Long Term Impact of Indigo Cultivation in British India

Soyra Gune*

10 September 2022

Abstract

This study analyzes the persistent effects of indigo cultivation in British India on human capital, public good access, agricultural employment and law enforcement using a dataset of 169 districts which were formerly part of British India. Indigo departs from previously studied cash crops as the state had a minimal role in its cultivation. Despite including an extensive list of controls in the OLS regression of the outcome variables on indigo cultivation, the estimates still could be suffering from endogeneity bias arising from non-random selection of districts into indigo cultivation and measurement error in the colonial data. To control for this endogeneity, I instrument indigo cultivation with a novel indigo suitability index. I find a positive persistent effect of indigo cultivation on agricultural employment most likely due to the failure of the indigo industry to develop production linkages with domestic industries. At the same time, there is a negative, persistent effect of indigo cultivation on number of schools in 2011 since labour intensive agriculture does not require development of human capital. The limited role of the state is reflected in the insignificant relationship between indigo cultivation and public good access. Lastly, the efforts of the planters to strengthen law enforcement to ensure honouring of indigo contracts does not translate into persistent effects on either police or crime in 1911. However, a negative, persistent effect of indigo cultivation is observed on police in 2011 suggesting that indigo districts are not preferred for security owing to them being agriculturally underdeveloped.

*An earlier version of this paper was submitted as the thesis requirement in the MSc Economics program at the University of Warwick. The thesis was completed under the supervision of Dr. Yannick Dupraz.

1 Introduction

Cash crop cultivation for export purposes is a defining characteristic of several colonial regimes. Often, the nature of cash crop cultivation influenced the type of institutions which arose. These institutions could have important implications for economic development (Vogel, 1994). However, the literature has found mixed results with regards to whether the persistent effects of cash crops are beneficial for economic development. Within these studies, the ability of cash crops to produce beneficial results depends on the extent to which it developed linkages with domestic industries (Hirschman, 1958) and the the role of the state in the cultivation of the crop.

Cash crops previously studied in the literature involve a high degree of state intervention either through licenses (Lehne, 2019) or agreements with planters (Dell and Olken, 2020). Indigo cultivation in British India departs from previously studied crops as it was largely cultivated by private planters who entered into contracts with peasants (Chattopadhay and Mamoon, 2009). The indigo dye manufactured from the crop was then exported to Europe to be used in the textile industry (Asiaticus, 1912). Additionally, direct contracts between entrepreneurs and peasants were earlier unheard of and considered risky in India's colonial land tenure system (Bhattacharya, 1977). Therefore, indigo could impact development through three channels. First, the minimal role of the state could indicate that the state did not divert resources away from public goods to aid indigo cultivation. Second, while the indigo dye was manufactured locally, the planters did not attempt to create linkages with other domestic industries or improve the quality of indigo (Kumar, 2012). Consequently, indigo cultivation did not create substantial learning opportunities in manufacturing which could have implications for agricultural employment. Third, to ensure that the peasants were honouring their indigo contracts, the planters strengthened the police and courts (Chaudhuri, 2008). Accordingly, there could be persistent effects on police and crime. Figure 1 summarizes the channels through which indigo cultivation could create persistent effects. Furthermore, indigo is no longer cultivated on a commercial scale anywhere in world (Nadri, 2016) which increases confidence in the fact that any effects observed can be associated with colonial indigo cultivation.

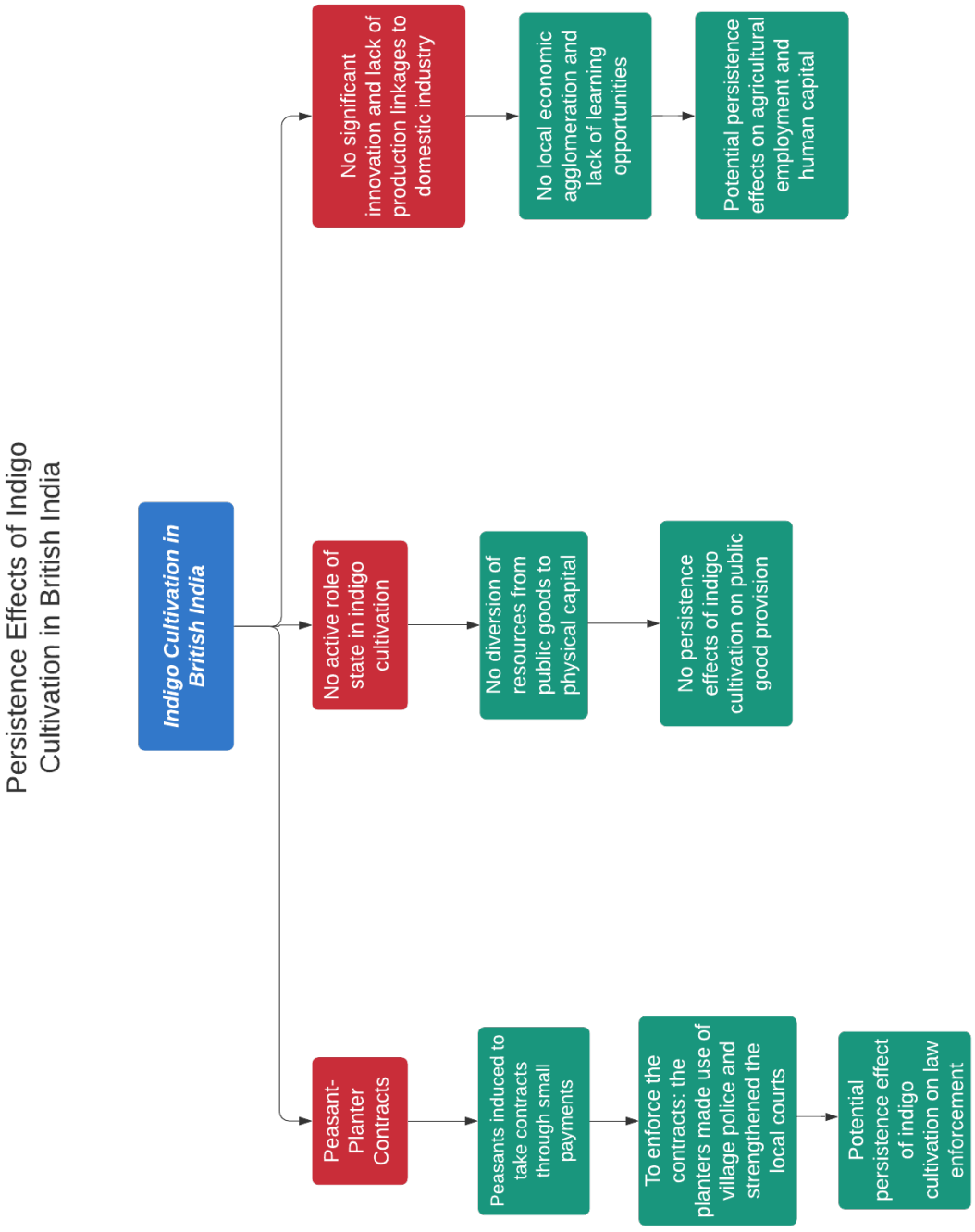

Figure 1: Channels through which indigo cultivation could create persistent effects.

To investigate whether indigo cultivation affects development through the aforementioned channels, this study will look at the persistent effects of indigo cultivation on public goods, human capital, agricultural employment and law enforcement in 1911 and 2011. For this purpose, I construct a dataset of 169 districts which were part of former British India. I begin the analysis by considering an OLS model where I regress the outcome variables on percentage of cultivated land under indigo in 1897. Other than a weakly significant association between indigo cultivation and education investment in 1911, the OLS estimates do not indicate that indigo cultivation has persistent effects. Despite including an extensive list of controls, the OLS estimates could be suffering from endogeneity bias arising from non-random selection into indigo cultivation and measurement error in the colonial data. I attempt to control for this endogeneity by instrumenting indigo cultivation in 1897 with a novel indigo suitability index. The 2SLS results indicate that the OLS estimates were biased towards zero.

I find a positive persistent effect of indigo cultivation on percentage of population employed as agricultural workers in 1911 and 2011. While there is no effect of indigo cultivation on literacy in 1911 and 2011, there is a negative persistent effect on number of schools per 10,000 in population in 2011. This suggests that there was a decline in the importance given to human capital development most likely due to the increased focus on labour intensive agriculture because of the high labour intensity of indigo cultivation. Additionally, the drop in number of schools was not caused by a change in education investment as the result for colonial educational investment in 1911 is insignificant. Simultaneously, I find no persistent effect of indigo cultivation on either investment in civil works in 1911 or access to public goods in 2011 (except for bank access in 2011) indicating that the minimal role of the state plausibly prevented a diversion of resources from public goods. Finally, the efforts of the planters to strengthen law enforcement to ensure that the indigo contracts were followed does not create persistent effects on either police or crime in 1911, though, there is negative, persistent effect of indigo cultivation on police in 2011. I postulate that the negative relationship between indigo cultivation and police in 2011 is potentially caused by labour intensive agriculture not being prioritized for security as more agriculturally developed districts might have a greater demand for security. I perform robustness checks to ensure that the results are not sensitive to different model specifications.

This study is divided into 9 sections: section 2 reviews the literature, sections 3 discusses the data, section 4 provides indigo cultivation's historical background, section 5 covers the OLS model and results, section 6 discusses the endogeneity bias in the OLS estimates, section 7 provides the 2SLS model and results along with robustness checks, section 8 discusses the results and section 9 concludes.

2 Literature Review

Broadly, this study contributes to the literature which explores the relationship between colonial institutions and economic development. Across countries, Acemoglu et al. (2001) and Easterly and Levine (2004) establish that settler mortality impacts colonial institutions which subsequently explain differences in income. Additionally, Nunn (2008) finds a negative relationship between slave trade participation and development. Dell (2010) focusing on the mining *mita* in Peru, shows a negative effect of the *mita* on household consumption through the channels of land rights and public good provisioning.

A subsection of this literature covers the colonial legacy of India. Banerjee and Iyer (2005) find that historical land tenure systems influence the level of agricultural investment. Similarly, Iyer (2010) shows that areas historically under direct British rule have lesser access to public goods compared to areas under indirect rule [1]. While there might be concern that these persistent effects cannot be reversed, Chaudhary and Garg (2015) show that colonial education investments only impact literacy till 1971 after which policy mutes the effect. Furthermore, certain papers examine how location-specific advantages such as health missions (Calvi and Mantovelli, 2018) and colonial railways (Donaldson, 2018) benefit economic development.

Finally, this study contributes to the literature on the impact of cash crop cultivation on economic development. Cash crop cultivation is a form of commodity production and therefore, it can effect development through its success in developing production linkages with

[1] Indirect rule refers to princely states which had some autonomy with regards to administration from the colonial state.

domestic industries (Hirschman, 1958). Through these linkages, it is able to spillover the benefits of cash crop cultivation to the rest economy. Dell and Olken (2020) find that the sugar industry in colonial Java led to economic agglomeration through the development of transport infrastructure which allowed for spillover effects to places close to sugar-producing sites. Subsequently, areas close to former sugar production sites are more industrialized and have higher income. In contrast, Roessler et al. (2020) find that the smallholder cash crop cultivation in West Africa by concentrating on commodity exports meant for Europe failed to develop production linkages with domestic industries. Therefore, while areas involved in cash crop cultivation observe higher levels of development, the spillovers effects to other areas are limited.

Additionally, cash crop cultivation can influence the type of colonial investment put in place to aid the extraction of the crop. Lehne (2019) finds that opium cultivation in British India led to a diversion of resources from human to physical capital. While the increased investment in physical capital does not have any persistent effects, the lack of investment in human capital is reflected through the negative relationship between opium cultivation and human capital. Furthermore, other than colonial investments, cash crops might affect the colonial institutions established to assist in the cultivation of the crop. Sokoloff and Engerman (2000) find that South America's suitability for cash crops over North America led to the creation of unequal land rights in South America. Similarly, Naritomi et al. (2012) show that Brazil's sugarcane boom is associated with higher land inequality because of extractive institutions. In comparison, Austin (2009) finds that the African cash crop revolution caused slavery's decline as the slave trade's violence disrupted cash crop exports. However, cash crop cultivation supporting institutions might come into conflict with pre-colonial institutions. Fenske (2014) establishes that the commercialization of land arising from the introduction of para rubber in Benin led to disputes over land which were influenced by the region's sociopolitical practices. Likewise, Gupta and Swamy (2017) find that migrant tea plantation workers in colonial Assam were less likely to take coercive contracts when they had information through family networks.

3 Historical Background

The East India Company (EIC) began exporting indigo dye to Europe in 1664 to be used in the textile industry (Asiaticus, 1912). However, when the EIC began mounting losses from the trade, it had to give up its monopoly in 1789 (Kumar, 2012). With the entry of European planters unaffiliated with the company and expertise of planters who had previously grown the crop in West Indies, India became the leading exporter of indigo in 1815 (Nadri, 2016).

3.1 Method of Cultivation

According to Kumar (2012), black, sandy and light soils, availability of water and cheap labour were needed for indigo cultivation. Planters incentivized peasants to take up contracts through advances (Chattopadhay and Mamoon, 2009). Peasants would take these contracts to pay back the loan they owed the zamindar [2] who leased the peasants land (Nadri, 2016) . These peasant-planter contracts were considered risky given India's land rights system (Chaudhuri, 2008). Indigo differed from other small landholding crops because of the power the planters had in the contract. Firstly, the planter decided which part of the peasant's land would be used for indigo (Roy, 2011). Secondly, peasants were cheated when they took the indigo to the factory as the remuneration they received for the crop was below the market price. After the seed and transport costs were deducted from the remuneration, the peasant often ended up in debt bondage to the planter as they were unable to pay back the advance (Bhattacharya, 1977). Thirdly, the planters employed rural police who had considerable power in areas where land rights were unequal to ensure that peasants were honouring their contracts (Chaudhuri, 2008) and strengthened local courts to ensure peasants could be held liable if they failed to honour their contracts (Chattopadhay and Mamoon, 2009).

3.2 Limited Learning Opportunities

The production process remained unchanged from the one used by Mughals previously. Indigo dye production was highly labour intensive from cultivation to factory work and there-

[2] Zamindars were landlords who were legally recognized under the British state under the Permanent Settlement Act of 1793 (Nadri, 2016).

fore, generated a continuous demand for labour (Kumar, 2012). Wages received by the factory workers were lower than what they would have received under the Mughal era. Additionally, no major breakthrough was made in production to improve the quality of indigo. Planters who took up indigo production mainly to make a quick fortune, often lacked the knowledge to improve the indigo's quality and misspent borrowed capital from the agency houses [3] (Nadri, 2016) . Finally, indigo dye was supplied only to European industries and therefore, the industry did not have any production linkages with local industries (Bhattacharya, 1977).

Because of the lack of innovation, Indian indigo was unable to compete with synthetic indigo when it was introduced to the world market in 1911 and subsequently, indigo is not grown on a commercial scale anywhere in the world (Nadri, 2016).

4 Data

The dataset used for this study consists of 169 districts formerly part of British India which fall under 21 states of India. [4]

4.1 Historical Data on Crop Cultivation

The Agricultural Statistics for India (1897-1902) provides district-level data on indigo cultivation for both British India and the princely states . Data on princely states[5] is limited and therefore, this study will focus on indigo cultivation in British India. However, comparing the results of British India to those of princely states would have allowed for discerning how much of the effect of indigo cultivation is through colonial institutions rather than the nature of the crop since princely states had some control over their administration. Additionally, the

[3] Agency houses were banks set up by the East India Company to function as banks for European investors whose needs could not be met by indigenous banks (Asiaticus, 1912)

[4] The represented states are Andhra Pradesh, Assam, Bihar, Chattisgarh, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Meghalaya, Orissa, Punjab, Rajasthan, Tamil Nadu, Tripura, Uttar Pradesh, Uttarakhand and West Bengal.

[5] Princely states are states in colonial India that had some administrative autonomy from the British government (Banerjee and Iyer, 2005).

Agricultural Statistics provides information on land use for other crops (rice, cotton, sugarcane, wheat, tobacco). To obtain the percentage of cultivated land under indigo, I divide the amount of land under indigo by the total cultivated land in the district and multiply it by one hundred.

From figure 2, it is evident that indigo cultivation in British India was concentrated in the provinces of Agra, Oudh, Bengal and Madras which is reflected in the literature (Kumar, 2012). Figure 3 maps indigo cultivation for princely states.

4.2 Outcome Variables

The outcome variables are from 1911 or 2011 and are measures of human capital, public goods, agricultural employment and law enforcement.

4.2.1 1911 Outcome Variables

The sources for 1911 outcome variables are the District Gazetteers and the 1911 colonial censuses. I digitize both these sources into workable datasets for the 169 historic districts. The District Gazetters provide district level data on colonial investments in education and civil works, number of criminals and rural police for either 1910 or 1911. However, the gazetteers for Naga Hills, Shimla, Sagar, Kadapa and Madras are not available. Additionally, data on investments in civil works is missing for Garo Hills, Angul, Mandla, Chanda, Bhandra and Gurgaon. From these districts, only Sagar, Kadapa and Gurgaon grew indigo.

The 1911 colonial censuses for British India provide district level data on literacy, percentage of population employed as agricultural workers and population. I choose agricultural workers here as these workers are often involved in labour intensive agriculture similar to the cultivation process of indigo. I divide colonial investments by district population in 1911 to obtain colonial investment per individual and divide rural police and number of crimes by population and multiply by 10,000 to obtain number of crimes and rural police per 10,000

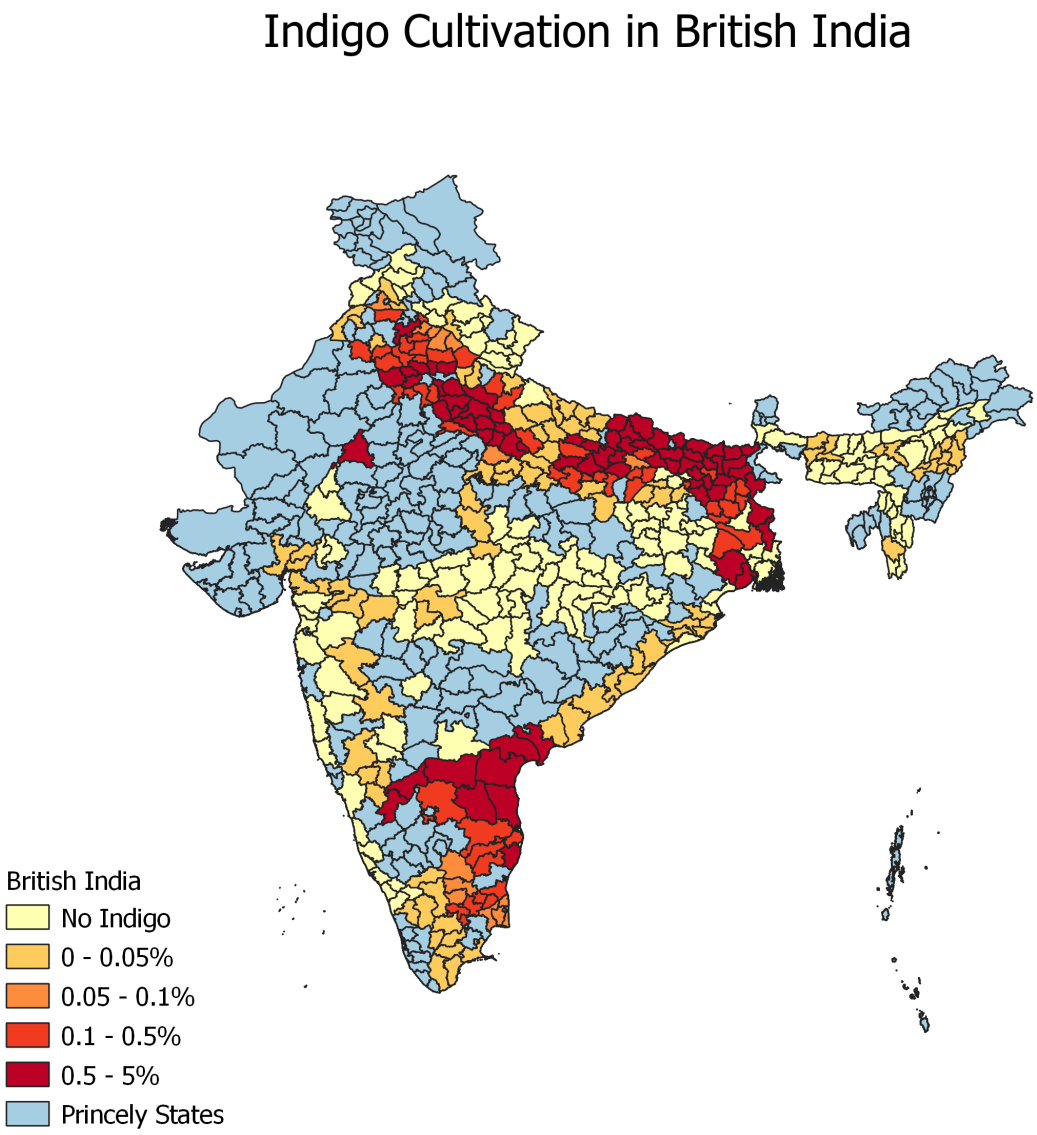

Figure 2: This map reflects the intensity of indigo cultivation for districts that were historically part of districts that grew indigo in British India. Districts marked in red historically were part of districts that had 0.5 to 5 percent of their cultivated land dedicated to indigo. Districts marked in blue were historically part of princely states. Map made using QGIS software and data on indigo cultivation taken from Agricultural Statistics of India (1897-1902).

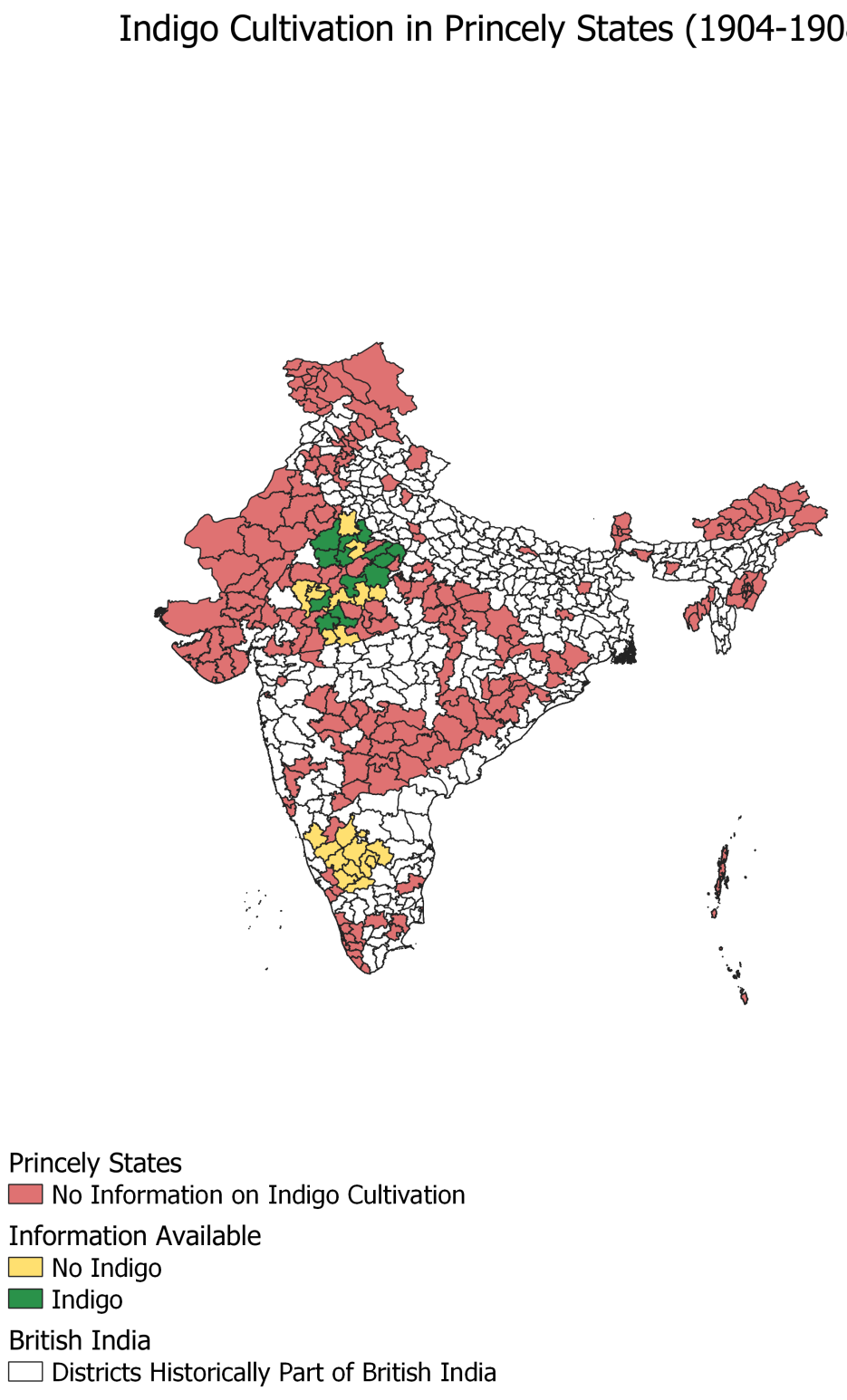

Figure 3: Districts which are marked in green are districts that historically were part of princely states that grew indigo while districts marked in yellow are districts that historically were part of princely states that did not grow indigo. Districts marked in maroon are districts which are historically part of princely states for which there is no information on indigo cultivation. Map made using QGIS software and data on indigo cultivation taken from Agricultural Statistics of India (1904-1908).

individuals in population.

4.2.2 2011 Outcome Variables

The 169 historical districts in 1897 cover 314 contemporary districts of India in 2011. A district in 1897 maps to one or more contemporary districts in 2011. The Administrative Atlas of India (2011) provides district level maps of India from 1891 to 2011. I employ the Atlas to map the 1897 district to the 2011 districts. The crosswalking method (Kumar and Somanathan, 2017) requires adding up the population of the contemporary districts in 2011 to find the population of the historic district in 2011. Then the variable can be aggregated to the historic district level by using the percentage of the historic district's population that comes from a particular contemporary district. For example, the historic district of Kanpur presently contains the districts of Kanpur Dehat and Kanpur Nagar. Kanpur Dehat contains 47.87 percent of Kanpur's 2011 population and Kanpur Nagar contains 52.13 percent of Kanpur's 2011 population. Kanpur Dehat's literacy in 2011 was 77.52 percent and Kanpur Nagar's literacy was 79.65 percent. To obtain the literacy rate for Kanpur in 2011, I multiply 0.4787 with 77.52 and multiply 0.5213 with 79.65 and then add the products to obtain the literacy of Kanpur in 2011 which was 78.62 percent. [6]

The District Census Handbooks for 2011 provide district level data on number of educational institutions per 10,000, percentage with access to public goods (medical care, drinking water, banks, pucca[7] road, public transport, agricultural credit, power supply), literacy rate, percentage of agricultural workers and population. I extract data on crime in 2011 from the National Crime Bureau database. Since there is no available data source on rural police, I substitute this with the district level constables data for 2011 from the Bureau of Police Research and Development. Constables hold the lowest police rank and therefore, can be expected to perform similar functions to rural police. Again, I divide the criminal offences and number of constables with the population in 2011 and multiply by 10,000 to obtain the per 10,000 individuals variables.

[6] Please refer to tables A1 to A7 in the appendix for the mapping of all historic districts in 1897 to contemporary districts in 2011.

[7] Pucca road refers to tarred roads in India.

4.3 Control Variables

The control variables can be divided into three categories: development, social and geographic.

4.3.1 Development and Social Controls

The district level social controls are taken from the colonial censuses of 1911 and broadly cover religion, caste and sex. The development controls include railways and proportion of population that is rural. The proportion of rural population is calculated using rural population data for 1911 from the colonial censuses and the railway dummy which equals one if the district had railways in 1911 is taken from Donaldson (2018). Donandlson's dataset does not contain data for the districts of Almora, Nainital, Ranchi and Washim.

4.3.2 Geographic Controls

The geographic variables mentioned here will be used both as controls in the OLS and to construct the indigo suitability index which will be discussed later. Fick and Hijmans (2017) provide 2.5 arc minute grid-level data on 19 bioclimatic variables. Furthermore, raster data on elevation and inland water is taken from the DIVA GIS database. Data on various soil variables is taken from the Food and Agricultural Organization database and latitude and longitude data is extracted from the World Bank's agriculture database. All geographic variables are aggregated to the contemporary district level using QGIS by taking the mean of the grid cells that fall under the contemporary district boundary. Then using the cross walking method , I aggregate the geographic data from the contemporary districts to the historic district. Using the Administrative Atlas, I construct a coastal dummy which equals one if the historic district was a coastal district.

4.4 Summary Statistics

Table 1 provides the summary statistics for the 169 districts. 61.8 percent of the districts were cultivating indigo in 1897. To help illustrate differences, table 2 and 3 divides the variables by indigo and non-indigo districts. From table 2, on average, it is seen that indigo districts in 1911 had lower literacy, crime, colonial investment and percentage of agricultural workers. Additionally, from table 2, it can be seen that indigo districts in 2011, on average, have lower literacy, crimes and number of schools and higher access to certain public goods, percentage of agricultural workers and more constables per individual relative to non-indigo districts. Table 3 performs a similar exercise for geographical variables. Indigo districts, on average, have sandier soil, less clay, more inland water sources, lower elevation, higher mean temperature and lower annual precipitation relative to non-indigo districts.

5 OLS Specification and Results

5.1 OLS Model

Below (eq 1) is the OLS specification:

$$y_{is} = \beta indigo_{is} + \theta_s + X_{is} + \epsilon_{is} (1)$$

Here, y_{is} is a measure of human capital, public goods, agricultural employment or law enforcement in 1911 or 2011 for district i in state s. $indigo_{is}$ is the percentage of cultivated land under indigo in 1897. β measures the change in the outcome variable associated with a one percentage point increase in percentage of cultivated land under indigo in 1897. θ_s is state level fixed effects. X_{is} is a vector of district level social, geographic and development controls.

The social controls broadly control for caste, sex and religion. Districts might have differences in demands for colonial investment and public goods based on the social composition of the district making it necessary to include social controls. Colonial officials might target investments to more developed districts and therefore, I include development controls at the

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Indigo Dummy	0.618	0.487	0	1	169
Indigo Percent	0.401	0.884	0	4.407	169
Indigo Suitability Index	5.58	8.987	2.037	94.896	169
1911 Variables					
Literacy	5.437	4.443	0.112	34.922	169
Crimes per 10,000	31.153	63.082	0.104	578.705	163
Educational Investment per Individual	0.139	0.456	0.003	4.425	164
Civil Works Investment per Individual	0.191	0.51	0.001	4.225	162
Rural Police per 10,000	20.734	21.841	0	125.836	162
Population	1054954	630634.5	39320	3201180	169
Share of SCs/STs in Population	0.146	0.086	0.002	0.501	169
Share of Brahmins in Population	0.066	0.074	0.002	0.428	169
Rail Dummy	0.88	0.327	0	1	166
Sex Ratio	1.055	0.136	0.447	1.693	169
Share of Hindus in Population	0.78	0.2	0.04	0.99	169
Share of Muslims in Population	0.124	0.127	0.001	0.722	169
Share of Christians in Population	0.011	0.023	0	0.146	169
Proportion Rural	0.754	0.247	0.106	1	169
Proportion Urban	0.246	0.247	0	0.894	169
Percentage Agricultural Workers	17.737	18.944	0.085	91.862	169
2011 Variables					
Number of Crimes per 10,000	14.834	9.829	0	83.166	169
Number of Schools per 10,000	11.42	4.401	0	31	169
Medical Care Access	64.349	21.059	0	100	169
Drinking Water	97.712	11.535	0	100	169
Post Office	45.419	20.947	0	100	169
Telephone	86.47	17.443	0	100	169
Transport	64.282	24.734	0	100	169
Banks	19.68	13.21	0	97.497	169
Agricultural Credit	27.57	22.802	0	97.39	169
Pucca Road	77.542	20.492	0	100	169
Power Supply	93.673	14.878	0	100	169
Population	2789463.411	1687400.354	142004	11060148	169
Literacy	73.141	8.779	48.485	94.097	169
Percentage of Cultivators	25.164	11.718	0.2	69.41	169
Percentage of Agricultural Workers	31.623	14.198	0.46	64.38	169
Percentage of Household Workers	8.194	6.62	1.11	36.12	169
Constables per 10,000	13.3	5.68	6.407	37.776	169

district level which include proportion of population that is rural and a railways dummy. State fixed effects control for unobserved differences at the state level. Here, there was an alternative of using province fixed effects in place of state but it is unlikely that the reorganization of states in 1956 which divided provinces into states is correlated with variation in indigo cultivation [8]. Additionally, bias could arise in the estimates if districts were selected for indigo cultivation based on geographic suitability as geographically endowed districts might be preferred for economic development. Therefore, to minimize this bias, I include a set of geographic controls that chiefly include precipitation, temperature, elevation, latitude, longitude, soil variables and inland water. I additionally include a coastal dummy control

[8]As a robustness check, I replace state fixed effects with province fixed effects. The same is report in Table 16 and 17.

Table 2: Summary statistics for 1911 and 2011 Variables by Indigo and Non-Indigo Districts

Variable	Mean	Std. Dev.	N	Mean	Std. Dev.	N
	Non-Indigo			Indigo		
1911 Variables						
Literacy 1911	6.546	6.099	64	4.761	2.853	105
Crimes per 10,000	35.679	81.996	61	28.446	48.671	102
Education Investment per Individual	0.162	0.415	62	0.126	0.481	102
Civil Works Investment per Individual	0.197	0.379	60	0.187	0.575	102
Rural Police per 10,000	27.248	31.132	59	17.002	12.778	103
Population 1911	815436	550459	64	1200946	634085.5	105
Share of SCs/STs in Population	0.11	0.086	64	0.167	0.079	105
Share of Brahmins in Population	0.056	0.077	64	0.072	0.071	105
Rail Dummy	0.738	0.444	61	0.962	0.192	105
Sex Ratio	1.043	0.127	64	1.062	0.141	105
Share of Hindus in Population	0.745	0.23	64	0.802	0.176	105
Share of Muslims in Population	0.098	0.132	64	0.14	0.122	105
Share of Christians in Population	0.016	0.033	64	0.007	0.013	105
Proportion Rural	0.632	0.29	64	0.828	0.183	105
Proportion Urban	0.368	0.29	64	0.172	0.183	105
Percentage Agricultural Workers	24.754	23.617	64	13.459	13.896	105
2011 Variables						
Crimes per 10,000	17.299	11.976	64	13.308	7.91	105
Medical Care Access	62.629	21.332	64	65.415	20.92	105
Drinking Water	95.823	17.475	64	98.882	4.968	105
Post office	42.124	20.647	64	47.459	20.968	105
Telephone	85.958	19.364	64	86.787	16.229	105
Transport	64.027	26.085	64	64.44	23.986	105
Banks	18.898	15.635	64	20.164	11.512	105
Agricultural Credit	27.742	25.808	64	27.463	20.855	105
Pucca Road	74.301	22.673	64	79.548	18.852	105
Power Supply	91.917	19.825	64	94.759	10.71	105
Population 2011	2373792	2020141	64	3046783	1393085	105
Literacy	77.144	7.794	64	70.663	8.465	105
Percentage Cultivators	25.911	14.04	64	24.701	10.063	105
Percentage Agricultural Workers	28.634	16.04	64	33.474	12.658	105
Percentage Household Workers	6.84	6.472	64	9.032	6.602	105
Constables per 10,000	15.263	5.701	64	12.084	5.34	105
Number of Schools per 10,000	12	5.38	64	11.061	3.651	105

since coastal districts might be both preferred for colonial investment and indigo cultivation. Finally, I cluster standard errors at the district level to allow for correlation at the district level.

5.2 OLS Results

Table 4 reports the results of the OLS regressions of outcomes relating to human capital. Focusing on column 1 of the the table which reports the results for literacy in 1911 while controlling for state fixed effects, there appears to be a significant, negative correlation between literacy in 1911 and indigo cultivation with a coefficient of -0.475. This coefficient does not change in size when social controls are added as seen in Column 2. However, when

Table 3: Summary statistics by Indigo and Non-Indigo District for Geographic Variables

Variable	Mean	Std. Dev.	N	Mean	Std. Dev.	N
	Non-Indigo			Indigo		
Latitude	22.675	4.584	64	23.912	5.51	105
Longitude	81.383	6.129	64	80.376	4.151	105
Percentage of Sandy Soil	43.713	18.891	64	51.122	18.824	105
Percentage of Clay Soil	35.506	14.638	64	25.268	12.261	105
Percentage of pH Water in Soil	6.623	1.091	64	6.774	0.87	105
Percentage of Organic Carbon in Soil	1.082	0.718	64	0.763	0.38	105
Percentage of Nitrogen in Soil	0.122	0.067	64	0.149	0.101	105
Percentage of BS in Soil	72.836	25.22	64	84.629	14.091	105
Cation Exchange Capacity of Soil	22.761	15.035	64	18.053	11.272	105
Cation Exchange of Clay Soil	54.419	19.498	64	56.121	12.739	105
Silt Soil	20.339	7.441	64	22.524	11.724	105
Percentage of CaCO3 in Soil	1.453	1.805	64	1.411	2.326	105
Bulk Density of Soil	1.344	0.237	64	1.38	0.185	105
Soil Curve Number	13.209	3.669	64	11.045	3.01	105
Major Inland Water Source	0.983	1.013	64	1.245	1.395	105
Annual Mean Temperature	24.663	2.761	64	25.483	1.445	105
Mean Temperature of the Warmest Quarter	29.744	3.139	64	31.19	1.671	105
Mean Temperature of the Coldest Quarter	19.069	3.784	64	18.409	3.307	105
Annual Precipitation	1651.94	873.74	64	1012.29	328.167	105
Precipitation of Wettest Month	466.305	275.258	64	289.538	86.124	105
Precipitation of Driest Month	4.944	3.941	64	3.698	3.455	105
Precipitation Seasonality	119.31	20.085	64	121.586	20.717	105
Precipitation of Wettest Quarter	1136.37	600.714	64	713.982	220.515	105
Precipitation of Driest Quarter	32.062	24.572	64	26.796	16.094	105
Precipitation of Warmest Quarter	454.692	546.495	64	267.978	202.837	105
Precipitation of Coldest Quarter	124.632	389.209	64	57.747	59.997	105
Isothermality	44.289	5.871	64	43.354	6.096	105
Temperature Seasonality	429.452	135.157	64	520.73	163.559	105
Maximum Temperature of the Warmest Month	36.96	4.678	64	38.958	2.714	105
Minimum Temperature of the Coldest Month	11.303	3.782	64	10.515	4.07	105
Temperature Annual Range	25.657	5.354	64	28.443	5.773	105
Mean Temperature of Wettest Quarter	26.394	2.314	64	28.211	1.585	105
Mean Temperature of Driest Quarter	21.289	4.013	64	24.127	3.513	105
Elevation	405.112	405.239	64	206.268	172.28	105
Coastal Dummy	0.123	0.331	64	0.124	0.331	105
Rice Percent	36.245	32.633	64	22.483	24.166	105
Wheat Percent	6.125	8.541	64	10.579	9.971	105
Jute Percent	0.447	1.784	64	0.104	0.451	105
Tobacco Percent	0.287	1.286	64	0.36	0.646	105
Cotton Percent	3.97	8.359	64	4.284	7.3	105

geographic and development controls are added the size of the coefficient decreases slightly and loses significance as evident in Column 3 and 4. Column 5-8 perform a similar exercise taking educational investment per individual in 1911 as the outcome. Concentrating on column 8 which includes all controls, a percentage point increase in cultivated land under indigo is associated with a 0.045 rupee decline in education investment per individual in 1911. This result is weakly significant. Column 9-12 take number of schools per 10,000 in population in 2011 as the outcome. Column 12 which includes all the controls reports that a one percentage point increase in cultivated land under indigo is associated with 0.482

more schools per 10,000 in population but this result is statistically insignificant. Column 13 to 16 take literacy in 2011 as the outcome variable and column 16 which includes all the controls reports that a percentage point increase in cultivated land under indigo is associated with a 0.257 percentage point increase in literacy in 2011. This result is insignificant. The coefficient on educational investment in 1911 is surprising considering that the state was not actively involved in indigo cultivation to change educational investment in response to indigo cultivation.

Table 5 provides OLS results of outcomes relating to law enforcement. Columns 1-4 use crimes per 10,000 in 1911 as the outcome and columns 5-8 use rural police per 10,000 in 1911 as the outcome. Concentrating on column 4 and 8 which include all controls, a percentage point increase in cultivated land under indigo is associated with a 9.638 and 1.873 fall in crime and rural police per 10,000 in population respectively. Turning to 2011 outcomes, columns 9-12 take crimes per 10,000 as the outcome variable and columns 13-16 take constables per 10,000 as the outcome variable. From column 12 and 16 which consider all controls, indigo cultivation is not significantly associated with either crimes or constables per 10,000 individuals in population in 2011.

From Table 6, Column 1-4 and 5-8 take percentage of district population employed as agricultural workers in 1911 and 2011 respectively, as the outcomes. From column 1-4, there appears to be a negative correlation between agricultural employment and indigo cultivation with column 4 stating that a percentage point increase in cultivated land under indigo is associated with a 1.6 percentage point decline in percentage employed as agricultural workers. However, this relationship is not significant when all controls are included. In contrast, column 5-8 show that there is a positive correlation between indigo cultivation and agricultural employment in 2011. Nonetheless, this correlation is not significant when all controls are included. Therefore, these results do not indicate a clear persistent relationship between agricultural employment and indigo cultivation.

Table 7 reports the results for outcomes relating to public goods. Broadly, it is evident that there is not a significant association between indigo cultivation and public goods. This is not

	Literacy 1911				Education Investment 1911				Number of Schools 2011				Literacy 2011			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Indigo 1897	-0.475**	-0.502***	-0.300	-0.290	-0.0334***	-0.0323***	-0.0424*	-0.0458*	-0.553	-0.307	0.520	0.482	-1.719***	-1.549**	0.125	0.257
	(-2.21)	(-2.61)	(-1.44)	(-1.56)	(-2.73)	(-2.67)	(-1.97)	(-1.96)	(-1.31)	(-0.85)	(1.37)	(1.22)	(-2.65)	(-2.22)	(0.22)	(0.45)
	[0.215]	[0.192]	[0.208]	[0.185]	[0.012]	[0.012]	[0.021]	[0.023]	[0.422]	[0.358]	[0.378]	[0.394]	[0.649]	[0.696]	[0.570]	[0.567]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Development Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Constant	4.896	-4.413	19.51	1.573	0.018	-0.154	-0.453	-0.201	11.18	13.59	-31.79	-41.84	68.41	63.90	202.2	231.5
	(5.90)	(-1.21)	(0.66)	(0.10)	(0.95)	(-0.75)	(-0.27)	(-0.11)	(16.14)	(3.82)	(-1.32)	(-1.61)	(44.93)	(11.84)	(5.13)	(5.91)
N	169	169	169	165	164	164	164	160	169	169	169	165	169	169	169	165

Table 4: This table presents the results of regressing outcome variables relating to human capital on percentage of land under indigo cultivation in 1897. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are literacy in 1911 (Column 1-4), education investment per individual in 1911 (Column 5-8), schools per 10,000 of population in 2011 (Column 9-12) and literacy in 2011 (Column 13-16). *t* statistics are reported in round parenthesis and standard errors which are clustered at the district level are reported in square parenthesis. Columns 1, 5, 9 and 13 only control for state fixed effects. Columns 2, 6, 10 and 14 control for stated fixed effects and social controls. Columns 3, 7, 11 and 15 control for state fixed effects, social and geographic controls. Columns 4, 8, 12 and 16 control for stated fixed effects and social, geographic and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

	Crime 1911				Rural Police 1911				Crime 2011				Constables 2011			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Indigo_1897	-4.118*	-4.886**	-8.738	-9.638	-1.811	-2.087	-1.873	-0.509	-0.444	-0.291	0.854	1.038	-1.351***	-0.904***	-0.225	-0.14
	(-1.93)	(-2.01)	(-1.46)	(-1.52)	(-1.63)	(-1.64)	(-0.72)	(-0.33)	(-0.78)	(-0.52)	(1.17)	(1.49)	(-3.58)	(-2.64)	(-0.50)	(-0.31)
	[2.13]	[2.431]	[5.999]	[6.347]	[1.11]	[1.271]	[2.615]	[1.552]	[0.567]	[0.555]	[0.728]	[0.696]	[0.376]	[0.342]	[0.452]	[0.457]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Development Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Constant	0.000	-0.003	0.061	0.059	0.002	0.001	-0.014	-0.004	0.000	0.001	-0.005	-0.005	0.001	0.001	0.007	0.006
	(1.02)	(-1.06)	(1.71)	(1.65)	(4.98)	(1.68)	(-1.29)	(-0.39)	(5.96)	(2.13)	(-1.30)	(-1.16)	(9.32)	(3.18)	(3.19)	(2.69)
N	163	163	163	159	162	162	162	158	169	169	169	165	169	169	169	165

Table 5: This table presents the results of regressing outcome variables relating to law enforcement on percentage of land under indigo cultivation in 1897. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are crimes per 10,000 in 1911 (Column 1-4), rural police per 10,000 in 1911 (Column 5-8), crimes per 10,000 in 2011 (Column 9-12) and constables per 10,000 (Column 13-16). *t* statistics are reported in round parenthesis and standard errors which are clustered at the district level are reported in square parenthesis. Columns 1, 5, 9 and 13 only control for state fixed effects. Columns 2, 6, 10 and 14 control for stated fixed effects and social controls. Columns 3, 7, 11 and 15 control for state fixed effects, social and geographic controls. Columns 4, 8, 12 and 16 control for stated fixed effects and social, geographic and development controls. * $p^{<}0.1$, ** $p^{<}0.05$, *** $p^{<}0.01$

	Agricultural Employment 1911				Agricultural Employment 2011			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Indigo 1897	-2.612***	-2.721***	-2.471	-1.681	2.760**	2.339**	0.495	0.449
	(-2.76)	(-2.68)	(-1.26)	(-1.48)	(2.07)	(2.03)	(0.49)	(0.44)
	[0.948]	[1.016]	[1.965]	[1.139]	[1.332]	[1.149]	[1.008]	[1.031]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic Controls	No	No	Yes	Yes	No	No	Yes	Yes
Development Controls	No	No	No	Yes	No	No	No	Yes
Constant	4.813	23.23	-46.37	28.19	25.53	61.79	-171.5	-203.0
	(1.58)	(2.68)	(-0.65)	(0.48)	(10.96)	(5.29)	(-2.69)	(-3.25)
N	169	169	169	165	169	169	169	165

Table 6: This table presents the results of regressing outcome variables relating to agricultural employment on percentage of land under indigo cultivation in 1897. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are percentage employed as agricultural workers in 1911 (column 1-4) and percentage employed as agricultural workers in 2011 (Column 5-8). *t* statistics are reported in round parenthesis and standard errors which are clustered at the district level are reported in square parenthesis. Columns 1 and 5 only control for state fixed effects. Columns 2 and 6 control for stated fixed effects and social controls. Columns 3 and 7 control for state fixed effects, social and geographic controls. Columns 4 and 8 control for stated fixed effects and social, geographic and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

unexpected given that the state was not involved in indigo cultivation. From column 1 in table 7, a percentage point increase in cultivated land under indigo cultivation is associated with a 0.014 rupee drop in civil work investment per individual in 1911. None of the variables reflecting access to public goods in 2011 have statistically or economically significant association with indigo cultivation. For example, the coefficient on access to pucca roads is -0.501 while the average is 77.542%.

6 Endogeneity Bias

Despite including an extensive list of controls, there is a concern of endogeneity because of the non-random selection of districts into indigo cultivation. As mentioned prior, availability of cheap labour and geographic suitability dictated where indigo was grown. Furthermore, there could be selection on unobserved variables which have not been controlled for. The historical background does not indicate whether this endogeneity could have led to a positive or negative selection. If areas which were geographically endowed selected into indigo cultivation, it could imply that there is positive selection as geographical endowment might be positively correlated with economic development. At the same time, if areas with cheap

	Civil Works Investment 1911	Medical Care 2011	Drinking Water 2011	Telephone 2011	Post Office 2011	Transport 2011	Banks 2011	Agricultural Credit 2011	Pucca Road	Power Supply 2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Indigo 1897	-0.014	0.267	0.507	1.738	1.346	0.026	0.077	0.619	-0.501	1.299
	(-0.44)	(0.14)	(0.86)	(0.87)	(0.92)	(0.02)	(0.09)	(0.37)	(-0.34)	(1.15)
	[0.032]	[1.879]	[0.588]	[1.989]	[1.468]	[1.714]	[0.831]	[1.676]	[1.471]	[1.132]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	7.227	56.68	82.30	367.6	292.3	355.9	27.76	346.8	227.8	256.5
	(2.15)	(0.53)	(2.12)	(4.77)	(2.75)	(3.82)	(0.47)	(3.89)	(2.26)	(3.74)
N	158	165	165	165	165	165	165	165	165	165

Table 7: This table presents the results of regressing outcome variables relating to public goods on percentage of land under indigo cultivation in 1897. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are civil works investment per individual (column 1), percentage with access to medical care (column 2), percentage with access to drinking water (column 3), percentage with access to telephones (column 4), percentage with access to post office (column 5), percentage with access to public transport in 2011 (column 6), percentage with access to banks in 2011 (column 7), percentage with access to agricultural credit in 2011 (column 8), percentage with access to pucca roads in 2011 (column 9) and percentage with access to power supply in 2011 (column 10). *t* statistics are reported in round parenthesis and standard errors which are clustered at the district level are reported in square parenthesis. All columns control for state fixed effects, social controls, geographic controls and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

labour selected into indigo cultivation, it could indicate that there was negative selection as these areas might be economically backward since labour is cheap.

Additionally, there could be measurement error in the independent variable that is indigo cultivation in 1897. A measurement error, if present, could induce correlation between the independent variable and the error term which will bias estimates towards zero (Pischke, 2007). A simple method to assess if there could be measurement error in the independent variable is row totals. For this, I calculate the total land under cultivation for each district under different crops in 1897 and compare it with the total provided by the Agricultural Statistics. While this method does not clearly indicate the presence of a measurement error, it does give an approximation of whether this error is present. Figure 4 plots the row totals against the reported totals. Most row totals are close to their reported total. However, for some districts, the reported totals exceed the row totals while for some districts the row totals are greater than the reported totals. Indeed, the preface to the Agricultural Statistics of India (1897-1902) mentions that in Bengal province, districts under North Bihar have estimates of land under cultivation which are not based on professional survey suggesting that the agricultural data for these districts might have measurement error.

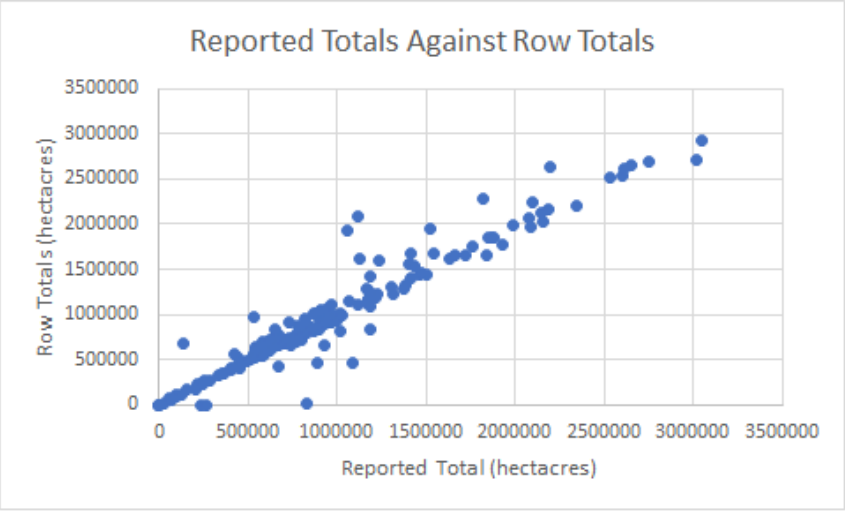

Figure 4: Scatter plot of the reported totals against the row totals. Data taken from Agricultural Statistics of India (1897-1902)

7 Instrumental Variable Strategy

Employing an instrumental variable strategy could deal with the endogeneity bias the OLS estimates are suffering from arising from selection (Angrist and Pischke, 2008) and measurement error (Pischke, 2007).

7.1 Constructing the Indigo Suitability Index

I instrument indigo cultivation with an indigo suitability index. The paper closest to this method is that by Roessler et al. (2020) where cash crop cultivation in Africa is instrumented with a agricultural suitability score. I have to construct the index since there is no prior index for the same. Given that the indigo suitability index will be used to instrument indigo cultivation in India, indigo data from India cannot be used for index construction. The Agricultural Statistics of India also provides data on indigo cultivation for present day Pakistan, Bangladesh and Burma which were part of colonial India in 1897. Assuming that indigo cultivation in these countries is uncorrelated with the subsequent partition of colonial India, data from the countries can be used to construct the index. There is data for 67 historic districts from these countries. The number of predictor variables in the model which largely cover latitude, longitude, inland water, elevation, precipitation, temperature and soil comes to 34. As the number of predictor variables is close to the number of observations, I employ a least absolute shrinkage and selection operator (LASSO) to select variables that explain a large part of the variation in indigo cultivation. I deal with the concern of overfitting by using 10-fold cross validation method which chooses the optimal lambda to penalize the coefficients on the variables (Fonti and Belister, 2017).

The LASSO model employing the optimal lambda selected 24 of 34 variables which are specified in Table 8. I then run a regression of percentage of land under indigo cultivation in 1897 for the 67 districts on the 24 selected variables. The coefficients (reported in Table 8) taken from the OLS regression are then used to construct the indigo suitability index for the Indian districts. The selection of variables relating to sandy soil, inland water and precipitation reflects the growing conditions for indigo mentioned in the literature (Kumar, 2012). Figure 5

maps the indigo suitability of districts historically part of British India. From the map, it is seen that districts historically falling under parts of Madras, Bengal, Agra and Oudh are the most suitable for indigo. Additionally, the summary statistics for the indigo suitability index for the 169 districts are mentioned in Table 1.

Selected Variable	***Coefficient from OLS Regression***
Latitude	-0.147
Longitude	0.086
Percentage of Sandy Soil	0.20
Percentage of Clay Soil	-0.027
Percentage of pH water in Soil	-0.284
Percentage of Organic Carbon in Soil	0.122
Percentage of Nitrogen in Soil	-0.395
Percentage of BS in Soil	-0.002
Cation Exchange Capacity of Soil	0.033
Cation Exchange Capacity of Clay Soil	0.022
Bulk Density of Soil	-0.749
Soil Curve Number	0.027
Major Inland Water Source	0.155
Mean Temperature of the Warmest Month	-0.375
Precipitation of the Wettest Month	0.001
Precipitation of the Driest Month	-0.011
Precipitation Seasonality	0.020
Precipitation of the Driest Quarter	-0.009
Isothermality	0.038
Temperature Seasonality	0.007
Minimum Temperature of the Coldest Month	0.100
Mean Temperature of the Wettest Quarter	0.224
Mean Temperature of the Driest Month	-0.037

Table 8: Variables Selected by LASSO Model

7.2 2 SLS Model

The indigo suitability index is then used to instrument percentage of cultivated land under indigo in 1897 in a 2SLS model. The first stage of the 2SLS is specified below:

$$indigo_{is} = \gamma suitability_{is} + \theta_s + X_{is} + \epsilon_{is} (2)$$

Here, $indigo_{is}$ is the percentage of cultivated land under indigo in district i in state s in 1897. $suitability_{is}$ is the suitability index. γ is the percentage point change in percentage of cultivated land under indigo from a one unit increase in the indigo suitability index. θ_s is state

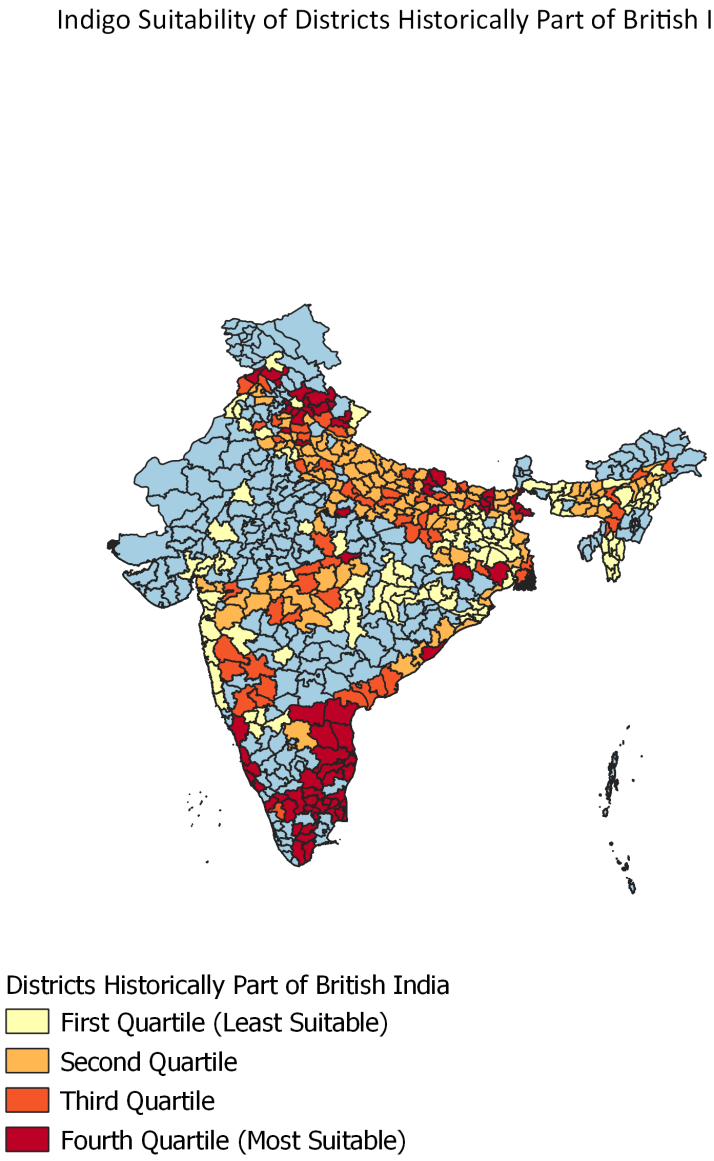

Figure 5: This map reflects the historic suitability of Indian districts for indigo. Districts marked in red were historically part of districts that were highly suitable for indigo cultivation and in the fourth quartile of indigo suitability. Districts marked in light yellow were historically part of districts that were least suitable for indigo and are in the first quartile of indigo suitability.

fixed effects. X_{is} is a vector of social, development and geographic controls. These controls are identical to the ones employed in the OLS. However, concerning the geographic controls, some of the controls used in the OLS model previously have been used in index construction. Therefore, if all the geographic variables used in the index construction are included as controls, the index will drop because of multicollinearity. Nonetheless, among the geographic variables included in the index are latitude and longitude which if not controlled for could inflate the t-statistic because of spatial correlation (Kelly, 2020). To abate this concern, I control for latitude, longitude, inland water, soil composition and precipitation along with geographic variables that were not used in index construction.

For the instrument to be valid, it needs to meet the relevance, independence and exclusion restriction conditions (Angrist and Pischke, 2008). The relevance condition specifies that the suitability index should have a causal effect on indigo cultivation. This condition is met through the first stage results reported in Table 10, 11, 12 and 13 where the correlation between the index and indigo cultivation is shown to be significant and positive. The independence condition requires that the instrument be as good as randomly assigned and can be assumed to be met given that planters could not assign indigo suitability to districts.

The third condition is the exclusion restriction which states that the suitability index can only influence the outcome variables through indigo cultivation in 1897. While indigo is no longer cultivated on a commercial scale anywhere in the world, the indigo suitability index can influence the outcome variables through other channels. The growing conditions for indigo might be similar to other crops which are still cultivated in India today suggesting that the indigo suitability index could be influencing the outcomes through the crop which has similar growing conditions to indigo. Consequently, I run OLS regression of percentage of cultivated land under other crops (rice, wheat, tobacco, jute, cotton) in 1897 on the indigo suitability index. Data on these crops is taken from the Agricultural Statistics of India. The results of the regression are reported in column 1-5 of Table 9 and indicate that the indigo suitability index does not have a significant correlation with any of the crops considered. Additionally, the indigo suitability index could have a correlation with economic activity as indigo suitability might attract economic activity irrespective of whether indigo is grown in the district. As a falsification check, I run a regression of the number of factories per

100 individuals on the indigo suitability index. Data on factories is taken from the colonial censuses of 1911 and serves as a proxy for economic activity. Column 6 of Table 9 report the results of this regression and there does not appear to be a significant correlation between indigo suitability and factories. Therefore, these falsification checks increase confidence that the exclusion restriction holds.

	Rice 1897	Wheat 1897	Jute 1897	Cotton 1897	Tobacco 1897	Factories 1911
	(1)	(2)	(3)	(4)	(5)	(6)
Indigo Suitability Index	-10.23	-0.558	0.308	-2.275	0.0438	-0.000318
	(-1.16)	(-0.30)	(0.97)	(-1.34)	(0.51)	(-0.45)
	[8.848]	[1.888]	[0.317]	[1.7]	[0.086]	[0.000]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	165	165	165	165	165	165

Table 9: This table presents the results of the falsifications checks. The outcome variables are percentage of land under rice in 1897 (column 1), percentage of cultivated land under wheat in 1897 (column 2), percentage of cultivated land under jute in 1897 (column 3), percentage of cultivated land under cotton in 1897 (column 4), percentage of cultivated land under tobacco in 1897 (column 5) and number of factories per 100 in population in 1911 (column 6). *t* statistics are reported in round parenthesis and standard errors which are clustered at the district level are reported in square parenthesis. All columns control for state fixed effects, social controls, geographic controls and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

7.3 Results for the 2SLS

Table 10,11,12 and 13 report the results for the 2SLS regressions. For a majority of the coefficients (except for crime in 1911 and telephone access in 2011), it is seen that their size has increased. This indicates that the OLS estimates were biased towards zero which suggests either positive selection or measurement error in the indigo cultivation variable. The F statistic ranges from 22.78 to 24.87 for the first stage regression of indigo cultivation in 1897 on the indigo suitability index. Due to the instrument just crossing the recommended threshold of 10, I report the conditional likelihood ratio confidence intervals for the endogenous regressor (indigo cultivation in 1897) (Moreira, 2003) (Mikusheva and Poi, 2006). The coefficient of the first stage ranges from 0.984 to 1.035 which implies that a unit increase in the indigo suitability index is associated with a 0.984 to 1.035 percentage point increase in percentage of cultivated land under indigo in 1897.

Table 10 reports the results for outcome relating to human capital. Column 1 and 2 state that a percentage point increase in cultivated land under indigo causes a 0.322 percentage point increase in literacy and 0.113 rupees more in education investment per individual in 1911. Similarly, column 3 and 4 indicate that a percentage point increase in cultivated land under indigo leads to 1.582 less schools per 10,000 in the population and a 1.029 percentage point fall in literacy in 2011. Out of these results, only the coefficient on schools per 10,000 in population in 2011 is significant. This result indicates while indigo cultivation did not influence the pattern of colonial investment in education, it did manage to influence the number of educational institutions presently.

	(1)	(2)	(3)	(4)
		First Stage		
		Indigo 1897		
Indigo Suitability Index	0.984***	1.024***	0.984***	0.984***
	(4.9)	(4.99)	(4.9)	(4.9)
	[0.2]	[0.205]	[0.2]	[0.2]
F Stat	24.06	24.87	24.06	24.06
		Second Stage		
	Literacy 1911	Education Investment 1911	Schools 2011	Literacy 2011
Indigo 1897	0.322	0.113	-1.582**	-1.209
	(0.52)	(1.62)	(-2.00)	(-0.76)
	[0.616]	[0.07]	[0.792]	[1.588]
Confidence Interval	[-1.291, 2.174]	[-0.074, 0.366]	[0.043, 0.324]	[-5.615, 2.57]
State Fixed Effects	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes
N	165	160	165	165

Table 10: This table reports the results of the 2SLS regression of outcome variables relating to human capital on percentage of cultivated land under indigo in 1897 where percentage of cultivated land under indigo is instrumented with an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. Both the results of the first and second stage are reported. The outcome variables are literacy in 1911 (column 1), education investment per individual in 1911 (column 2), number of schools per 10,000 in 2011 (column 3) and literacy in 2011 (column 4). For the first stage, the F statistic is reported as well. The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. Conditional likelihood ratio confidence intervals for the endogenous regressor (indigo cultivation in 1897) are reported based on Moreira (2003) and Mikusheva and Poi (2006). All regressions reported control for stated fixed effects, social controls, geographic controls and development controls. * p0.1, ** p<0.05, *** p<0.01

Table 11 reports the results for law enforcement outcomes and it is seen that the results remain largely unchanged in significance from the OLS results except for constables in 2011. From column 4, it is seen that a percentage point increase in cultivated land under indigo

causes a 1.96 decline in constables per 10,000 in population which is economically significant as it indicates a 14.7% fall relative to the mean. The insignificant results for rural police and crime in 1911 imply that the efforts of the planters to strengthen law enforcement did not produce persistent effects. However, the significant, negative result for constables is surprising given that indigo cultivation increased police and not the other way around. It is is possible that the declining role of indigo growing districts as economic centers because of their focus on labour intensive agriculture, required the shifting of physical capital to districts focused on industrial activity and advanced agriculture. In the discussion section, I report correlations between constables and various measures of economic activity and agricultural development in 2011 to further understand the mechanisms behind this result.

	(1)	(2)	(3)	(4)
		First Stage		
		Indigo 1897		
Indigo Suitability Index	1.035***	1.025***	0.984***	0.984***
	(4.98)	(4.92)	(4.9)	(4.9)
	[0.207]	[0.208]	[0.2]	[0.2]
F Stat	24.84	24.24	24.06	24.06
		Second Stage		
	Crime 1911	Rural Police 1911	Crime 2011	Constables 2011
Indigo 1897	-4.778	-2.275	-2.149	-1.96**
	(-0.32)	(-0.69)	(-1.23)	(-2.27)
	[15.008]	[3.294]	[1.749]	[0.865]
Confidence Interval	[-36.301, 28.353]	[-11.705, 6.765]	[-7.092, 1.82]	[-4.717, -0.014]
State Fixed Effects	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes
N	159	158	165	165

Table 11: This table reports the results of the 2SLS regression of outcome variables relating to law enforcement on percentage of cultivated land under indigo in 1897 where percentage of cultivated land under indigo is instrumented with an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. Both the results of the first and second stage are reported. The outcome variables are number of crimes per 10,000 in 1911 (column 1), rural police per 10,000 in 1911 (column 2), number of crimes per 10,000 in 2011 (column 3) and number of constables per 10,000 in 2011 (column 4). For the first stage, the F statistic is reported as well. The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. Conditional likelihood ratio confidence intervals for the endogenous regressor (indigo cultivation in 1897) are reported based on Moreira (2003) and Mikusheva and Poi (2006). All regressions reported control for stated fixed effects, social controls, geographic controls and development controls. * p<0.1, ** p<0.05, *** p<0.01

The results for agricultural employment are reported in Table 12. Column 1 shows that a percentage point increase in cultivated land under indigo causes a 4.265 percentage point in-

crease in percentage employed as agriculture workers in 1911. Similarly, column 2 states that a one percentage point increase in cultivated land under indigo leads to a 5.87 percentage point increase in percentage employed as agricultural workers in 2011. Both these results are statistically significant and potentially imply that indigo cultivation by being highly labour intensive and not involving innovation limited opportunities to take on more complex economic activities such as manufacturing resulting in a persistent effect on agricultural employment.

	(1)	(2)
	First Stage	
	Indigo 1897	
Indigo Suitability Index	0.998***	0.998***
	(4.77)	(4.77)
	[0.209]	[0.209]
F Stat	22.78	22.78
	Second Stage	
	Agricultural Employment 1911	Agricultural Employment 2011
Indigo 1897	4.265**	5.87**
	(1.81)	(2.62)
	[4.047]	[3.975]
Confidence Interval	[0.793, 11.216]	[1.14, 12.245]
State Fixed Effects	Yes	Yes
Social Controls	Yes	Yes
Geographic Controls	Yes	Yes
Development Controls	Yes	Yes
N	165	165

Table 12: This table reports the results of the 2SLS regression of outcome variables relating to agricultural employment on percentage of cultivated land under indigo in 1897 where percentage of cultivated land under indigo is instrumented with an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. Both the results of the first and second stage are reported. The outcome variables are percentage employed as agricultural workers in 1911 (column 1) and percentage employed as agricultural workers in 2011 (column 2). For the first stage, the F statistic is reported as well. The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. Conditional likelihood ratio confidence intervals for the endogenous regressor (indigo cultivation in 1897) are reported based on Moreira (2003) and Mikusheva and Poi (2006). All regressions reported control for stated fixed effects, social controls, geographic controls and development controls. * p<0.1, ** p<0.05, *** p<0.01

Finally, Table 13 reports the results when taking outcomes related to public goods. Column 1 shows that a percentage point increase in percentage of cultivated land under indigo causes 0.244 rupees more to be spent per individual on civil works. This result is not statistically significant which increases confidence in the claim that the minimal involvement of the state prevented diversion of investment away from public goods. Column 2 to 10 present the

results for various public goods in 2011. Here, only the result for percentage with access to banks is significant. From 7, it is seen that a percentage point increase in cultivated land under indigo causes a 7.887 percentage point increase in percentage with access to banks. This result might appear unexpected because of the minimal role of the state, however, planters might have been incentivized to improve banking infrastructure to aid the indigo industry.

7.4 Robustness Checks

It was earlier reported that there was a choice of using province fixed effects. I argued that employing state fixed effects in place of province would not matter as long as indigo cultivation is uncorrelated with the reorganization of states in 1956. To increase confidence in this claim, I replace state fixed effects with province fixed effects. Table 14 and 15 report the results with province fixed effects. The coefficients in size and significance are similar to those reported with state fixed effects. Therefore, the results are robust to replacing state fixed effects with province fixed effects.

Additionally, Kelly (2019) finds that historical persistence studies tend to have inflated t-statistics which are caused by standard errors which have not been corrected for spatial correlation. Here, there might be potential spatial correlation between a district and its neighbours as similar factors might have influenced whether a district and its neighbours were considered suitable for indigo. I calculate the distance for each historical district to its first and second closest neighbour. The maximum distance between a district and its first and second closest neighbour is 213.74 and 307.89 kilometres respectively. As a robustness check, I report the Conley standard errors which allow for spatial correlation within the specified bandwith. Results with the bandwith set to 213.74 and 307.89 kilometres are reported using the acreg command (Collela et al., 2019) which provides the second stage results and the Conley standard errors. The results for the same are reported in Table 16 and 17.

Compared to the main results which were reported in table 10, 11, 12 and 13, the t-statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	First Stage									
	Indigo 1897									
Indigo Suitability Index	1.026***	0.984***	0.984***	0.984***	0.984***	0.984***	0.984***	0.984***	0.984***	0.984***
	(4.97)	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)
	[0.206]	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]
F Stat	24.67	24.06	24.06	24.06	24.06	24.06	24.06	24.06	24.06	24.06
	Second Stage									
	Civil Works 1911	Medical Access 2011	Drinking Water 2011	Telephone 2011	Postoffice 2011	Transport 2011	Banks 2011	Power Supply 2011	Agricultural Credit 2011	Pucca Road 2011
Indigo 1897	0.244	6.294	1.152	-0.615	4.129	1.464	7.887***	-3.691	6.661	4.84
	(1.5)	(1.42)	(0.77)	(-0.21)	(1.27)	(0.4)	(2.62)	(-1.55)	(1.63)	(1,29)
	[0.162]	[4.426]	[1.491]	[2.95]	[3.252]	[3.688]	[3.015]	[2.38]	[4.08]	[3.765]
Confidence Interval	[0.015, 0.572]	[-3.432, 18.795]	[-3.558, 6.278]	[-9.412, 7.582]	[-3.000, 12.503]	[-7.899, 11.76]	[2.642, 16.21]	[-1.727, 16.876]	[-3.706, 15.245]	[-13.674, 3.742]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	158	165	165	165	165	165	165	165	165	165

Table 13: This table reports the results of the 2SLS regression of outcome variables relating to public goods on percentage of cultivated land under indigo in 1897 where percentage of cultivated land under indigo is instrumented with an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. Both the results of the first and second stage are reported. The outcome variables are civil works investment per individual in 1911 (column 1), percentage with access to medical care in 2011 (column 2), percentage with access to drinking water in 2011 (column 3), percentage with access to telephones in 2011 (column 4), percentage with access to post office in 2011 (column 5), percentage with access to public transport in 2011 (column 6), percentage with access to banks in 2011 (column 7), percentage with access to power supply in 2011 (column 8), percentage with access to agricultural credit in 2011 (column 9) and percentage with access to pucca roads in 2011 (column 10). For the first stage, the F statistic is reported as well. The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. Conditional likelihood ratio confidence intervals for the endogenous regressor (indigo cultivation in 1897) are reported based on Moreira (2003) and Mikusheva and Poi (2006). All regressions reported control for stated fixed effects, social controls, geographic controls and development controls. * p<0.1, ** p<0.05, *** p<0.01

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
						First Stage				
						Indigo 1897				
Indigo Suitability Index	1.004***	1.049***	1.004***	1.004***	1.068***	1.052***	1.004***	1.004***	1.019***	1.019***
	(4.85)	(4.95)	(4.85)	(4.85)	(4.96)	(4.9)	(4.85)	(4.85)	(4.72)	(4.72)
	[0.207]	[0.212]	[0.207]	[0.207]	[0.215]	[0.215]	[0.207]	[0.207]	[0.216]	[0.216]
F Stat	23.52	24.52	23.52	23.52	24.56	23.98	23.52	23.52	22.25	22.25
						Second Stage				
	Literacy 1911	Education Investment 1911	Schools 2011	Literacy 2011	Crime 1911	Rural Police 1911	Crimes 2011	Constables 2011	Agricultural Employment 1911	Agricultural Employment 2011
Indigo 1897	0.312	0.083	-1.423*	-0.921	2.431	-1.591	-1.156	-2.195**	4.112*	5.502**
	(0.53)	(1.25)	(-1.93)	(-0.61)	(0.21)	(-0.51)	(-0.69)	(-2.3)	(1.77)	(2.47)
	[0.591]	[0.066]	[0.737]	[1.505]	[11.612]	[3.135]	[1.668]	[0.955]	[2.33]	[2.225]
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	165	160	165	165	159	158	165	165	165	165

Table 14: This table reports the results of the 2SLS regression of outcome variables on percentage of land under indigo cultivation which is instrumented by the indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are literacy in 1911 (column 1), education investment per individual in 1911 (column 2), schools per 10,000 in population (column 3), literacy in 2011 (column 4), crimes per 10,000 in 1911 (column 5), rural police per 10,000 in 1911 (column 6), crimes per 10,000 in 2011 (column 7), constables per 10,000 in 2011 (column 8), percentage employed as agricultural workers in 1911 (column 9) and percentage employed as agricultural workers in 2011 (column 10). The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. All regressions reported control for province fixed effects, social controls, geographic controls and development controls. * p<0.1, ** p<0.05, *** p<0.01

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					First Stage					
					Indigo 1897					
Indigo Suitability Index	1.048***	1.004***	1.004***	1.004***	1.004***	1.004***	1.004***	1.004***	1.004***	1.004***
	(4.93)	(4.85)	(4.85)	(4.85)	(4.85)	(4.85)	(4.85)	(4.85)	(4.85)	(4.85)
	[0.213]	[0.207]	[0.207]	[0.207]	[0.207]	[0.207]	[0.207]	[0.207]	[0.207]	[0.207]
F Stat	24.31	23.52	23.52	23.52	23.52	23.52	23.52	23.52	23.52	23.52
					Second Stage					
	Civil Works 1911	Medical Access 2011	Drinking Water 2011	Telephone 2011	Post Office 2011	Transport 2011	Banks 2011	Agricultural Credit 2011	Pucca Road 2011	Power Supply 2011
Indigo 1897	0.227	5.056	1.114	-0.552	3.798	0.972	7.748**	5.986	3.396	-3.849
	(1.39)	(1.2)	(0.77)	(-0.19)	(1.20)	(0.27)	(2.58)	(1.5)	(1.09)	(-1.58)
	[0.164]	[4.222]	[1.445]	[2.935]	[3.175]	[3.592]	[3.002]	[3.99]	[3.112]	[2.429]
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	158	165	165	165	165	165	165	165	165	165

Table 15: This table reports the results of the 2SLS regression of outcome variables on percentage of land under indigo cultivation which is instrumented by an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are civil works investments per individual in 1911 (column 1), percentage with access to medical care in 2011 (column 2), percentage with access to drinking water in 2011 (column 3), percentage with access to telephone in 2011 (column 4), percentage with access to post office in 2011 (column 5), percentage with access to public transport in 2011 (column 6), percentage with access to banks in 2011 (column 7), percentage with access to agricultural credit in 2011 (column 8), percentage with access to pucca road in 2011 (column 9) and percentage with access to power supply (column 10). The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. All regressions reported control for province fixed effects, social controls, geographic controls and development controls. * p<0.1, ** p<0.05, *** p<0.01

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Second Stage									
	Literacy 1911	Education Investment 1911	Schools 2011	Literacy 2011	Crime 1911	Rural Police 1911	Crimes 2011	Constables 2011	Agricultural Employment 1911	Agricultural Employment 2011
Panel A:Distance Threshold (213.74 km)										
Indigo 1897	0.322	0.113	-1.582*	-1.209	-4.778	-2.275	-2.149	-1.196**	4.112*	5.502*
	(0.53)	(1.15)	(-1.72)	(-0.68)	(-0.36)	(-0.62)	(-1.44)	(-2.00)	(1.76)	(1.81)
	[0.61]	[0.098]	[0.917]	[1.77]	[13.172]	[3.683]	[1.495]	[0.984]	[2.339]	[3.034]
Panel B: Distance Threshold (307.89 km)										
Indigo 1897	0.322	0.113	-1.582*	-1.209	-4.778	-2.275	-2.149	-1.196**	4.112**	5.502*
	(0.48)	(1.08)	(-1.72)	(-0.79)	(-0.58)	(-0.68)	(-1.25)	(-0.85)	(2.05)	(1.88)
	[0.672]	[0.105]	[0.917]	[1.536]	[8.2]	[1.716]	[1.716]	[0.992]	[2.004]	[2.918]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	165	160	165	165	159	158	165	165	165	165

Table 16: This table reports the results of the 2SLS regression of outcome variables on percentage of land under indigo cultivation which is instrumented by the indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are literacy in 1911 (column 1), educational investment per individual in 1911 (column 2), number of schools per 10,000 in population in 2011 (column 3), literacy in 2011 (column 4), crimes per 10,000 in 1911 (column 5), rural police per 10,000 in 1911 (column 6), crimes per 10,000 in 2011 (column 7), constables per 10,000 in 2011 (column 8), percentage employed as agricultural workers in 1911 (column 9) and percentage employed as agricultural workers in 2011 (column 10). The t statistics are reported in round parenthesis and the Conley standard errors are reported in square brackets. Conley standard errors are computed using the acreg command recommended by Colella et al. (2019). Panel A takes 213.74 kilometres as the threshold and Panel B takes 307.89 kilometres as the threshold. All regressions reported control for state fixed effects, social controls, geographic controls and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Second Stage									
	Civil Works 1911	Medical Access 2011	Drinking Water 2011	Telephone 2011	Post Office 2011	Transport 2011	Banks 2011	Agricultural Credit 2011	Pucca Road 2011	Power Supply 2011
Panel A:Distance Threshold (213.74 km)										
Indigo 1897	0.244	6.294	1.152	-0.615	4.129	1.464	7.887**	6.661	4.84	-3.691
	(1.47)	(1.13)	(0.76)	(-0.23)	(1.16)	(0.37)	(2.11)	(1.32)	(1.27)	(-1.26)
	[0.166]	[5.552]	[1.507]	[2.711]	[4.129]	[3.99]	[3.774]	[5.041]	[3.806]	[2.938]
Panel B: Distance Threshold (307.89)										
Indigo 1897	0.244	6.924	1.152	-0.615	4.129	1.464	7.887*	6.661	4.84	-3.691
	(1.52)	(1.23)	(0.77)	(-0.23)	(1.16)	(0.38)	(1.91)	(1.22)	(1.20)	(-1.19)
	[0.16]	[5.117]	[1.503]	[2.653]	[3.562]	[3.885]	[4.127]	[5.44]	[4.031]	[3.092]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	158	165	165	165	165	165	165	165	165	165

Table 17: This table reports the results of the 2SLS regression of outcome variables on percentage of land under indigo cultivation which is instrumented by the indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. The outcome variables are literacy in 1911(column 1), education investment per individual (column 2), schools per 10,000 in population (column 3), literacy in 2011 (column 4), crimes per individual in 1911 (column 5), rural police per individual (column 6), crimes per individual in 2011 (column 7), constables per individual in 2011 (column 8), percentage employed as agricultural workers in 1911 (column 9) and percentage employed as agricultural workers in 2011 (column 10). The *t* statistics are reported in round parenthesis and the Conley standard errors are reported in square brackets. Conley standard errors are computed using the acreg command recommended by Colella et al. (2019). Panel A takes 213.74 kilometres as the threshold and Panel B takes 307.89 kilometres as the threshold. All regressions reported control for state fixed effects, social controls, geographic controls and development controls. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

in table 16 and 17 are relatively smaller. This suggests the existence of spatial correlation between a district and its neighbouring districts. However, the coefficients in size and significance predominantly remain similar to those reported in the main results. This indicates that the results are robust to accounting for spatial correlation between districts and their nearest neighbours.

From the main results and the various robustness checks performed, it is seen that the results for agricultural employment in 1911 and 2011, schools per 10,000 in population in 2011, access to banks in 2011 and constables per 10,000 in population are robust to various specifications.

8 Discussion of Results

Out of the three channels mentioned in the introduction through which indigo cultivation could have persistent effects, the channel at play based on the results appears to be the failure of the indigo industry to create production linkages with local industries and lack of innovation. Vogel (1994) commented that agriculture in rural areas needs to foster production linkages to the rest of economy for rural industrialization to be possible. Similarly, Hirschman (1958) pointed out that the failure of agriculture to develop production linkages prevents capital formation which is needed for economic development. The failure to develop production linkages is reflected in the positive relationship between indigo cultivation and employment as agricultural workers. Agricultural workers typically are involved in labour intensive agriculture reflecting the failure of indigo cultivation to move towards complex agricultural practices. Additionally, the negative persistent effect of indigo cultivation on schools per 10,000 indicates that the focus of former indigo district on labour intensive agriculture has reduced the importance of educational institutions needed to develop human capital. At the same time, the lack of persistent effect on literacy in 1911 and 2011 suggests that the negative persistent effect is mainly driven by higher education institutions. This intuition is confirmed in table 18 which considers the impact of indigo cultivation on different levels of educational institutions in 2011. A negative, significant effect is observed for secondary, senior secondary and colleges. Data on educational institutions is taken from the

District Census Handbooks.

	(1)	(2)	(3)	(4)	(5)
			First Stage		
			Indigo 1897		
Indigo Suitability Index	0.984***	0.984***	0.984***	0.984***	0.984***
	(4.9)	(4.9)	(4.9)	(4.9)	(4.9)
	[0.2]	[0.2]	[0.2]	[0.2]	[0.2]
F Stat	24.06	24.06	24.06	24.06	24.06
			Second Stage		
	Primary	Middle	Secondary	Senior Secondary	College
Indigo 1897	-0.557	-0.207	-0.279*	-0.387***	-0.151**
	(-1.47)	(-0.74)	(-1.70)	(-2.85)	(-2.15)
	[0.379]	[0.279]	[0.164]	[0.135]	[0.07]
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Social Controls	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Development Controls	Yes	Yes	Yes	Yes	Yes
N	165	165	165	165	165

Table 18: This table reports the results of the 2SLS regression of number of educational institutions per 10,000 on percentage of cultivated land under indigo in 1897 where percentage of cultivated land under indigo is instrumented with an indigo suitability index. Percentage of land under indigo cultivation takes a value between 0 and 100. Both the results of the first and second stage are reported. The outcome variables are number of primary schools per 10,000 (column 1), number of middle schools per 10,000 (column 2), number of secondary schools per 10,000 (column 3), number of senior secondary schools per 10,000 (column 4) and number of colleges per 10,000 (column 5). For the first stage, the F statistic is reported as well. The *t* statistics are reported in round parenthesis and the standard errors which are clustered at the district level are reported in square brackets. All regressions reported control for stated fixed effects, social controls, geographic controls and development controls. * p0.1, ** p<0.05, *** p<0.01

Concurrently, the limited role of the state prevented the creation of enclave economies as seen in the case of cash crop cultivation in West Africa (Roessler, 2020). In West Africa, areas involved in cash crop cultivation developed into enclaves and were preferred for colonial infrastructural investments leading to these areas being more economically developed that neighbouring areas. Alternatively, the lacking presence of the state distinguishes indigo cultivation from opium cultivation where the state actively diverted resources from human capital to physical capital (Lehne, 2019). The minimal role of the state in indigo cultivation is evident in the lack of relationship between indigo cultivation and access to public goods like drinking water. Nonetheless, the significant, positive result for access to banks indicates that planters need for capital to finance production created persistent effects on financial infrastructure.

Furthermore, the planter's efforts to strengthen law enforcement to ensure following of contracts does not appear to have persistent effects in 1911. Lehne (2019) reports similar results in the case of opium cultivation in India, although, the results for rural police are significant in 1911 and lose significance in 2011 suggesting that the role of the state in opium cultivation allowed for the effect on police to persist slightly. However, unlike, Lehne's (2019) results which lose significance for police in 2011, the coefficient on constables per 10,000 in population is negative and significant. Literature which looks at the role of police and security in developing countries (Marenin, 2014) indicates that a functional police force is required for economic development. If a positive relationship is expected between development and police force then it is possible that former indigo growing districts are not prioritized for security as a result of being considered economically backward because of their focus on labour intensive agriculture. This hypothesis is supported by the insignificant result for crimes in 2011 which suggests that the negative, significant effect on constables is not caused by criminal activity. As a preliminary analysis to test this hypothesis, I run Pearson's correlations between constables per 10,000 and various measures of economic activity in 2011 (secondary sector GDP, tertiary sector GDP, market density, road density). I obtain the district level measures of economic activity from ICRISAT [9]. However, the ICRISAT database only covers 147 of the historical districts in the study and therefore, these correlations should be taken as approximations of the association between police and economic development.

Table 19 reports the Pearson correlation coefficients. Constables per 10,000 is significantly

	Secondary Sector GDP	Tertiary Sector GDP	Market Density	Road Density
Constables 2011	0.183**	0.144*	0.244***	0.004***
N	147	147	147	147

Table 19: The table reports the Pearson's correlations between constables per 10,000 in 2011 and various measures of economic activity. The measures of agricultural development are all in 2011 and were extracted from the ICRISAT database. The measures of economic activity are secondary sector GDP, tertiary sector GDP, market density and road density.

but weakly correlated with the various measures of economic activity which suggests that there might not be an economically significant relationship between constables and economic activity.

[9] International Crops Research Institute for Semi-Arid Tropics

Another explanation behind the results for constables could be that police forces are sent to districts with developed agriculture. Since former indigo growing districts are involved in labour intensive agriculture, they might not be prioritized as they might not contribute significantly to agricultural output. Security might be needed in districts which have large landholdings dedicated to major crops such as rice, sugar, wheat and cotton. To test whether such a relationship might exist, I report Pearson correlations between constables per 10,000 in 2011 and measures of agricultural development in 2011 (percentage employed as agricultural workers, percentage with access to agricultural credit, primary sector GDP, irrigated land under rice, sugarcane, cotton and wheat). Data on irrigated land under various crops and primary sector GDP is taken from ICRISAT and therefore, again limited to 147 districts.

	Agricultural Workers	Agricultural Credit	Primary Sector GDP	Rice Irrigation	Sugarcane Irrigation	Cotton Irrigation	Wheat Irrigation
Constables 2011	-0.319**	0.262**	0.080	0.2190**	-0.165*	0.128	-0.076
N	169	169	147	147	147	147	147

Table 20: The table reports the Pearson's correlations between constables per 10,000 in 2011 and various measures of agricultural development. The measures of agricultural development are all in 2011 and were extracted from the ICRISAT database and 2011 district census handbooks. The measures of agricultural development are percentage employed as agricultural workers, access to agricultural credit, primary sector GDP and irrigated land under rice, sugarcane, cotton and wheat.

Table 20 reports the Pearson correlation coefficients between agricultural development and constables per 10,000 in population in 2011. Constables in 2011 is moderately and negatively correlated with percentage employed as agricultural workers. At the same time, it is positively and moderately correlated with percentage with access to agricultural credit. Agricultural credit can be taken as a sign of agricultural development since it indicates that farmers are not relying on informal sources such as moneylenders which small landholders tend to employ (Wahi, 2019). Therefore, these correlations suggest that constables is positively associated with agricultural development. As a result, the significant, negative coefficient on constables in 2011 could imply that indigo districts are not preferred for security since these districts are agriculturally underdeveloped. However, these correlations should be taken as a preliminary analysis and additional research is required to understand the mechanisms through which indigo cultivation impacts constables in 2011.

9 Conclusion

This study attempts to estimate the long term impact of indigo cultivation on human capital, agricultural employment, public goods and law enforcement. The results suggest that indigo cultivation led to an increased reliance on labour intensive agriculture and subsequently, a decline in institutions needed to develop human capital. However, the minimal role of the state prevented a diversion of resources away from public goods. Furthermore, the limited role of the state did not allow changes to law enforcement made by the planters to persist in 1911. However, there is a negative effect of indigo cultivation on constables in 2011 suggesting that former indigo cultivating districts are not preferred for security because of being agriculturally backward.

This study primarily contributes to the literature on the relationship between colonial cash crop cultivation and economic development. Indigo differs from earlier studied crops because of the lack of state intervention and the use of direct peasant to planter contracts which were considered risky given India's land tenure system (Chaudhuri, 2008). Additionally, this study contributes a novel indigo suitability index. Future work on indigo cultivation can look at the relationship between colonial indigo cultivation and democratic participation and conflict because of the peasant-planter conflicts caused by the contracts (Bhattacharya, 1977).

The central limitation of this study is that the archival datasets could be suffering from measurement error. As mentioned prior, with regards to Agricultural Statistics of India (1897-1902), certain districts might have agricultural data not collected from professional survey creating measurement error. A solution to this could be using the colonial land settlement reports for British India which report agricultural data at the subdistrict level and are more likely to be accurate. However, due to limited access to archives, these reports could not be accessed. A second concern is the predictability power of the indigo suitability index. While the index is able to predict indigo cultivation in British India, it was constructed using data for Pakistan, Bangladesh and Myanmar which are geographically similar to certain regions of India. It is unclear if this index will be able to predict indigo cultivation is areas that are geographically unlike British India. While indigo is no longer cultivated at a commercial scale in the world presently, the index could be tested using archival data from either USA or South

America where indigo was historically cultivated as well. However, these data limitations can be dealt with through access to archival data sources.

Word Count: 7999 (excluding references, footnotes, tables, figures, appendices, equations)

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91.5 (2001): 1369-1401.

Asiaticus. "The Rise and Fall of the Indigo Industry in India." *The Economic Journal* (1912): 237-247.

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Austin, Gareth. "Cash Crops and Freedom: Export Agriculture and the Decline of Slavery in Colonial West Africa." *International Review of Social History* 54.1 (2009): 1-37.

Banerjee, Abhijit, and Lakshmi Iyer. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95.4 (2005): 1190-1213.

Bhattacharya, Subhas. "The Indigo Revolt of Bengal." *Social Scientist* (1977): 13-23.

Calvi, Rossella, and Federico G. Mantovanelli. "Long-Term Effects of Access to Health Care: Medical Missions in Colonial India." *Journal of Development Economics* 135 (2018): 285-303.

Chattopadhyay, Kunal, and Muntassir Mamoon. "Indigo Rebellion." *The International Encyclopedia of Revolution and Protest* (2009): 1-2.

Chaudhary, Latika, and Manuj Garg. ”Does History Matter? Colonial Education Investments in India.” *The Economic History Review* 68.3 (2015): 937-961.

Chaudhuri, Binay Bhushan. *Peasant History of Late Pre-colonial and Colonial India*. Vol. 8. Pearson Education India, 2008.

Colella, Fabrizio, Lalive, R., Sakalli, S. O. and Thoening, M.. ”Inference with Arbitrary Clustering.” *SSRN Electronic Journal* (2019).

Dell, Melissa. ”The Persistent Effects of Peru’s Mining Mita.” *Econometrica* 78.6 (2010): 1863-1903.

Dell, Melissa, and Benjamin A. Olken. ”The Development Effects of the Extractive Colonial Economy: the Dutch Cultivation System in Java.” *The Review of Economic Studies* 87.1 (2020): 164-203.

Donaldson, Dave. ”Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.” *American Economic Review* 108.4-5 (2018): 899-934.

Easterly, William, and Ross Levine. ”The European Origins of Economic Development.” *Journal of Economic Growth* 21.3 (2016): 225-257.

Fenske, James. ”Trees, Tenure and Conflict: Rubber in Colonial Benin.” *Journal of Development Economics* 110 (2014): 226-238.

Fick, Stephen E., and Robert J. Hijmans. ”WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 37.12 (2017): 4302-4315.

Fonti, Valeria, and Eduard Belitser. ”Feature Selection Using Lasso.” VU Amsterdam *Research Paper in Business Analytics* 30 (2017): 1-2

Gupta, Bishnupriya, and Anand V. Swamy. ”Reputational Consequences of Labor Coercion:Evidence from Assam’s tea Plantations.” *Journal of Development Economics* 127 (2017): 431-439.

Hirschman, Albert O. *The Strategy of Economic Development.*. Yale University Press, New Haven, 1958.

India. Office of the Registrar General, et al. *Administrative Atlas of India*. Office of the Registrar General Census Commissioner, India, Ministry of Home Affairs, Government of India, 2011.

Iyer, Lakshmi. "Direct Versus Indirect Colonial Rule in India: Long-Term Consequences." *The Review of Economics and Statistics* 92.4 (2010): 693-713.

Kelly, Morgan. "The Standard Errors of Persistence". *SSRN Electronic Journal* (2019).

Kumar, Prakash. *Indigo Plantations and Science in Colonial India*. Cambridge University Press, 2012.

Kumar, Hemanshu and Rohini Somanthan. Creating Long Panels Using Census Data (1961-2001) *Economic and Political Weekly*. 52 (2017).

Lehne, Jonathan. "The Long-Run Effects of Opium Cultivation in British India." 2019. Paris School of Economics. PhD Dissertation.

Marenin, Otwin. "Styles of Policing and Economic Development in African States." *Public Administration and Development* 34.3 (2014): 149-161.

Moreira, Marcelo J. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71.4 (2003): 1027-1048.

Nadri, Ghulam A. *The Political Economy of Indigo in India, 1580-1930: A Global Perspective*. Brill, 2016.

Naritomi, Joana, Rodrigo R. Soares, and Juliano J. Assunção. "Institutional Development and Colonial Heritage within Brazil." *The Journal of Economic History* 72.2 (2012): 393-422.

Nunn, Nathan. "The Long-Term Effects of Africa's Slave Trades." *The Quarterly Journal of Economics* 123.1 (2008): 139-176.

Pischke, Steve, ”Lecture Notes on Measurement Error” Spring 2007, London School of Economics.

Mikusheva, Anna, and Brian P. Poi. ”Tests and Confidence Sets with Correct Size when instruments are Potentially Weak.” *The Stata Journal* 6.3 (2006): 335-347.

Roessler, Philip, Pengl, Y. I., Marty, R., Titlow, K. S. and van de Walle. ”The Cash Crop Revolution, Colonialism and Legacies of Spatial Inequality: Evidence from Africa”. No. 2020-12. Centre for the Study of African Economies, University of Oxford, 2020.

Roy, Tirthankar. ”Indigo and Law in Colonial India.” *The Economic History Review* 64 (2011): 60-75.

Sokoloff, Kenneth L., and Stanley L. Engerman. ”Institutions, Factor Endowments, and Paths of Development in the New World.” *Journal of Economic Perspectives* 14.3 (2000): 217-232.

Vogel, Stephen J. ”Structural Changes in Agriculture: Production Linkages and Agricultural Demand-led Industrialization.” *Oxford Economic Papers* 46.1 (1994): 136-156.

Wahi, Namita. ”Understanding Land Conflict in India and Suggestions for Reform.” *CEPR* (2019).

Appendix

A Mapping Historic Districts in 1897 to Contemporary Districts in 2011

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Cachar	Cachar	48.07
	Hailakandi	17.96
	Karimganj	33.98
Goalpara	Bongagigon	22.91
	Dhubri	38.25
	Goalpara	19.48
	Kokrajhar	20.35
Kamrup	Barpeta	33.95
	Kamrup	42.1
	Nalbari	23.86
Darrang	Darrang	48.07
	Sonitpur	51.93
Nagaon	Marigaon	25.22
	Nagaon	74.78
Sibsagar	Golaghat	28.05
	Jorhat	35.62
	Sibsagar	35.87
Lakhimpur	Dibugarh	0.17
	Tinsukia	0.15
	Dhemaji	32.32
	Lakhimpur	67.34
Khasi and Jaintia Hills	Ri-Boi	12.26
	Jaintia Hills	18.77
	West Khasi Hills	18.32
	East Khasi Hills	62. 91
Garo Hills	East Garo Hills	25.26
	West Garo Hills	74.74
Burdwan	Burdwan	100
Bankura	Bankura	100
Birbhum	Birbhum	100
Medinipur	Medinipur	100
Hooghly	Hugli	100
24 Parganas	Kolkata	19.83
	North Parganas	44.15
	South Parganas	36
Nadia	Nadia	100
Murshidabad	Murshidabad	100
Darjeeling	Darjeeling	100
Jalpaiguri	Jalpaigur	100
Tripura	North Tripura	25.69
	South Tripura	24.05
	West Tripura	50.26
Patna	Patna	36.72
	Nalanda	63.28
Gaya	Aurangabad	22.8
	Gaya	39.21
	Jehnabad	17.93
	Nawada	20.05
Shahapur	Bhojpur	50.65
	Rohtas	49.35
Darbhanga	Darbhanga	31.01
	Madhubani	36.15
	Samastipur	32.84
Muzaffarpur	Muzaffarpur	39.44
	Sitamarhi	32.69
	Vaishali	27.87

Table A1: Crosswalking Contemporary Districts to Historic Districts

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Saran	Gopalganj	25.88
	Saran	39.55
	Siwan	34.17
Champaran	Paschim Champaran	44.79
	Purba Champaran	55.21
Monghyr	Begusarai	29.48
	Jamui	24.23
	Lakhisarai	14.89
	Munger	21.74
	Sheikhpura	9.77
Bhagalpur	Banka	40.36
	Bhagalpur	59.64
Purnia	Katihar	28.24
	Purnia	30.04
	Araria	25.31
	Kishanganj	15.93
Malda	Malda	100
Santhal Parganas	Deoghar	19.07
	Dumka	32.7
	Godda	19.19
	Sahibganj	16.44
	Pakaur	12.65
Cuttack	Cuttack	37.18
	Jagatsingh	16.91
	Jajpur	25.1
	Kendrapara	20.81
Balasore	Baleshwar	60.54
	Bhadrak	39.46
Angul	Angul	100
Puri	Puri	100
Hazaribagh	Chatra	21.55
	Hazaribagh	64.57
	Kodarma	13.88
Ranchi	Lohardaga	74.8
	Gumla	33.13
	Ranchi	59.39
Palamow	Garhwa	32.69
	Palamu	67.31
Manbhum	Dhanbad	28.5
	Bokaro	10.6
	Purulia	60.9

Table A2: Crosswalking Contemporary Districts to Historic Districts

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Singhbum	Paschim Singhbum	51.82
	Purbhi Singhbum	48.18
Dehradun	Dehradun	100
Saharanpur	Saharanpur	68.13
	Hardwar	31.87
Muzaffarnagar	Muzaffarnagar	100
Meerut	Baghpat	29.88
	Meerut	70.12
Bulandshahr	Bulandshahr	86.61
	Gautum Buddha Nagar	13.39
Aligarh	Aligarh	74.32
	Hathras	25.68
Muttra	Mathura	85.47
	Hathras	14.53
Agra	Agra	79.11
	Firozabad	20.89
Farrukhabad	Farrukhabad	52.63
	Kannauj	47.37
Mainpuri	Mainpuri	61.52
	Firozabad	38.48
Etwah	Etwah	52.93
	Auralia	47.07
Etah	Etah	55.3
	Kasanganj	44.7
Bareilly	Bareilly	100
Bijnor	Bijnor	100
Budaun	Budaun	100
Moradabad	Jyotiba Phule Nagar	28.04
	Moradabad	71.96
Shahjahanpur	Shahjahanpur	100
Pilibhit	Pilibhit	100
Kanpur	Kanpur Dehat	47.87
	Kanpur Nagar	52.13
Fatehpur	Fatehpur	100
Banda	Banda	67.99
	Chitrakoot	32.01
Hamirpur	Hamipur	60.31
	Mahoba	39.69
Allahabad	Allahabad	79.22
	Kaushambi	20.88
Jhansi	Jhansi	66.57
	Lalitpur	33.43

Table A3: Crosswalking Contemporary Districts to Historic Districts

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Jalaun	Jalaun	100
Benares	Varanasi	51.6
	Chaundali	26.23
	Sant Ravidas Nagar	22.17
Mirzapur	Mirzapur	61.85
	Sonbhadra	38.15
Jaunpur	Jaunpur	100
Ghazipur	Ghazipur	100
Ballia	Ballia	95.11
	Mau	4.89
Gorakhpur	Gorakhpur	64.83
	Maharaganj	35.17
Basti	Basti	61.43
	Siddharth Nagar	38.57
Azamgarh	Azamgarh	60.9
	Mau	29.1
Almora	Almora	72.7
	Bageshwar	27.3
Garhwal	Garhwal	36.58
	Tehri Garhwal	32.94
	Rudra Prayag	12.89
	Uttarkashi	17.57
Nainital	Champawat	2.67
	Nainital	27.28
	Udham Singh Nagar	60.05
Lucknow	Lucknow	100
Unnao	Unnao	100
Rae Bareli	Rae Bareli	100
Sitapur	Sitapur	100
Hardoi	Hardoi	100
Kheri	Kheri	100
Faizabad	Faizabad	46.14
	Ambedkar Nagar	53.86
Gonda	Balrampur	38.73
	Gonda	61.7
Bahraich	Bahraich	66.59
	Shrawsti	33.41
Sultanpur	Sultanpur	100
Pratapgarh	Pratapgarh	100
Barabanki	Barabanki	100
Ajmer	Ajmer	100

Table A4: Crosswalking Contemporary Districts to Historic Districts

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Hissar	Hissar	100
Rohtak	Bhiwani	1.43
	Jhajjar	38.84
	Rohtak	42.96
	Sonipat	16.05
Gurgaon	Gurgaon	38.84
	Faridabad	41.91
	Mahendragarh	19.63
Karnal	Jind	6.56
	Karnal	49.89
	Kurukshetra	43.55
Ambala	Amabala	61.37
	Yamunanagar	38,63
Shimala	Shimla	33.16
	Solan	66.84
Kangra	Hamirpur	19.97
	Kangra	60.34
	Una	19.69
Hoshiarpur	Hoshiarpur	89.26
	Nawanshahr	10.57
Jalandhar	Jalandhar	100
Ludhiana	Ludhiana	100
Ferozpur	Firozpur	54.83
	Faridkot	45.17
Amritsar	Amritsar	100
Gurdaspur	Gurdaspur	100
Ahmadabad	Ahmadabad	100
Kaira	Kheda	51.93
	Anand	48.07
Panchmahals	Panch Mahals	56.9
	Dohad	43.1
Khandesh	Jalgaon	53.33
	Dhule	25.68
	Nandurbar	20.78
Nasik	Nasik	100
Ahmadnagar	Ahmadnagar	100
Pune	Pune	100
Solapur	Solapur	100
Satara	Satara	100
Belgaum	Belgaum	100
Bijapur	Bijapur	52.52
	Bagalkot	47.48
Dharwar	Dharwad	40.96
	Gadag	23.6
	Haveri	35.43
Tanna	Thane	100

Table A5: Crosswalking Contemporary Districts to Historic Districts

Historical District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Kolaba	Raigarh	100
Ratnagiri	Ratnagiri	65.35
	Sindhudurg	34.65
Kanara	Uttar Kannada	100
Sagar	Sagar	100
Damoh	Damoh	100
Jabalapur	Jabalpur	66.72
	Katni	33.28
Mandla	Mandla	60.36
	Dindori	39.64
Seoni	Seoni	100
Narsinghpur	Narsimhapur	100
Hoshangabad	Harda	8.43
	Hoshangabad	91.57
Nimar	Khargone	47.52
	Khandwa	33.24
	Burhanpur	19.23
Betul	Betul	100
Chhindwara	Chhhindwara	100
Wardha	Wardha	100
Nagpur	Nagpur	100
Chanda	Chandrapur	69
	Gadchiroli	31
Bhandra	Bhandara	48.46
	Gondiya	51.54
Balaghat	Balaghat	100
Raipur	Dhamtari	15.04
	Mahasamund	20.25
	Raipur	64.72
Bilaspur	Bilaspur	44.68
	Janjgir Champa	29.27
	Kawardha	4.29
	Korba	21.77
Buldhana	Buldhana	100
Wani	Yavatmal	100
Sambalpur	Bargarh	44.76
	Debagarh	8.06
	Jharsuguda	16.39
	Sambalpur	30.16
Akola	Akola	60.16
	Washim	38.94
Amravati	Amravati	100
Ganjam	Ganjam	85.6
	Gajapati	14.4
Visakhapatnam	Visakhapatname	72.57
	Vizianagram	27.43
Godavari	West Godavari	43.31
	East Godavari	56.68

Table A6: Crosswalking Contemporary Districts to Historical Districts

Historic District in 1897	Contemporary District in 2011	Percentage of 2011 Population
Krishna	Krishna	48.09
	Guntur	51.9
Nellore	Nellore	68.77
	Prakasam	31.23
Kurnool	Kurnool	100
Ballari	Bellary	87.61
	Davanagere	12.39
Anantapur	Anantapuram	100
Kadapa	Cuddapah	100
North Arcot	Vellore	59.55
	Tiruvannamalai	40.45
Chingleput	Kanchipuram	100
Madras	Chennai	100
Salem	Salem	40.51
	Dharmapuri	17.52
	Krishnagiri	21.87
	Namakkal	20.09
Coimbatore	Coimbatore	58.91
	Erode	41.09
Trichinopoly	Tiruchirappalii	53.08
	Ariyalur	15.38
	Karur	20.64
	Peramblur	10.9
Tanjore	Thanjavur	45.32
	Nagapattinam	30.4
	Thiruvarur	24.28
Madurai	Theni	30.42
	Madurai	45.58
	Dindigul Anna	23.99
Tinnovelly	Tirunelveli	62.21
	Thoothukuddi	37.79
Nilgirs	Nilgiri	100
Malabar	Kannur	23.93
	Kozhikode	29.28
	Wayanad	7.75
	Malappuram	39.02
South Canara	Dakshina Kannada	45.68
	Udupi	25.73
	Kasaragod	28.58

Table A7: Crosswalking Contemporary Districts to Historic Districts