, 6(3), 355-376.
- Faber, B. (2014). "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System." *The Review of Economic Studies*, 81, 1046–1070.
- Fenske, James. 2013. "Does Land Abundance Explain African Institutions?" *Economic Journal* 123(573): 1363–90.
- Gelman, A., and Imbens, G. (2014). "Why high-order polynomials should not be used in regression discontinuity designs." (No. w20405) National Bureau of Economic Research.
- Gennaioli, N. and I. Rainer. 2007. "The Modern Impact of Precolonial Centralization in Africa." *Journal of Economic Growth* 12 (3): 185–234.
- Gruber, J. (1994). "The incidence of mandated maternity benefits." *The American Economic Review*, 622-641.
- Guiso, L., P. Sapienza, L. Zingales (2008). "Social Capital as Good Culture." *Journal of the European Economic Association*; 6 (2-3): 295-320.
- Hornung, E. (2015). "Railroads and growth in Prussia." *Journal of the European Economic Association*, 13(4), 699-736.
- Huffman, S. (2015). "Mapping the genetic relationships of the world's languages." Retrieved from http://www.worldgeodatasets.com/files/8113/0193/1121/Mapping_Genetic_Relationships_World_Languages.pdf.
- Hurd, J. (1983), "Chapter VIII.2: Railways," in *Cambridge Economic History of India*, volume 2, ed. by D. Kumar. Cambridge: Cambridge University Press.
- Imbens, G., and K. Kalyanaraman (2012) "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79 (3): 933-59.
- IRVAPP (2015). "La riforma del maestro unico," *Il Mulino*, forthcoming.

- Iyer, L. (2010). “*Direct versus indirect colonial rule in India: Long-term consequences.*” *The Review of Economics and Statistics*, 92(4), 693-713.
- Jedwab, R., Kerby, E., and Moradi, A. (2015). “*History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya.*” *The Economic Journal*, 127 (August), 1467–1494.
- Jedwab, R., and Moradi, A. (2016). “*The permanent effects of transportation revolutions in poor countries: evidence from Africa.*” *Review of Economics and Statistics*, 98(2), 268-284.
- La Porta R, Lopez-de-Silanes F, Shleifer A, and Vishny R. (1997). “*Legal determinants of external finance.*” *Journal of Finance*. 52:1131–50.
- La Porta R, Lopez-de-Silanes F, Shleifer A, and Vishny R. (1998). “*Law and finance.*” *Journal of Political Economy*. 106:1113–55
- Lange, M. (2009). “*Lineages of despotism and development: British colonialism and state power.*” University of Chicago Press.
- Lee, D. S., and Card, D. (2008). “*Regression discontinuity inference with specification error.*” *Journal of Econometrics*, 142(2), 655-674.
- Lee, D. S., and Lemieux, T. (2010). “*Regression discontinuity designs in economics.*” *Journal of economic literature*, 48(2), 281-355.
- MacPherson, W. J. (1955). “*Investment in Indian railways, 1845–1875.*” *The Economic History Review*, 8(2), 177-186.
- Meghir, C. and Palme M. (2005). “*Educational reform, ability, and family background.*” *American Economic Review*, 414-424.
- Michaels, G. (2008). “*The effect of trade on the demand for skill: Evidence from the interstate highway system.*” *The Review of Economics and Statistics*, 90(4), 683-701.
- Michalopoulos, S. and E. Papaioannou. 2013. “*Pre-colonial Ethnic Institutions and Contemporary African Development.*” *Econometrica* 81 (1): 113–52.
- Michalopoulos, S. and E. Papaioannou. 2014. “*National Institutions and Subnational Development in Africa.*” *Quarterly Journal of Economics* 129 (1): 151–213.
- Mathur, Y. B. (1973). “*British administration of Punjab, 1849-75.*” Delhi: Surjeet Book Depot.
- Nichols, R. (2001). “*Settling the frontier: Land, law and society in the Peshawar Valley, 1500—1900.*” Oxford University Press.

- Nichols, R. (2013). *"The Frontier Crimes Regulation: A History in Documents."* Oxford University Press.
- North, D. C. (1990). *"Institutions, institutional change and economic performance."* Cambridge university press.
- Nunn, N. (2009). *"The Importance of History for Economic Development."* Annual Review of Economics 1: 65-92.
- Tiwari, P. (1921) *"Indian railways: Their historical, economical and administrative aspects."* Shiwahare.
- Schlotter, M., Schwerdt, G., and Woessmann, L. (2011). *"Econometric methods for causal evaluation of education policies and practices: a non-technical guide."* Education Economics, 19(2), 109-137.
- Schwartzberg, J. E. (1978). *"A Historical Atlas of South Asia."* University of Chicago Press.