Displacement and Development: Long Term Impacts of the Partition of India *

Prashant Bharadwaj† Rinchan Ali Mirza‡

Preliminary Draft as of September 2016. Please do not circulate.

Abstract

The partition of British India in 1947 resulted in one of the largest and most rapid migrations in human history. This paper examines how areas affected by the partition fare in the long run. Using migrant presence as a proxy for the intensity of the impact of the partition, and district level data on agricultural output between 1911-2009, we find that areas that received more migrants have higher average yields, are more likely to take up high yielding varieties (HYV) of seeds, and are more likely to use agricultural technologies. These correlations are more pronounced after the Green Revolution in India. Using pre-partition data, we show that migrant placement is uncorrelated with soil conditions, agricultural infrastructure, and agricultural yields *prior* to 1947; hence, the effects are not solely explained by selective migration into districts with a higher potential for agricultural development. Migrants moving to India were more educated than both the natives who stayed and the migrants who moved out. Given the positive association of education with the adoption of high yielding varieties of seeds we highlight the presence of educated migrants during the timing of the Green Revolution as a potential pathway for the observed effects.

*This project has benefitted from comments by Latika Chaudhary, James Fenske, Bishnupriya Gupta, Takashi Kurosaki, and participants at the World Economic History Congress, 2015.

†University of California San Diego & NBER. Corresponding author: Economics Dept, 9500 Gilman Dr. #0508, La Jolla, CA 92093-0508, US. Phone: +1-858-822-6760. E-mail: prbharadwaj@ucsd.edu

‡University of Oxford. Faculty of History. George Street. Oxford. OX1 2RL. UK. Phone: +44-747-021-1816. E-mail: rinchan.mirza@history.ox.ac.uk. Website: www.rinchanmirza.com

1 Introduction

The end of the British Empire in India in 1947 was marked with an unprecedented mass migration of nearly 17 million people and a human rights disaster involving nearly a million deaths in the wake of the riots that ensued between Hindus and Muslims on either side of the newly created India-Pakistan border. The emergence from nearly a century of colonial rule left an indelible mark on Indian history as historical events undoubtedly shape modern day institutions and development (Acemoglu, Hassan, and Robinson, 2011; Nunn, 2008; Banerjee and Iyer, 2005; Acemoglu, Johnson, and Robinson, 2002; Chaney and Hornbeck, 2015; Dippel, 2014; Dell, 2010). This paper seeks to examine the legacy of the partition on an important aspect of economic progress – agricultural development.

Using migrant presence as a proxy for the intensity of the impact of the partition, this paper highlights important correlations between areas that received migrants and subsequent agricultural development in those areas. Documenting these correlations is an important contribution as mass migrations, institutional upheaval, and partitions are a reality even today.[1] It is crucial therefore to understand how communities and areas develop long after such events take place. While affected areas suffer in the short run, it is critical to document whether the legacy of such events forever change the long run trajectory of these places.

We find that the partition had a statistically significant but moderate impact on agricultural development in the decades after India's independence. Between 1957 and 2009, areas that received migrants saw average annual wheat yields increase by 3.2% compared to areas that received no migrants. We find similar results when examining annual revenue per hectare;[2] this measure is used so as to not be reliant on any specific crop for our productivity measures. In the case of wheat, we find that most of the effects of

[1] The current (as of September 2015) refugee crisis in Europe is a relevant example of a mass migration with the potential to affect labor markets and economic development of receiving countries. The most recent example of a partition is that of Sudan where in a referendum held in January 2011 in the south showed that an overwhelming 98.8 percent of the population were in favor of secession. As a consequence, constitutional declaration of the independence of South Sudan took place on 9 July 2011. The other recent example is the Dayton peace agreement of November 1995, which led to the partition of Bosnia and brought an end to the Bosnian War. Yet another prominent example is the partition of Cyprus into Greek and Turkish speaking separate territorial units after the Turkish invasion and occupation of Northern Cyprus in 1974 (Christopher, 2011; Kumar, 2004; Kliot and Mansfield, 1997).

[2] This measure uses data on the production of wheat, rice, sugar, jowar, maize, bajra, barley, cotton, groundnut, jute gram, potato, ragi, rapeseed, mustard, sesame, soybean, sugarcane, sunflower, tobacco, tur and other pulses.

partition are concentrated after the green revolution. The green revolution transformed Indian agriculture in the 1960s, making crops less susceptible to destruction via pests and droughts, increasing yields and increasing land-based investments like irrigation. We find that migrant presence is also strongly correlated with the use of tractors (a 10% increase in migrants is correlated with a 2% increase in the use of tractors between 1957-1987) and fertilizers (phosphorous and nitrogen), which is in line with the idea that partition-affected districts were more likely to adopt HYVs and other technologies related to the green revolution. These results are not just driven by migration into agriculturally suitable states like the Punjab – the results are robust to the inclusion of state fixed effects and state specific time trends.

A key aspect of our empirical framework uses agricultural data from before 1951, and employs a difference in differences design for a subset of districts for whom such data is available to examine the impact of partition affected districts on long run agricultural outcomes. A key concern with examining simple correlations of migrant presence and outcomes is that despite the uncertainty and chaos of partition, migrants might have moved to places pre-disposed to agricultural growth. Hence, the ability to use extensive pre-partition agricultural data goes a long way in ensuring that districts that were affected by partition related migration are not on differential trends until the start of the green revolution. While limited in our ability to examine *trends* along certain variables, we use available data to examine at least in levels whether migrants went to more endowed districts along dimensions that might matter for agricultural development. For example, canals and tube wells were important characteristics that allowed for the spread of high yielding varieties of seeds; however, we find no correlation between pre-partition canal irrigation and migrant presence. We also find no correlation between migration and the presence of other infrastructure variables such as banks, post offices, length of roads, and hospitals by 1961 (pre green revolution). Migrant presence is also uncorrelated with the *growth* in the literate population prior to partition. This mitigates the concern that even if migrants did not choose districts based on agricultural yields, they might have chosen districts based on some characteristic that happened to be extremely important for the spread of the green revolution (like roads, banks, or schooling).

However, we might still be concerned that migrant presence might be generally related to district characteristics or trends that affect agricultural yields. Our results on migrant presence and yields are however *only* present for crops that were affected by the green revolution. Non green revolution crops, such as certain millets, sugarcane, etc. do not show any changes in yields with migrant presence. Finally, we are able to account for an important institutional feature of the British colonial system that has been shown to affect agricultural yields and the take up of the green revolution - the British taxation system on agricultural lands. Using data from Banerjee and Iyer (2005), we are able to control for these features, and find that adding these controls does not affect out main estimates.

While we believe these results to be important, we want to be upfront about the scope and limitations of this paper. This research is motivated by the goal of linking partition to subsequent economic development (as measured by income, health and human capital); however, in this paper we specifically (and only) examine agricultural outcomes. There are two main reasons for this: first, agricultural outcomes are available at a yearly level, at fine levels of administrative disaggregation, and over a long period of time – the same is not true of many other variables of interest to development economists like health, income, etc. Second, agriculture was, and still is, an important part of employment and economic output in India.[3]

A second limitation of this study is that the partition was an event that resulted in many changes: two way migrations along two new borders, new governments, mass deaths, demographic changes, and loss and restructuring of land, just to name a few. It is difficult therefore to have a single variable that captures all of these forces, or even obtain data on most of these individual changes. Our way of measuring the intensity of the impact of partition is to use displaced person population in 1951. By this, we assume that areas (districts in our case) that received migrants due to partition were more "affected" along various dimensions, like the ones we just mentioned, by partition than districts that did not receive any migrants. While we use displaced person population as our metric for the intensity of the impact of partition, we do not wish to interpret our results as solely the effects of partition induced *migration*. For example, districts with more migrants could have received more government aid in the years after partition, and our effects should be interpreted as capturing the effect of both migration and government assistance. While this is a potential downside of our study, we wish to point out that this issue is present in all studies of mass migrations. Mass migrations or refugee movements, by their very nature, induce all kinds of responses on the part of sending and receiving governments

[3] In 2014 approximately 17 percent of Indian GDP was made up of the agricultural sector and for the decade prior to that it fluctuated between 18 and 17 percent. In 2012 as much as 47 percent of the total Indian workforce was employed in agriculture (data from World Bank Economic Indicators).

and communities. Hence, these effects are not to be confused with or compared to those arising from the economic migration of people.

While it is imperative to uncover the precise mechanisms or governmental policies (or lack thereof) that led to these long run effects, we are unable to do so in this paper. Data limitations, the sheer magnitude of the event, and the two way nature of the migratory flows makes it nearly impossible to make precise statements about any *one* leading factor. We do, however, provide some preliminary evidence that the composition of migrants played a qualitatively important role in the future agricultural development of more affected areas. Migrants who moved to India were more educated than the natives who stayed behind (Bharadwaj, Khwaja, and Mian, 2009). Given the positive correlation between education and the better use and take up of agricultural technologies (Feder, Just, and Zilberman, 1985), the demographic changes induced by partition could be a plausible mechanism for the effects seen. The migrants were also more likely to have been involved in money lending and other commercial aspects of farming. Since credit is an important aspect of agriculture and especially so for the take up of newer technologies it is likely that the presence of migrants during the green revolution helped along this dimension as well. Note that since partition resulted in two way migration flows (Muslims leaving India, replaced by Hindus and Sikhs arriving from Pakistan/East Pakistan), the main mechanism for agricultural development is unlikely to be the same as identified by Hornbeck and Naidu (2014) in the case of the American South.

This paper contributes to the economics literature on the long term impacts of historical events in general (see Nunn (2009) for a review), and also to the literature more focussed on the impacts of history and colonization in India (Jha, 2013; Chaudhary and Rubin, 2011; Donaldson, 2010; Iyer, 2010). Most closely related is the work of Banerjee and Iyer (2005), who show that different institutions (specifically practices regarding land rights) during the colonial period had a profound impact on agricultural development long after the British left India. They find that these institutions played an important role after the green revolution, where individual rights to ownership of land were a crucial aspect of districts that were able to take advantage of HYV seeds, fertilizers, and other agricultural technologies. This paper also builds on and extends the research that is directly related to the partition of India (Bharadwaj, Khwaja, and Mian, 2009; Jha and Wilkinson, 2012; Bharadwaj and Fenske, 2012). While these papers contribute in important ways to our understanding of the event by analyzing the demographic consequences of partition, the role of combat experience during WWII on ethnic cleansing during the partition, and the impact of partition related migratory movement on jute cultivation, they do not examine long run consequences. Hence, the main contribution of this work is to examine how partition (as measured by the presence of displaced populations) impacted long run economic outcomes such as agricultural development.

2 Background

2.1 Partition of India - adapted from Bharadwaj, Khwaja and Mian 2009

While the possibility that British India would be partitioned once it gained independence from the British was present several years prior to the actual event, with the Muslim League formally calling for a separate Muslim state by 1940, its details weren't worked out till a few months prior to partition and the actual plan was not made public till a few days after the two countries had been declared independent.

The partition plan of June 3rd, 1947 laid the foundations for the redrawing of the boundaries of the Punjab and Bengal. Sir Cyril Radcliffe chaired both the Bengal and Punjab Boundary Commissions. Radcliffe was a lawyer by profession and unfamiliar with boundary making ? his selection as chairman of the boundary commission was based on his impartial relations with India. Along with this impartiality, however, came a lack of intimate knowledge of the people and the land he was about to carve up (Kudaisya & Yong Tan, 2000: 84). His first meeting with the then Viceroy, Lord Mountbatten took place on June 8th and Radcliffe was shocked when he was told that he had only five weeks to draw the lines. The ambiguities associated with the terms that would determine the boundary making process further complicated Radcliffe's task. While the political parties had agreed on partition, they had vaguely laid down that boundaries would be demarcated by contiguous majority areas of Muslims and Non-Muslims as well as considering "other factors". This clause – "other factors" – probably caused the most controversy during the entire process of boundary making. The idea of "contiguous areas" was also vague as it was not certain whether this meant districts or tehsils.

The boundary decisions were kept secret until the last minute and this heightened speculation regarding Radcliffe's methods of demarcating the border. It was also alleged that Radcliffe used the 1941 census to calculate religious majorities in various districts. Since the decision for a separate Muslim state was released in 1940, many feared that the 1941 census was rigged and under reported the presence of certain religious groups. In a hasty 2 months, undivided India was carved into the independent states of India and Pakistan. The Radcliffe Award, as the boundary commission's reports were called, caused more controversy than the peace they were intended for. In some ways, "no man made boundary has caused so many troubles and effectively impeded the advent of peace in South Asia as the Punjab boundary resulting from the Commission's verdict" (Cheema, 2000:1). The Commission's report was made public on August 17th, 1947, two days after Indian independence.

When the Radcliffe award was finally made public, there were voices of dissent everywhere. Radcliffe, well aware of the criticisms he would face, admitted: "The many factors that bore upon each problem were not ponderable in their effect upon each other. The effective weight given to each other was a matter of judgment, which under the circumstances threw it upon me to form; each decision at each point was debatable and formed of necessity under great pressure of time, conditions, and with knowledge that, in any ideal sense, was deficient." (Kudaisya & Yong Tan, 2000: 93) The myriad factors that Radcliffe had to consider in a short period of time, made the boundary decision process illogical and inconsistent at times. He later lamented, "Nobody in India will love me for the award about the Punjab and Bengal and there will be roughly 80 million people with a grievance who will begin looking for me" (Khilnani 1997: 201). In short the boundary decisions were bound to cause problems.

As a result, on August 17th, 1947 when the award was made known, thousands found themselves on the wrong side of the border, particularly in the state of Punjab. There were neither provisions nor preparations for the affected populations to be evacuated, until it was too late (Kudaisya, 2000: 98). The widespread violence, migrations and human suffering was unprecedented. Even before the declaration of independence, the violence in Punjab had started to take its toll. In March 1947 the scale of rioting was such that several thousand villagers in Lahore and Rawalpindi districts were forced to leave their villages. Together these factors meant that the majority of migratory flows took place under a relatively short span of time. Since the boundaries were not declared till later and there was a lot of uncertainty regarding them, it was unlikely that people moved much before partition.

2.2 Green Revolution in India

2.2.1 Development of High Yielding Varieties

The Green Revolution originates from the crossbreeding experiments carried out at the International Rice Research Institute (IRRI), in the Philippines in 1961, and its sister institution, the International Centre for Maize and Wheat Improvement (CIMMYT) in Mexico in 1967. The objective of the experiments was to develop ?shorter, stiff strawed varieties of the wheat and rice crops that devoted much of their ?energy to producing grain and relatively little to producing straw or leaf material? (Evenson and Gollin, 2003a, p .758). The two most successful hybrid varieties that came out of these early experiments were the IR8 and the Norin10-Brevor. The first was a hybrid rice variety that was developed by cross breeding an Indonesian variety called "Peta" with a semi-dwarf variety from Taiwan called "Dee-Geo-Woo-Gen" (Gollin, Hansen, and Wingender, 2016, p. 8). The second was a hybrid wheat variety that was a cross between a short variety developed in Japan in the 1930s called "Norin10", and an American variety called "Brevor" (Gollin, Hansen, and Wingender, 2016, p. 9).

The development of HYVs of crops other than rice and wheat took longer and was not as impressive as that of rice and wheat. This was because scientists had already developed a critical mass of knowledge for rice and wheat in particular, which did not exist for other crops (Evenson and Gollin, 2003a). As late as the 1980s only a few HYVs of crops like sorghum and millet had been developed (Evenson and Gollin, 2003a, p. 758). The differences in the initial stock of scientific knowledge of crops meant that the benefits of HYV adoption in terms of increasing agricultural productivity were largely concentrated to households producing wheat and rice.

2.2.2 Diffusion of rice and wheat HYVs in India

A selection of the hybrid varieties developed at the IRRI and CIMMYT were imported into India where they were further crossed with local varieties to adapt to local conditions. Out of these crosses came the locally adapted rice varieties of "Padma" and "Jaya" and wheat varieties of "Kalyan Sona" and "Sonalika". It was the mass scale release of such locally adapted varieties in India in the late 1960s that marked the start of the Green Revolution.

The wheat varieties of Kalyan Sona and Sonalika were an immediate success and were quickly adopted in the three main wheat-growing regions of India: the Northwest Plains, the Northeast Plains and the Central Peninsular zone.[4] In particular, the production of wheat went up from twelve million tons to twenty million tons between 1966-67 and 1969-70?an increase of 40% in a span of just three years (Chakravarti, 1973, p. 321). The success of the varieties was due to their robustness to the varying conditions under which wheat is grown in India (Munshi, 2004, p. 187). Building on the success of the early wheat varieties, agricultural scientists began concentrating their research on developing new varieties for what were termed "marginal environments". Marginal environments included low rainfall areas with limited or no irrigation infrastructure. As a consequence of research efforts, the new wheat varieties were able to increase adoption in marginal environments in the later phases of the Green Revolution (Byerlee and Moya, 1993, p. XI).

In contrast to the early wheat varieties, the early rice varieties of Padma and Jaya were only marginally successful in penetrating the rice growing areas of India. Both varieties were unsuitable in a variety of stress conditions such as water logging, salinity and drought (Munshi, 2004, p. 190). They were also found to be susceptible to a number of pests and diseases prevalent in the rice growing areas (Munshi, 2004, p. 190). Due to the failure of the early rice varieties Indian agricultural scientists concentrated their research efforts on developing varieties that were both 'suited to specific local conditions of areas where rice was grown (Munshi, 2004, p. 190) and also incorporated resistance to pests and diseases (Evenson and Gollin, 2003a, p. 759).

In addition to the differences in the technological development discussed above, other important factors influenced the adoption of rice and wheat HYVs across regions of India. An important factor in the adoption of the HYVs was the provision of timely irrigation (Rud, 2012, p. 353). This was because the uninterrupted supply of water at specific periods of growth, development and flowering was crucial to the successful performance of the HYVs. Hence, pre-existing patterns of irrigation and climate were the main drivers behind the diffusion of the HYVs (Gollin et. al., 2016, p. 5). The importance of irrigation can be gauged by the fact that states like Punjab, Haryana and Tamil Nadu, where the HYV adoption was already very high in the 1970s, all had in common a well developed

[4] The states of Punjab, Haryana, (western) Uttar Pradesh, Delhi and Rajasthan make up the Northwest Plains region. The Northeast Plains region includes (eastern) Uttar Pradesh, Bihar, Orissa and Bengal. Finally, the Central Peninsular zone is made up of the states of Madhya Pradesh and Gujarat.

irrigation infrastructure dating back to the colonial period (Evenson and Gollin, 2003b, p. 91). On the other hand, states like Gujarat, Maharashtra, Orissa, West Bengal, Bihar, Kerala and Rajasthan, which did not have an extensive colonial canal network, lagged behind considerably in terms of HYV adoption as late as 1973-74 (NCAI Part 1, 1976, p. 284).

Recognizing the importance of irrigation, the post-independence Indian State substantially expanded the irrigation infrastructure built during the colonial era. A number of new canal irrigation projects like the Bhakra-Nangal, the Damodar Valley and the Hirakud were taken up in the period immediately after Independence (NCAI Part V, 1976, p. 14). The net irrigated area increased from 20.9 million hectares in 1950-51 to 27.7 million hectares in 1960-61 as a result of such projects (NCAI Part 1, 1976, p. 201). An important feature of such an expansion was that the historically canal-irrigated states experienced greater growth in irrigated area relative to other states. The main reason being that regions outside the historically canal irrigated states did not possess the topography or the river systems that were crucial to the construction of irrigation canals.

However, beginning in the late 1960s with the advent of the Green Revolution, more minor irrigation projects were undertaken to rapidly expand irrigation beyond the historically canal-irrigated states. The expansion came in the shape of electrified tubewells using groundwater as opposed to river water for irrigation purposes (Rud, 2012, p. 353). Such minor works were categorised as high priority from the end of the third Five Year Plan in 1965-66. Bharadwaj (1990) notes that the rate of increase in irrigation by tubewells was higher than that of canals and accelerated remarkably between during the period between 1969 and 1980.

Several reasons are behind the shift in focus from canals towards electrified tubewells. One was the greater control electrified tubewells offered in terms of flow and timing of water supplies (NCAI Part V, 1976, p. 20). Another was the cost effectiveness that made tubewells particularly attractive for small and medium sized farmers in areas without canal irrigation. Finally, the state-financed extension of the electricity network across rural India and provision of credit to farmers were also important factors behind the increased use of electrified tubewells (NCAI Part V, 1976, p. 20).

3 Data and Empirical Framework

3.1 Post-Partition Data

For our post-partition analysis the data comes from three different sources: the 1951 census of India, the Indian Agriculture and Climate Dataset (i.e. IACD) and the Village Dynamics in South Asia Dataset (i.e. VDSA). The 1951 census data was used to construct a measure of displacement that was then related to measures of agricultural development from 1957 to 2009 that were constructed from data in the IACD and VDSA datasets.[5] An important task in relating the two measures was to make district boundaries comparable between 1951, the year in which displacement data was recorded, and the first year for which data is available in the combined IACD and VDSA panel dataset (i.e.1957).[6] For those districts that were partitioned between 1951 and 1957 we used a mapping procedure to achieve such a task. Our procedure involved the following steps. We first identified the districts that were created between 1951 and 1957. We called these are our child districts. We then identified the 1951 districts from which our child districts were created between 1951 and 1957. We called these our parent districts. We then recorded the areas of all our child and parent districts. Next, we divided the area of the child district by the area of its corresponding parent district to determine the proportion of the 1951 parent district that was made up of the child district. Finally we use the resulting proportions to estimate 1951 numbers for the child districts that were created between 1951 and 1957.

3.1.1 1951 Census of India

The 1951 census of India was carried out in the last three weeks of February 1951 with enumerators revisiting households from the 1st to the 3rd of March of the same year. It is significant for having recorded the initial and the most substantial phase of migration inflows that resulted from partition. A total of 7.3 million displaced in-migrants were enumerated, of whom 4.7 and 2.55 million had come from West and East Pakistan,

[5] In constructing our agricultural development measures from 1957 to 2009 we combined the IACD data from 1957 to 1965 with the VDSA data from 1966 to 2009. For the period where there was an overlap between the IACD and the VDSA (i.e. 1966 to 1987) we carried out empirical exercises to show that the data contained in both of them were not significantly different from each other.

[6] The district boundaries were kept constant for the period 1957 to 2009 in the combined IACD and VDSA panel. Therefore, making the 1951 district boundaries comparable with those in the first year of the combined IACD and VDSA panel (i.e. 1957) also makes them comparable with the boundaries in all the subsequent years of the panel (i.e. from 1958 to 2009).

respectively, and 0.05 million did not specify their place of origin (Visaria, 1969). Information on the in-migrants was disaggregated by gender, age, occupation and region of origin. In the case of sex, separate inflows were recorded for both males and females. According to Bharadwaj, Khwaja, and Mian (2009) the percentage of men in the inflows was, on average, 1.09 percentage points lower than the residents. For age structure, the refugees were classified in ten-year age groups going from ages 5-14 through 65-74. The region of origin for each in-migrant was identified as being either West or East Pakistan. In addition to demographic characteristics, there was also data on the occupation of in-migrants. Appendix II of Table IV of the census provides a detailed occupational classification of the in-migrants. Here again according to Bharadwaj, Khwaja, and Mian (2009) the in-migrants tended to engage more in non-agricultural professions relative to the resident population.

The 1951 census provides the best estimate to date of the spatial distribution of the immigration from Pakistan into India due to partition. That said, it does have some drawbacks. Firstly, the data on region of origin does not provide enough granularity to identify the district of West or East Pakistan from which an in-migrant came from. Secondly, substantial changes in the administrative machinery and the relatively unsettled conditions in those districts that received in-migrants casts doubt over the quality and coverage of the data (Visaria, 1969). On the other hand the multiple counting of persons crossing the border into India more than once caused an over reporting of in-migrants (Visaria, 1969). Finally, the high mortality rate amongst the refugees who arrived between 1947 and 1951 meant that the true scale of partition related displacement could not be established (Visaria, 1969).

3.1.2 Indian Agriculture and Climate Dataset

The Indian Agriculture and Climate Dataset is a panel dataset that covers 271 districts across thirteen states of India and includes annual data on agricultural, economic, climate and edaphic variables for the period 1957 to 1987. The states covered are Haryana, Punjab, Uttar Pradesh, Gujarat, Rajasthan, Bihar, Orissa, West Bengal, Andhra Pradesh, Tamil Nadu, Karnataka, Maharashtra and Madhya Pradesh. One of the key concerns that the compilers of the dataset addressed was to keep district boundaries constant between 1957 and 1987 so as to make the data comparable over time. They did so by taking into account all the changes in district boundaries that occurred between 1957 and

1987. More, specifically they preserved the original district boundaries by consolidating new districts created after the start date of the panel (i.e. 1957) into previous parent districts. For this reason the actual number of districts at the end of the panel period (i.e. 1987) is larger than the 271 districts contained in it.

In particular, the dataset includes annual information on the quantity produced of each crop (in tons), the area planted to each crop (in hectares), the area planted to high yield varieties of each major crop (in hectares) and the price of each crop. The quantity and price of the various inputs used in agriculture such as bullocks, tractors, and fertilizer (in tons) is also given. The climatic variables included are average monthly rainfall (in millimetres) and average monthly temperature (in degree celsius) for the period 1957 to 1987. Data from the population census is available on the number of persons, literacy, number of cultivators and the number of agriculture laborers. Finally, there is a set of 21 indicator variables each specifying a different soil quality type in the dataset.

3.1.3 Village Dynamics in South Asia Dataset

The Village Dynamics in South Asia Dataset is a panel dataset that covers 594 districts across nineteen states of India and includes annual district level data on agricultural, socioeconomic, climate, edaphic variables and agro-ecological variables for the period 1966 to 2009. It builds and expands on the thirteen states given in the IACD by including the six additional states that are Assam, Himachal Pradesh, Kerala, Chhattisgarh, Jharkhand and Uttarakhand. The dataset uses 1966 as the base year for its districts. Hence, data from child districts formed after 1966 are given back to their respective parent districts to form a comparable sample of districts from 1966 to 2009 that is based on 1966 district boundaries. This is the same process of consolidating child districts into their parent districts that is used by the IACD dataset.

Specifically, the VDSA dataset includes annual information on crop area (in hectares) and production (in tons), price of crops (in rupees), area planted to high yield varieties of each major crop (in hectares), irrigated area, livestock, agricultural implements, annual rainfall (in millimetres), fertilizer consumption (in tons) and operational holdings. It also contains population census data on number of persons, literacy, number of cultivators and the number of agricultural laborers.

3.1.4 Combining the IACD with the VDSA datasets

In constructing our panel we combined the data on the thirteen states contained in the IACD dataset from 1957 to 1965 with the data on the same thirteen states in the VDSA dataset from 1966 to 2009. For the period where the two datasets overlapped (i.e. 1966 to 1987) we used the data from the VDSA dataset. A concern here was that for the overlapping period the data in the IACD dataset could be significantly different from the data in the VDSA dataset. We carried out two empirical exercises to show that this is not the case. Firstly, in Figures A.1 and A.2 we show that the correlation between the data on the annual wheat yields and the annual proportion of wheat HYV in the two datasets are quite high. Secondly, in Appendix Table 4b we show regressions for annual wheat yields and annual proportion wheat HYVs that exclude observations that are zero in one of the datasets and non-zero in the other. As is clear from the results, dropping observations that are not similar across the two datasets does not reduce the significance or the magnitude of our results.

3.2 Pre-Partition Data

3.2.1 Agricultural Statistics of British India

For our pre-partition analysis we use the Agricultural Statistics Reports of British India to extract information on yields for each of the major crops: Wheat, Rice, Sugar and Maize. The reports were produced on an annual basis by the Department of Revenue and Agriculture of the Colonial government. They contained information on yields for all major crops and most other crops for districts of British India and a select group of princely states. Although the reports came out on an annual basis, the yield numbers were only revised every five years. Therefore the pre-partition panel dataset we constructed contains information on yields for only four years during the period stretching from 1910 to 1940.[7] The colonial government started recording rough estimates of acreage and production of the major crops from as early as 1861. However, a concerted effort to systematically collect such information on most crops only began in 1891-92 (Heston, 1973). Our selection of 1910 as the starting point of our pre-partition panel was determined by the substandard quality of data prior to that date.

[7]To be more precise the exact years are 1911, 1921, 1932 and 1938.

3.3 Empirical Specification

3.3.1 Sub sample difference in differences

Ideally, the empirical specification used in this paper would account for pre-existing trends in agricultural development in areas that eventually received migrants relative to areas that did not receive migrants. As mentioned in Section 2, for a smaller sample of our data, we were able to obtain agricultural data starting in 1911 although not at the yearly level. For this subsample of districts (Table 1B compares the districts in this sub sample to the overall sample), we estimate the following regression:

$$Y_{ist} = \beta D_{is}^{51} + \theta Post_t + \gamma D_{is}^{51} \times Post_t + \mu Z_{is} + \zeta_s \times t + \zeta_s + \alpha_t + \epsilon_{ist} \quad (1)$$

Y_{ist} represents the yields of a specific crop (say wheat) in district i, in state s, at time t. D_{is}^{51} is the main independent variable of interest, the log number of displaced persons (in-migrants) in the district i (this is measured at a single time period in 1951), and its interaction with time (either via a single "Post" dummy, or simply year dummies in a more flexible specification), which is captured by the coefficient γ. In order to control for the overall size of the district in 1951 we also crucially include the log of the population of that district in 1951 (Pop_{is}^{51}). Z_{ist} is a vector of controls representing agricultural characteristics of the district like soil types (soil types do not vary over time in the district), altitude, latitude and longitude. As mentioned earlier, the IACD data contains information on soil types at the district level, and we control for these as soil conditions might play an important role in the adoption of agricultural technology and agricultural productivity. Finally, we control for broader time-invariant characteristics at the state level with state fixed effects (ζ_s), for country level year specific effects with calendar year fixed effects (α_t), and also for state-specific time varying characteristics with state-time trends ($\zeta_s \times t$) that are split to capture state trends pre and post independence. We cluster our standard errors at the district level.

3.3.2 Full sample panel regressions

Our main estimating equation for samples where the data does not extend to the pre-partition time period is the following:

$$Y_{ist} = \beta D_{is}^{51} + \delta Pop_{is}^{51} + \gamma Den_{is}^{61} + \mu Z_{ist} + \zeta_s \times t + \zeta_s + \alpha_t + \epsilon_{ist} \quad (2)$$

Y_{ist} represents the outcome of interest in district i, in state s, at time t. We examine 2 *crop specific* agricultural outcomes: yield and HYV adoption (defined as acreage using HYV seeds divided by the total amount of land under cultivation). In order to compare districts that grow different crops, we use an overall revenue based measure as well. This measure uses a single calendar year price (in our case 1960)[8] for each crop produced in the district, then dividing by the area under cultivation in that district to construct our "revenue per acre" measure. In some specifications, we also use other measures of technology adoption like tractors and fertilizer use per acre, and pre-partition yields to examine the role of migrant placement based on land productivity prior to partition. While our main specification uses the data in panel form, an analogous specification would be to collapse the data at the district level by taking averages for the entire period for which we have agricultural data, or for specific decades or years. This would analyse cross-sectional variation. Not surprisingly, the results with the cross sectional approach are similar and presented in the appendix. The main advantage of the panel form is in our examination of the effect of displaced persons after the green revolution. In some specifications, we interact D_{is}^{51} with the calendar year in the district when the acreage under HYV exceeds 5% (our approximate measure of when the green revolution started in that district). The interaction thus represents the differential impacts due to partition on agricultural outcomes after the start of the green revolution and in many ways is similar to the difference in difference specification used for the sub sample analysis.

Although we claim to not highlight a causal link between migration and agricultural development in this paper, it is useful (perhaps for future work in this area) to consider some of the biases that might be present when estimating equations 1 and 2. A leading candidate for a variable that we do not measure, but that could be correlated with both displaced persons and agricultural development is government intervention or aid for migrants. There were many programs set up by the government to help with refugee resettlement (land redistribution for example) and these programs could have had direct

[8]We do this to avoid the fact that production in any given year can affect prices.

bearing on agriculture as well. Our estimates on displaced persons in the above estimating equations therefore represent a reduced form or "net" effect of migration and associated changes due to migration on agricultural development. Such an interpretation is still useful, as rarely in the world would a mass movement of people take place without other, simultaneous responses (either by governments or by people in receiving countries).

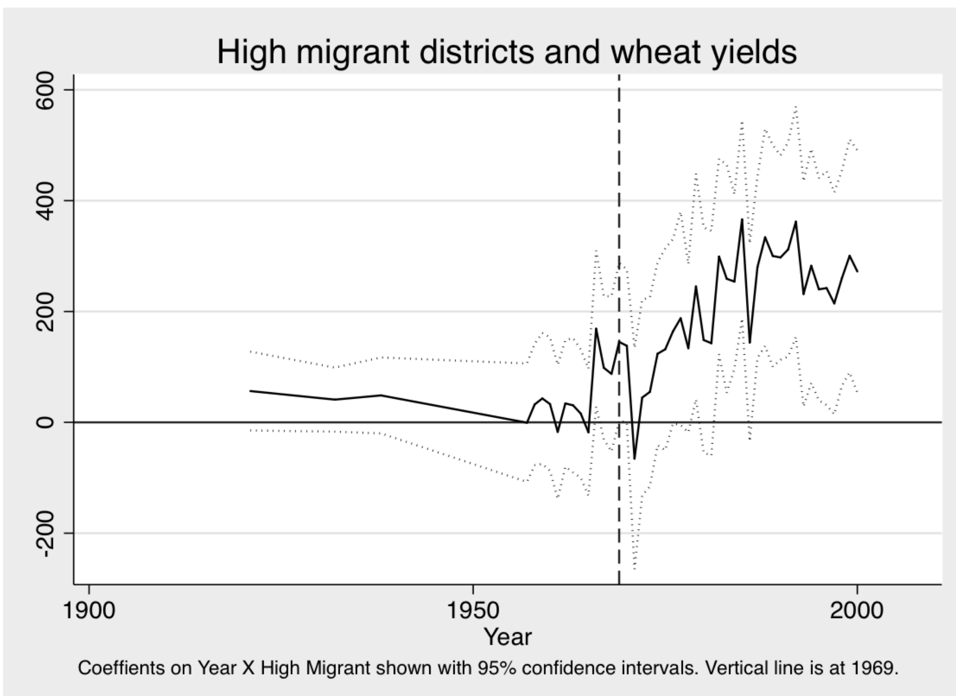
4 Results

The results from estimating equation 1 is presented in Table 2. We use two variables to capture the effects of partition - log number of migrants as well as a dummy that captures a "high migrant" district, which is defined as a district above the 75th percentile in terms of the fraction of its population in 1951 that is composed of migrants. Table 2 shows that places that received more migrants did better after partition, and more specifically, after the green revolution in India. The "post green revolution" dummy takes on a value of 1 after 1972, which is the first year when India's overall HYV adoption was greater than equal to 10%. We show robustness to these definitions in a later table. In particular Column 8 of Table 2 shows a stark result when we examine the effects by decade. We find a large and statistically significant effect during the decade of 1977-1987 (the height of the green revolution period in India) in those districts that had more migrants.

The figure below is analogous to Table 2, but more flexible in its specification. To create the figure below, we simply interact year dummies with our measure for migration and plot the resulting year and migration interaction coefficients and their associated confidence intervals; it is important to note that this specification still controls for state fixed effects so we not simply comparing one state (say the Punjab) with another (say Bihar). This figure captures the essence of the paper - that migration in 1951 seems uncorrelated with trends in yields for wheat prior to partition, and indeed even for many years *after* the partition. There is a clear "take off" however occurring in the high migration areas after the green revolution occurs in India.

Table 3 examines whether other crops responded in similarly to migrant presence. Our main argument is that migrant presence enabled the take up of better crops and technologies once the green revolution made it possible to do so. Hence, for crops *not* affected by the green revolution, we would not expect to see an increase in yields, unless migrants were somehow better at farming *all* crops. Table 3 shows that this is broadly not the

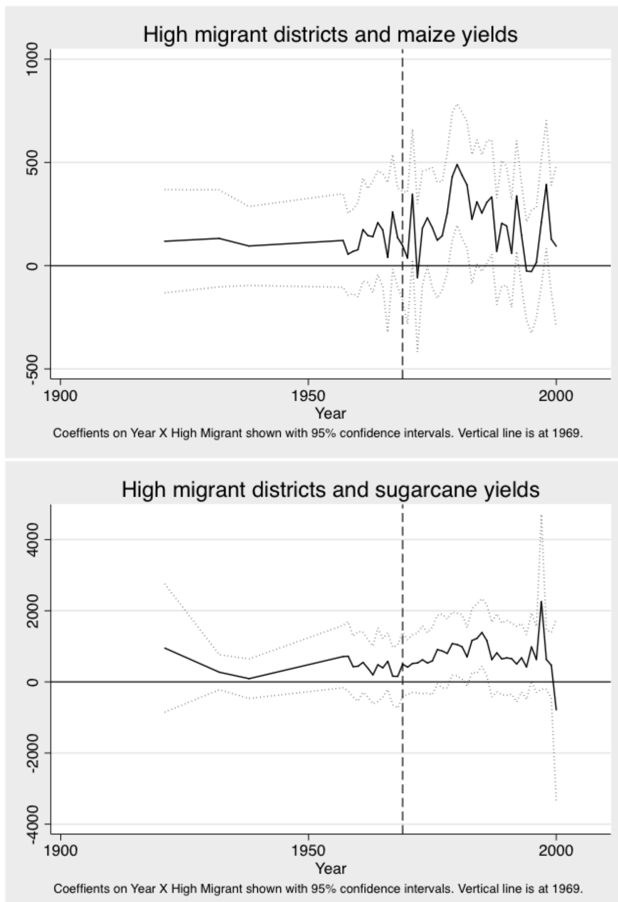
Figure 1: Migrant Presence and Wheat Yields



Notes: Each point on the graph is the interaction coefficient from a regression where year dummies are interacted with a migration dummy. The regression control for the main effects, and for state fixed effects along with controls for soil types, latitude, longitude and altitude at the district level. This regression is based on a sub sample of districts for whom comparable agricultural data was available starting in 1911 as described in the text.

case across eight other crops for which we have consistent data. Figure 2 follows the same methodology as Figure 1 and show that a) there were no pre trends in the yields of other crops prior to partition, and b) that even after partition and the advent of the green revolution, there was little change to the yields of non-green revolution crops in high migrant areas.

Figure 2: Migrant Presence and Other Crops



Notes: See Figure 1 and text for details.

The outcome variable in Table 4 is the revenue per acre using 1960 prices. As mentioned earlier, the data is in panel form and hence, we cluster the standard errors at the district level (comparable estimates from a cross section where the average over the entire period is used is presented in Appendix Table 1). In column 1 of Table 4, we estimate equation 1 with no controls for soil conditions and population density, but including controls for state and year fixed effects, as well as state specific time trends. So as to control for the size of the district, we include log of the population in 1951 in all our specifications as well. Column 1 shows that a 10% increase in migrants, is correlated with an increase in revenue by 1.4 Rupees per hectare. Given the average revenue per hectare of approximately 485 Rupees, this is a rather small increase. This average effect likely hides important heterogeneity since some districts had much larger inflows than others. In the Punjab for example, districts like Gurdaspur, Kapurthala and Jalandhar, after partition, were made up of 34, 28 and 25 percent migrants, respectively. On the other hand, districts like Mysore, Bangalore and Hassan in Karnataka all had close to 0 percent migrants.

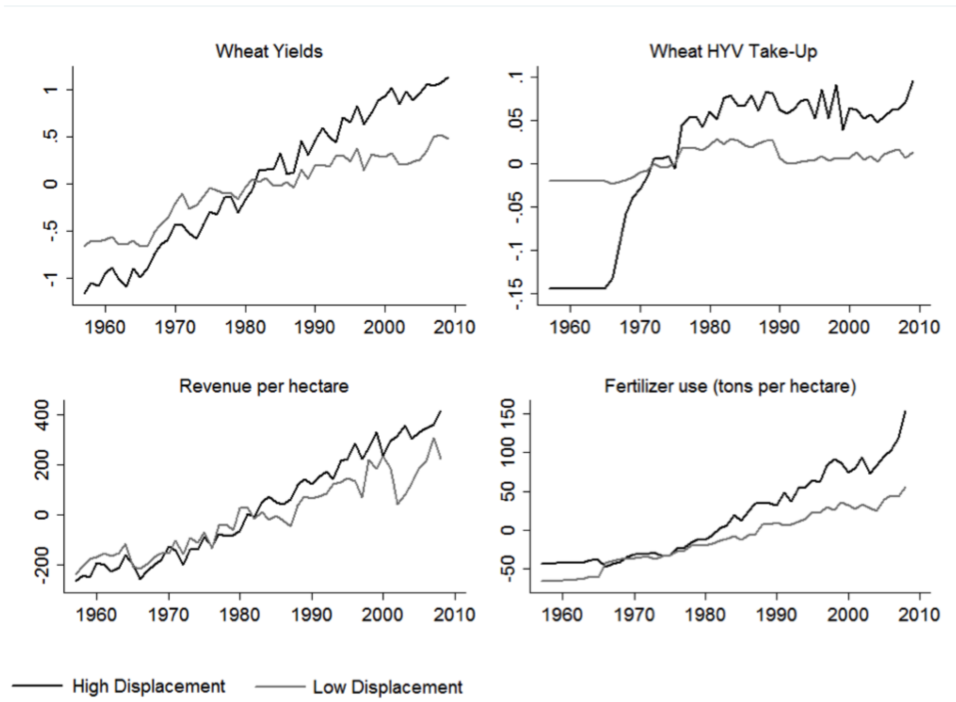
In Columns 2 and 3, we sequentially add the controls of soil quality, population density and rainfall. We do this primarily to asses whether migrant selection into districts was systematically correlated with these variables, which might also affect the outcome of interest. For example, the estimate on log migrants in Column 2 is very similar to the estimate in Column 1 (indeed, they are not statistically significantly different from one another), suggesting that migrant selection on the basis of soil quality and suitability for agriculture is not a concern in our case. Similarly, adding population density and rainfall keeps the results largely stable (although the estimates are not statistically significant in columns 2 and 3).

Table 5 examines whether the effect of migrant presence is greater after the green revolution starts in that particular district. We define the start of the green revolution as the calendar year after which 5% or more of acreage is under HYV seeds.[9] Note that we do not interpret the timing of the green revolution as exogenous. In fact, as we show in Table 6, migrant presence was correlated with the take up of HYV seeds. Instead, our preferred interpretation is that once a certain fraction of acreage is under HYV, the presence of migrants helps revenue from agriculture improve *even more*. The positive and significant coefficients in Table 5 on the interaction confirm this.

Table 6 examines wheat yields and the take up of HYV varieties of wheat as the dependent variables of interest. Both yields and the take up of HYV are significantly correlated with migrant presence, and the effects are only larger once the green revolution occurs in that district (the cross sectional results for take up of HYV are presented in Appendix Table 2). Visually, this is confirmed in Figure 1. It is important to note here that the individual figures *do* take into account state fixed effects. The figures show that high displacement

[9]The results on take up of HYV varieties of wheat are similar if we define the green revolution timing to be based on a *national* level; that is, defining green revolution start as the first year when more than 5% of crops nationally were HYV.

Figure 3: Migrant presence and agricultural outcomes net of state fixed effects



Notes: In High Displacement districts the proportion of in-migrants is either equal to or above the 75th percentile based on the full sample and in the Low Displacement districts it is either equal to or below the 25th percentile. The variable on the y-axis has been stripped of the state fixed effects. For instance in the case of the top left plot we first regressed the wheat yields on state dummies. We then predicted the residual values for wheat yields and plotted these residuals by year, distinguishing between High and Low Displacement districts. We employed the same procedure for all plots in this figure.

areas and low displacement areas within a state were quite similar until the mid-late 1960s, after which the high displacement districts see greater revenue, wheat yields, tractor use, and acreage under HYV seeds. This is broadly consistent with the timing of the green revolution (Foster and Rosenzweig, 1996). Table 6 Column 1 suggests that, compared to districts that received no migrants, districts that received migrants saw yields increase by 3.2%. As expected, this effect is stronger after the green revolution occurs in a given district. Column 3 of Table 6 suggests that districts with migrants saw an increase in HYV use of 6% compared to districts with no migrants. These results are similar when we specify the right hand side variable in terms of proportion migrants (rather than log number of migrants) as shown in Appendix Table 4a. While Columns 1 and 3 are not statistically significant, the migrant proportion interacted with the green revolution dummy is statistically significant. Appendix Table 4b shows that these results are also robust to exclusion of mismatching data across the overlapping years in the VDSA and IACD data sets.

Table 7 confirms the graphical result seen for tractor and fertilizer use in regression form - tractor use per acre is 20% higher in areas with migrants compared to areas without, and is even more so after the green revolution; nitrogen fertilizer use is 4.6% higher and phosporous fertilizer use almost 7% in districts with some migrants.

An important consideration here is whether migrant presence was correlated with characteristics of a district that made it predisposed to better agricultural development. For example, if a district had better access to waterways or better soil, and if migrants concentrated in these districts, it would be difficult to separate the forces of selective migration into districts from the effects of migrant presence and other partition related forces on agricultural outcomes. We have shown earlier in Table 4 and 5 that soil conditions do little to change the overall estimates, suggesting that migrant presence is not correlated systematically in ways with soil conditions that matter for agricultural output. Tables 2 and 3 reinforce this point that even if districts varied on the basis of agricultural suitability in other unobserved ways, they did not at least result in differential *trends* in agricultural outcomes prior to partition. In Table 8, we conduct a more direct test by regressing various characteristics of districts on migrant presence in 1951. The aim of this table is to show that migrant presence in 1951 is uncorrelated with key district level variables such as pre-partition investment in canal infrastructure, or pre-green revolution investments in banks, post office, hospitals, etc. Only two variables show any significant correlations and these are schools per capita and length of roads in the district (although this correlation is only significant at the 10% level). The correlation with schools per capita is *negative* which goes against the idea that migrants moved to places that were better off and therefore perhaps better able to take advantage of the green revolution. If anything, this correlation suggests that *despite* migrants moving to areas with fewer schools per capita, they are able to take advantage of higher yields after the green revolution.

5 Mechanisms

Our empirical analysis in the previous section has shown a positive correlation between the number of migrants and long-run agricultural development. In this section we will elucidate two channels through which such a relationship operated. Firstly, migrants were more literate than the natives of the districts in which they settled. This in turn meant that they were capable farm managers who were more likely to adopt newer agricultural technologies. Secondly, for a substantial period prior to partition the migrants had been involved in lending to small-scale farmers for agricultural purposes. Therefore, it is also likely that they influenced agricultural development through their contribution towards credit expansion.

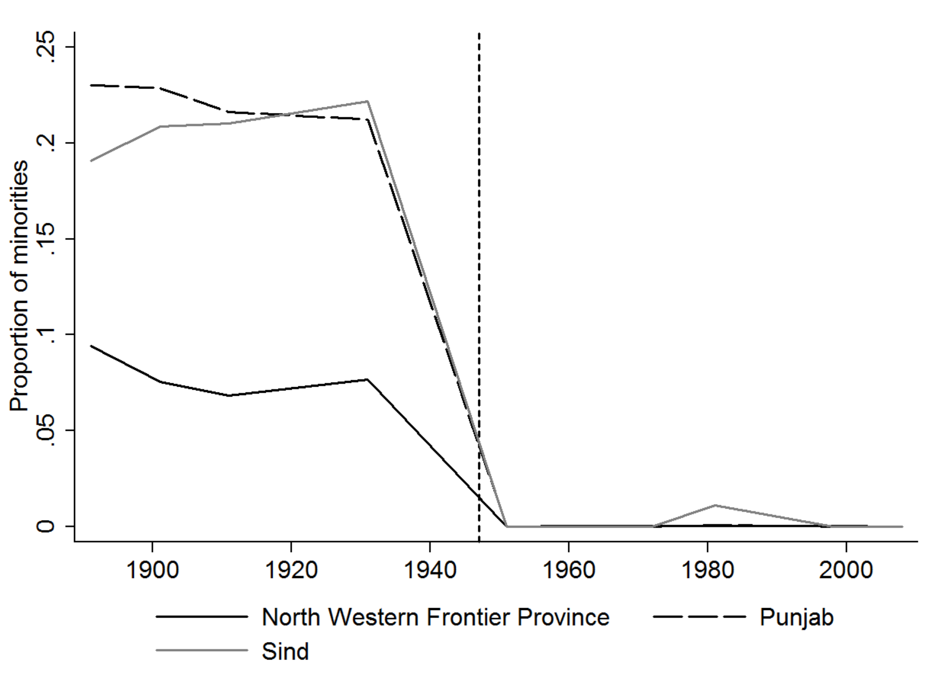
5.1 Migrants and Human Capital

The migrants who came to post-independence India were the Hindus and Sikhs who were expelled from areas that later became post-independence Pakistan[10]. Figure 2 provides a visual illustration of their expulsion from three regions of colonial India that went to post-independence Pakistan – Western Punjab, Sind and North Western Frontier Province. Between 1931, the last reliable census prior to partition, and 1951, the first census after partition, the percentage of Hindus and Sikhs in the regions drop from 20% to 0.3%. Such a sudden and universal drop is evidence of there being no selective out-migration from the districts that went to post-independence Pakistan.

An important characteristic of the Hindu and Sikh migrants who came to India was their above average literacy rates in the areas from which they were expelled. Figure 3 compares the literacy rate of Hindus and Sikhs with those of the Muslims in the districts

[10]Ideally, we would have liked to distinguish between the affect of Hindu and Sikh migrants in our paper. However, the 1951 census of India does not record the religion of the displaced migrants.

FIGURE 4: Drop in proportion of Hindus and Sikhs around partition



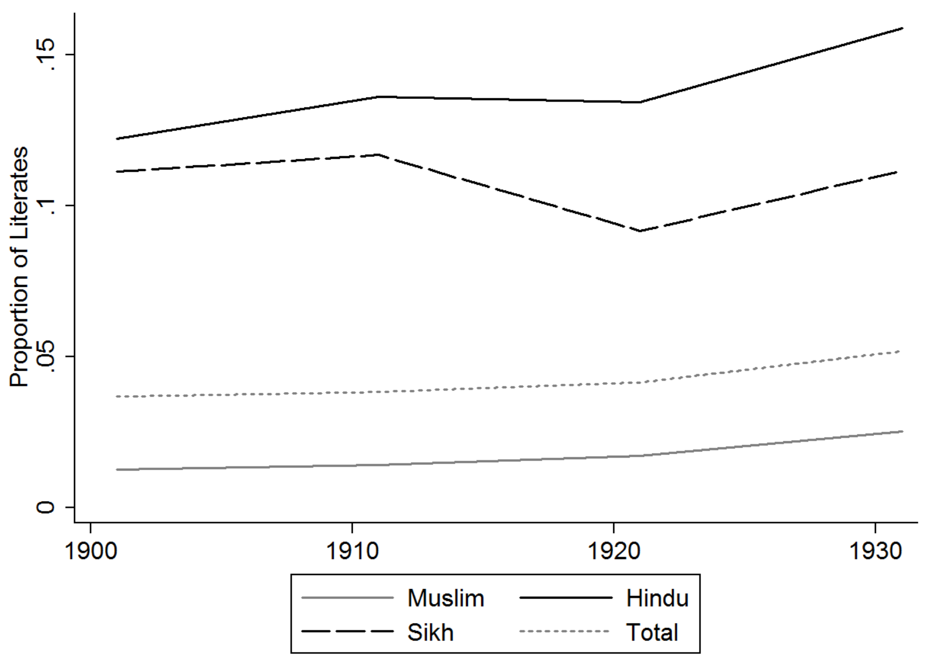
Notes: The figure excludes the 1941 census numbers that are widely regarded as being unreliable. Furthermore, most of the out migration of minorities shown in the figure took place in the brief period between 1947 and 1951.

that later became post-independence Pakistan for the period prior to partition. The stark difference between the literacy rates is quite revealing. The Hindus and Sikhs out performed the Muslims in terms of literacy throughout the four pre-partition census years of 1901, 1911, 1921 and 1931. It is therefore plausible that the migrants, at least in part, contributed towards agricultural development through their impact on literacy. Simple correlations in Appendix Table 6a show that migrant presence is indeed correlated with increased literacy of Indian districts in the years after partition. The correlation coefficient between log number of migrants in 1951 and rural male literacy in 1961 is 0.1204 and is significant at the 10% level. It increases in both magnitude and significance between 1961 and 1991.

There are several papers that correlate education with agricultural technology adoption and crop yields (Schultz, 1964; Gerhart, 1974; Jamison, Lau, et al., 1982; Rosenzweig, 1978; Ram, 1976; Sidhu, 1976). The argument usually put forward is that the adoption of agricultural technology requires the ability to perceive, interpret, and respond to new events in the context of risk, and that such ability is derived through human capital (Schultz, 1964). The underlying hypothesis of such an argument is that education increases the ability of farmers to "understand and evaluate the information on new products and processes", thereby incentivizing them to adopt new technologies (Feder, Just, and Zilberman, 1985). Rosenzweig (1978) finds that the probability of adopting high yield varieties of grain in the Indian Punjab is positively related to farmer education and farm size. Sidhu (1976) in another study on the Indian Punjab finds that the education of farmers has a positive impact on both the crop yields and gross sales revenue from the lands that were cultivated in the early stages of the Green Revolution. Finally, Ram (1976) in yet another study on India show that the contribution of farm operators to production was positively related to their education. Feder, Just, and Zilberman (1985) provide a comprehensive review of the broader literature connecting human capital with agricultural technology adoption. In our context, we show using simple correlations in Appendix Table 6b, that literacy at the district level is positively correlated to the take up of high yielding variety of seeds in the years subsequent to the partition. The correlation coefficient between take up of high yielding variety of all major crops and rural male literacy in 1971 is 0.2691, is 0.0843 in 1981 and is 0.2640 in 1991.

To substantiate our claim that the migrants influenced agricultural development through their impact on literacy we present anecdotal evidence which suggests that the migrants were literate cultivators who were known for their superior farming practices. Our evi-

Figure 5: Hindu, Sikh and Muslim literacy prior to partition in districts that went to post-independence Pakistan.



Notes: The figure is based on the three colonial regions of Western Punjab, Sind and North West Frontier Province, all of which became part of post-independence Pakistan dence comes from those districts of colonial India that later became post-independence Pakistan and from where the migrants emigrated. For instance, the Hindu Jats[11] of the Lyallpur district were considered by the colonial authorities as being the "most useful class of peasants"[12]. The Hindu and Sikh Jats of the Sialkot district were deemed to be far superior cultivators than their Muslim counterparts[13]. The gazetteer of the Lahore district notes that the Hindu and Sikh Jats were "good husbandsmen"[14]. The Sikh Virakhs[15] of the Montgomery district were considered first-rate cultivators[16]. Most emphatically, the (1881) census of the Punjab states that a substantial proportion of the Sikh Jats belonging to the Lahore and Gujranwala districts were "stalwart, sturdy yeomen of great independence, industry, and agricultural skill" who collectively formed "perhaps the finest peasantry in India"[17].

Official colonial documents also acknowledge the superior position held by the migrants in terms of education in the districts from which they came from. For instance, literacy was highest "among Hindus and Sikhs, among the non-Christian population" of the Attock district[18]. In the Lahore district the pre-eminence of the Hindus in education was deemed "remarkable" and the considerable progress that had been made in "education of Sikh males" recognized[19]. Interestingly, the 1929 Muzaffargarh district gazetteer went so far as to suggest that "no special measures were necessary in the case of Hindus and Sikhs" as they were "ready to take advantage of every opportunity" of providing education to their children[20]. A more systematic record of statements contrasting the pre-partition literacy rate of Hindus and Sikhs with those of the Muslims in the districts that went to post-independence Pakistan is given in Appendix Table A7. Other sources, outside of the official colonial publications, also point to the contribution the migrants had made to education. Raychaudhuri, Habib, and Kumar (1983) when discussing the aftermath of partition in post-independence Pakistan observe that the event led to the sudden departure of teachers and instructors who mainly came from the Hindu and

[11]An agricultural caste of the Punjab
[12]Gazetteers, Punjab District. Gazetteer of the Chenab Colony, 1904. Vol.A. Page 51
[13]Gazetteers, Punjab District. Gazetteer of the Sialkot District, 1893-94. Page 75
[14]Gazetteers, Punjab District. Gazetteer of the Lahore District, 1883-84. Page 65
[15]An agricultural caste of the Punjab
[16]Gazetteers, Punjab District. "Gazetteer of the Montgomery District, 1898-99. Page 86
[17]Report on the Census of the Panjab Taken on the 17th of February 1881. Page 229
[18]Gazetteers, Punjab District. Gazetteer of the Attock District, 1907. Page 304
[19]Gazetteers, Punjab District. Gazetteer of the Lahore District, 1893-94. Page 84
[20]Gazetteers, Punjab District. Gazetteer of the Muzaffargarh District, 1929. Page 291

Sikh communities[21]. The First Five Year Plan of the Planning Commission of Pakistan acknowledges the damage done to the educational sector by the "sudden departure of Hindu teachers and instructors" who had manned the staff of the technical institutions, schools, colleges and universities in the country[22]. The Hartog (1929) committee report that reviewed the growth of education in late colonial India notes that in the Western Punjab and the North Western Frontier Province–both regions that later became part of post-independence Pakistan–the Hindus and Sikhs had done "good service to the cause of education by the maintenance of a large number of schools and colleges"[23].

In line with the evidence presented above, we posit that at least part of the impact of migrants on agricultural outcomes that we document statistically is mediated through their impact on literacy. Literacy, however, is not the only dimension of human capital along which the migrants could have contributed. We, therefore, consider occupation as another dimension of human capital through which the migrants could have influenced agricultural development. It is well documented that lack of credit is a constraint farmers in developing countries face in adopting new technologies (Bhalla, 1979; Pitt and Sumodiningrat, 1991; Lipton, 1976). Often, introducing a high yielding variety of a crop or purchasing a tractor requires having access to loans because farmers simply do not have adequate savings to make such investments on their own. Access to credit then acts as a supplement to savings that can be used to invest in technology. The provision of credit also reduces the risks farmers face in their lives as it cushions them from extreme fluctuations in agricultural output. The reduction in risk in turn makes them more likely to adopt newer, more riskier, technologies.

From the anecdotal evidence we have gathered we know that large proportions of the migrants were involved in small-scale money lending to farmers for agricultural purposes in the districts from which they emigrated. They provided a "much needed source of credit for cultivation"(Raychaudhuri, Habib, and Kumar, 1983) for local farmers who would otherwise not have had access to formal credit markets. Most of them belonged to the three great Hindu and Sikh mercantile castes of India–Khatris, Aroras and Baniahs–that dominated commercial activity. Figure 3 provides a snapshot of the predominance

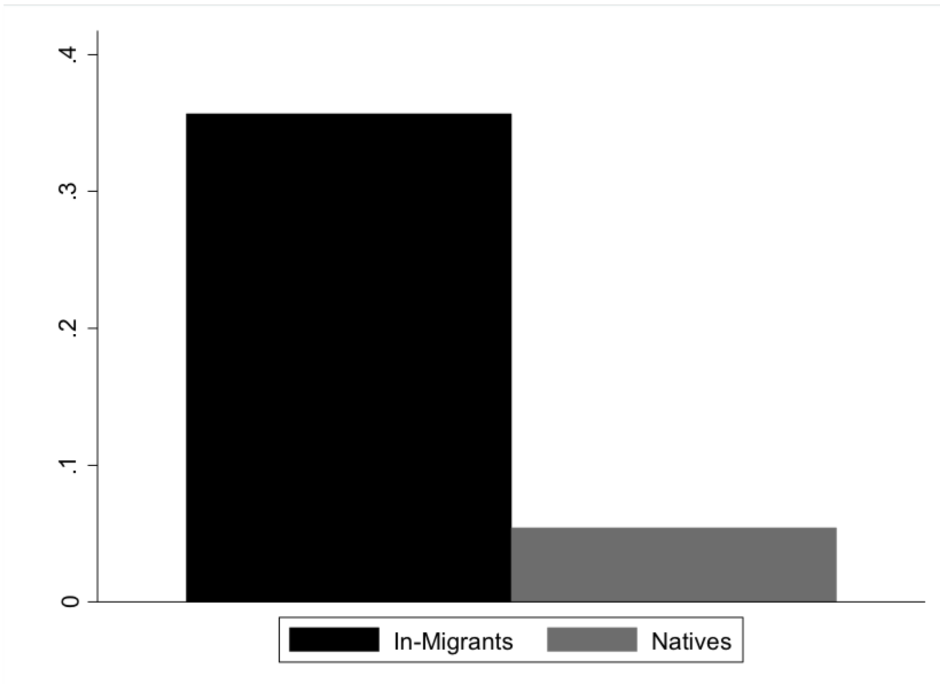
[21]Raychaudhuri, Tapan, Irfan Habib, and Dharma Kumar, eds. The Cambridge economic history of India. Vol. 2. CUP Archive, 1983. Page 998

[22]Planning Commission. Government of Pakistan. The First Five Year Plan 1955-60. (1957). Page 7

[23]Hartog, P.J., 1929. Interim Report of the Indian Statutory Commission: Review of Growth of Education in British India by the Auxiliary Committee Appointed by the Commission . Vol. 3407. HM Stationery Office. Page 246

of the migrants in commercial occupations at the time of partition. It compares the proportion of migrants engaged in commerce against the same proportion for the natives based on actual data on both groups from the 1951 census of India. Again, the stark contrast between the two groups in terms of their involvement in commerce is clearly apparent.

FIGURE 6: Proportion in the commercial sector at partition.



Notes: The bar for migrants is the proportion of the displaced persons in 1951 that were previously engaged in commerce. This data is given in Appendix II of Table IV of the 1951 census of India. The bar for natives is the proportion of the non-displaced persons that were previously engaged in commerce. This data is also available in the 1951 census of India.

As was the case with their educational superiority the higher concentration of the migrants in commercial occupations was also noted in official publications dating from the colonial period. The notes, again, pertain to the Hindu and Sikh communities in areas that later became post-independence Pakistan. For instance, the (1881) census of Punjab states that the Hindus and Sikhs were mostly traders.[24] Hindus from the Arora caste controlled "almost the whole of the trade, moneylending, and banking" in the Muzaffargarh district.[25] The Hindu Aroras were also considered as being the "chief moneylenders

[24]Report on the Census of the Panjab Taken on the 17th of February 1881. Page 125-138.
[25]Gazetteers, Punjab District. Gazetteer of the Muzaffargarh District, 1929. Page 78.

and capitalists" and the "chief creditors of the agriculturists" in the Jhang district.[26] Yet again, the Hindus and Sikhs from the Arora caste were identified as being the main moneylenders in the Montgomery district[27]. In the Attock district "almost the whole trade and money-lending business" was divided by the the three most numerically important Hindu castes amongst themselves[28].

In line with the above evidence we conclude that in addition to literacy the other channel through which the migrants influenced agricultural development was commercial occupations.

6 Conclusion

In this study, we examine the impact of partition on agricultural productivity and the take up of agricultural technology post-partition. Using migrant presence as a proxy for the intensity of displacement, we find that areas with more migrants have higher average yields, are more likely to take up High Yielding Varieties (HYV) of seeds, and are more likely to use agricultural technologies within the first 60 years after partition in India. We further show, using pre-partition agricultural data, that the effects are not solely explained by selective migration into districts with a higher potential for agricultural development. We then argue that the greater levels of education of the migrants and their higher concentration in commerce relative to both the natives who stayed and the migrants who moved contributed to agricultural development post partition.

While our work highlights important correlations in this area, it should not be interpreted as the causal effect of partition induced migration. The main reason for this is that the partition simultaneously resulted in many changes, migration being just one component. Hence, isolating the effect of migration alone is a rather impossible task. Despite these caveats, we believe this paper makes an important contribution towards understanding the long run trajectory of places affected by the partition in India. More studies are needed in this area as partitions or wars accompanied by mass human movements are still very much a part of the current global political environment, and understanding their lasting impacts on growth and economic development will be crucial.

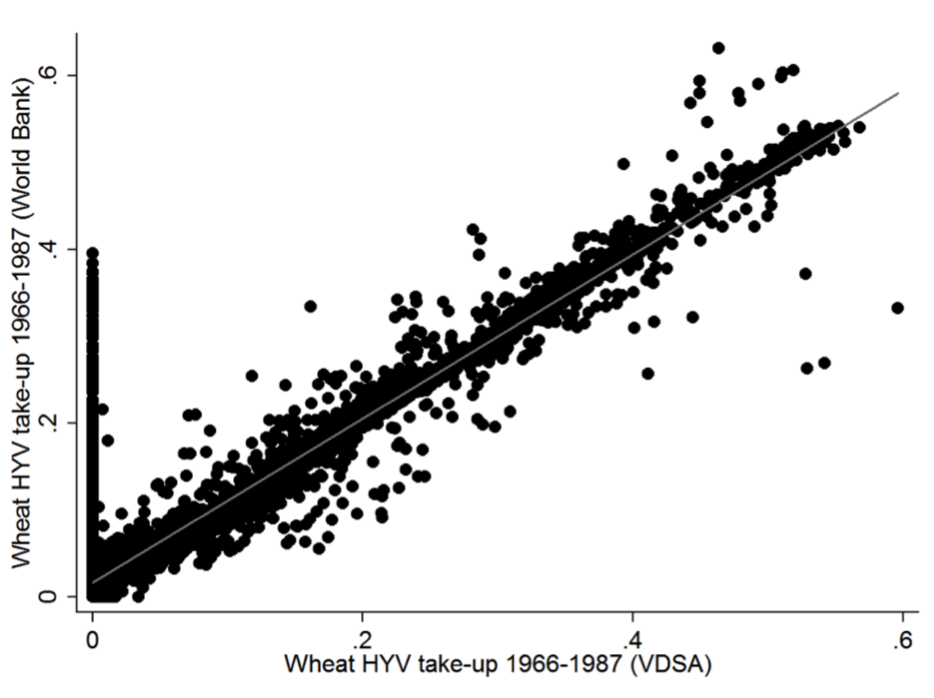
[26] Gazetteers, Punjab District. Gazetteer of the Jhang District, 1883-84. (1884). Page 68.
[27] Gazetteers, Punjab District. Gazetteer of the Montgomery District, 1883-84. (1884). Page 69-70
[28] Gazetteers, Punjab District. Gazetteer of the Attock District, 1930. Page 115.

A Appendix

Figure A.1: IACD to VDSA Wheat Yields Comparison (1966-87).

Notes: There were 19 cases in which for the same year and district the annual wheat yield was zero in the World Bank dataset but was non-zero in the VDSA dataset. 68% of these cases came from the Andhra Pradesh state and 32% came from the Karnataka state. On the other hand there were 5 cases in which for the same year and district the annual wheat yield was zero in the VDSA dataset but was non-zero in the World Bank dataset. 80% of these cases came from the Karnataka state and 20% came from the Maharashtra state.

Figure A.2: IACD to VDSA Wheat HYV Take Up Comparison (1966-87).



Notes: There were 22 cases in which for the same year and district the annual fraction of HYV of wheat was zero in the World Bank dataset but was non-zero in the VDSA dataset. 82% of these cases came from the Maharashtra state, 9% came from the Gujarat state, 4.5% from Rajasthan state and 4.5% from Tamil Nadu state. On the other hand there were 475 cases in which for the same year and district the annual fraction of HYV of wheat was zero in the VDSA dataset but was non-zero in the World Bank dataset. 98.3% of these cases came from the Uttar Pradesh state, 0.9% came from the Andhra Pradesh state, 0.2% came from the Gujarat state, 0.4% came from the Madhya Pradesh state and 0.2% came from the Orissa state.

References

Acemoglu, D., T. A. Hassan, and J. A. Robinson (2011): "Social Structure and Development: A Legacy of the Holocaust in Russia," *The Quarterly journal of economics*, 126(2), 895–946.

Acemoglu, D., S. Johnson, and J. A. Robinson (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *The Quarterly Journal of Economics*, 117(4), 1231–1294.

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

Bhalla, S. S. (1979): "Farm and Technical Change in Indian Agriculture," .

Bharadwaj, P., and J. Fenske (2012): "Partition, migration, and jute cultivation in India," *Journal of Development Studies*, p. 1.

Bharadwaj, P., A. I. Khwaja, and A. R. Mian (2009): "The partition of India: demographic consequences," *Available at SSRN 1294846*.

Chaney, E., and R. Hornbeck (2015): "Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos," *The Economic Journal*.

Chaudhary, L., and J. Rubin (2011): "Reading, writing, and religion: Institutions and human capital formation," *Journal of Comparative Economics*, 39(1), 17–33.

Christopher, A. J. (2011): "Secession and South Sudan: an African precedent for the future?," *South African Geographical Journal*, 93(2), 125–132.

Dell, M. (2010): "The persistent effects of Peru's mining mita," *Econometrica*, 78(6), 1863–1903.

Dippel, C. (2014): "Forced coexistence and economic development: Evidence from native american reservations," *Econometrica*, 82(6), 2131–2165.

Donaldson, D. (2010): "Railroads of the Raj: Estimating the impact of transportation infrastructure," Discussion paper, National Bureau of Economic Research.

Feder, G., R. E. Just, and D. Zilberman (1985): "Adoption of agricultural innovations in developing countries: A survey," *Economic development and cultural change*, pp. 255–298.

Foster, A. D., and M. R. Rosenzweig (1996): "Technical change and human-capital returns and investments: evidence from the green revolution," *The American economic review*, pp. 931–953.

Gerhart, J. D. (1974): "The diffusion of hybrid maize in western Kenya," *PhD Dissertation, Princeton University.*

Hartog, P. J. (1929): *... Interim Report of the Indian Statutory Commission:(Review of Growth of Education in British India by the Auxiliary Committee Appointed by the Commission)...*, vol. 3407. HM Stationery Office.

Heston, A. (1978): "A Further Critique of Historical Yields Per Acre in India," *Indian Economic & Social History Review*, 15(2), 187–210.

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

Hornbeck, R., and S. Naidu (2014): "When the levee breaks: black migration and economic development in the American South," *The American Economic Review*, 104(3), 963–990.

Iyer, L. (2010): "Direct versus indirect colonial rule in India: Long-term consequences," *The Review of Economics and Statistics*, 92(4), 693–713.

Jamison, D. T., L. J. Lau, et al. (1982): *Farmer education and farm efficiency.* Johns Hopkins University Press.

Jha, S. (2013): "Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia," *American Political Science Review*, 107(04), 806–832.

Jha, S., and S. Wilkinson (2012): "Does combat experience foster organizational skill? Evidence from ethnic cleansing during the partition of South Asia," *American Political Science Review*, 106(04), 883–907.

Kliot, N., and Y. Mansfield (1997): "The political landscape of partition: the case of Cyprus," *Political geography*, 16(6), 495–521.

KUMAR, R. (2004): "Settling Partition Hostilities: Lessons Learned, Options Ahead," *The Fate of the Nation-state*, p. 247.

LIPTON, M. (1976): "Agricultural finance and rural credit in poor countries," *World Development*, 4(7), 543–553.

NUNN, N. (2008): "The Long-Term Effects of Africa's Slave Trades," *The Quarterly Journal of Economics*, 123(1), 139–176.

——— (2009): "The importance of history for economic development," Discussion paper, National Bureau of Economic Research.

PITT, M. M., AND G. SUMODININGRAT (1991): "Risk, schooling and the choice of seed technology in developing countries: a meta-profit function approach," *International Economic Review*, pp. 457–473.

RAM, R. (1976): "Education as a quasi factor of production: the case of India's agriculture," Ph.D. thesis, University of Chicago, Department of Education.

RAYCHAUDHURI, T., I. HABIB, AND D. KUMAR (1983): *The Cambridge Economic History of India: Volume 2, C. 1751-c. 1970*, vol. 2. CUP Archive.

ROSENZWEIG, M. R. (1978): "Schooling, allocative ability and the green revolution," *Paper presented at the meetings of the Eastern Economic Association, Washington, D.C.*

SCHULTZ, T. W. (1964): *Transforming Traditional Agriculture*. Yale University Press, New Haven, CT.

SIDHU, S. S. (1976): "The Productive Value of Education in Agricultural Development," Discussion paper, University of Minnesota, Department of Applied Economics.

VISARIA, P. M. (1969): "Migration between India and Pakistan, 1951–61," *Demography*, 6(3), 323–334.

Table 1A. Summary Statistics

Variable	Mean	Median	Standard Deviation	Number of Observations
Annual Wheat Yields (1957-2009) (tons per hectare)	1.489351	1.291846	0.8928221	12763
Annual Take-Up of HYVs of Wheat (1957-2009) (proportion)	0.0702903	0.0048164	0.1275124	11573
Annual Revenue per hectare based on 1960 prices (rupees per hectare)	493.3906	386.9451	356.9391	11500
Annual consumption of Nitrogen Fertilizer (tons)	40.24067	16.26249	118.8009	13001
Annual consumption of Phosphorus Fertilizer (tons)	15.81228	5.905565	50.42723	13001
Annual consumption of Potassium Fertilizer (tons)	10.2566	1.257571	99.38923	13001
Annual number of tractors per 1000 hectares (1957-1987)	1.811449	0.3952479	4.89581	8370
Log Migrants (1951)	7.443211	7.548975	2.576294	270
Log of Population (1951)	13.79153	13.84592	0.638946	270
Annual Rainfall (1957-2009) (millimeters)	1039.807	945	573.9935	13971
Population Desnity (1961)	1.789214	1.250549	2.117194	270

Table 1B: Summary Statistics comparison between full and restricted sample

	Full Sample (1957-2009)					Restricted Sample (1911-2009)						
Variable	Actual	Mean	Median	Std. Dev	Observations	Actual	Proportion of Full Sample	Mean	Median	Std. Dev	Observations	Proportion of Full Sample
Number of Migrants (1951)	5862319	…	…	…	265	3329178	**57%**	…	…	…	67	25%
Population (1951)	302980160	…	…	…	265	94666560	**31%**	…	…	…	67	25%
Annual Wheat Yields (lbs per acre)	…	1328.77	1152.56	796.56	12763	…	…	1341.08	1106.13	838.29	3870	30%
Annual Rice Yields (lbs per acre)	…	1213.27	1009.41	736.97	13396	…	…	894.47	378.50	1468.52	3427	26%
Annual Maize Yields (lbs per acre)	…	1190.27	973.29	817.38	12758	…	…	1175.99	1049.76	656.69	3508	27%
Annual Jowar Yields (lbs per acre)	…	631.79	590.52	388.95	11840	…	…	652.58	632.14	386.15	2950	25%
Annual Bajra Yields (lbs per acre)	…	653.52	575.19	418.15	8514	…	…	658.07	580.00	416.91	2395	28%
Annual Rainfall (millimeters)	…	1040.65	946.00	574.73	13926	…	…	1013.19	949.30	437.26	3470	25%
Altitude (meters)	…	353.38	357.00	140.20	13878	…	…	424.72	396.00	212.03	3939	28%
Latitude (degrees)	…	22.78	23.45	4.85	13878	…	…	24.61	24.96	3.92	3939	28%
Longitude (degrees)	…	78.82	78.11	4.36	13878	…	…	79.43	78.43	4.57	3939	28%

Notes: The unit of observation is a district and year. Data on annual rainfall does not change between 1957 and 1965 and is only available from 1957 onwards. Restricted sample is the sample of 67 districts for which data is consistently available from 1911 to 2009. Full sample is the sample of districts in the dataset from 1957 onwards.

Table 2: Wheat Yields and Migration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Migrants				High Migrant dummy			
Post Partition X Var	32.3	32.24			150.1**	150.0**		
	(21.82)	(23.33)			(60.99)	(65.21)		
Post GR X Var			53.10**				196.3***	
			(25.69)				(72.67)	
1957-1967 X Var				-6.926				2.094
				(15.78)				(44.99)
1967-1977 X Var				8.937				67.05
				(22.20)				(57.10)
1977-1987 X Var				51.11				183.5**
				(31.00)				(81.00)
1987-1997 X Var				63.16*				251.4***
				(33.19)				(93.12)
1997-2009 X Var				42.26				224.1**
				(34.76)				(106.0)
Controls	State linear trends	State X Year FE			State linear trends	State X Year FE		
Observations	3,785	3,785	3,785	3,785	3,785	3,785	3,785	3,785
R-squared	0.893	0.925	0.926	0.926	0.914	0.946	0.948	0.927

Robust standard errors in parentheses, clustered at the district level
*** p<0.01, ** p<0.05, * p<0.1

Controls: In addition to the controls stated on the table, all regressions control for state fixed effects, district soil, latitude, longitude and altitude, and population of district in 1961 as controls.

Table 3: Migrant presence and other crops

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sugarcane	Maize	Sorghum	Other pulses	Other millets	Chickpea	Rapeseed	Linseed
Post Partition X High Median Migrant dummy	363.2**	55.96	-93.09	21.37	-124.8	49.80	173.0*	1.915
	(165.3)	(121.9)	(85.85)	(63.30)	(102.6)	(51.67)	(103.6)	(18.45)
Observations	3,656	3,441	2,918	2,367	1,130	3,108	2,948	1,786
R-squared	0.533	0.493	0.326	0.418	0.190	0.055	0.524	0.479

Robust standard errors in parentheses, clustered at the district level
*** p<0.01, ** p<0.05, * p<0.1
Controls: State linear trends, state fixed effects, and controls for soil type, latitude and longitude included in all regressions.

Table 4. Annual Revenue Per Hectare based on 1960 prices and Log Migrants at Partition

	Annual revenue (in Rupees) per hectare based on 1960 prices		
	(1)	(2)	(3)
Log Migrants (1951)	14.14733*	12.438	10.34717
	(7.633068)	(8.528403)	(8.046685)
Mean Outcome	485.1667	484.1981	477.7716
No of observations	11149	11119	10965
Controls			
Soil type Dummies	No	Yes	Yes
Population Density (1961)	No	No	Yes
Annual Rainfall	No	No	Yes

Notes: This table shows regressions of annual revenue per hectare based on 1960 crop prices for the period 1957 to 2009 on the log number of Migrants in 1951. The unit in which revenue is measured is rupees. All the above regresssions include the log of population in 1951, state fixed effects, year fixed effects and state-specific time trends. Additionally, Column 2 includes 21 dummies which caputre soil type and Column 3 includes both the soil 21 type dummies, annual rainfall and population density in earliest available year which is 1961.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Table 5. Annual Revenue Per Hectare based on 1960 prices and Interaction between Log Migrants at Partition & Post Green Revolution Dummy

	Annual revenue (in Rupees) per hectare based on 1960 prices		
	(1)	(2)	(3)
Log Migrants (1951)	9.734309	8.217483	6.462203
	(6.471033)	(7.520569)	(7.129287)
Log Migrants (1951) X Green Revolution	7.208264*	6.968991*	6.42985*
	(3.835457)	(3.733785)	(3.637053)
Mean Outcome	485.1667	484.1981	477.7716
No of observations	11149	11119	10965
Controls			
Soil Type Dummies	No	Yes	Yes
Population Density (1961)	No	No	Yes
Annual Rainfall	No	No	Yes

Notes: This table shows regressions of annual revenue per hectare based on 1960 crop prices for the period 1957 to 2009 on the interaction between the post green revolution dummy and log number of Migrants in 1951 and the log number of Migrants in 1951. The unit in which revenue is measured is rupees. All the above regressions include the interaction between the post green revolution dummy (defined as taking a value 1 on or after the first year in which the proportion of HYV area for all major crops exceeds 5% and 0 otherwise) and the log of population in 1951, the post green revolution dummy, state fixed effects, year fixed effects and state-specific time trends. Additionally, Column 2 includes 21 dummies which caputre soil type and Column 3 includes both the soil 21 type dummies, annual rainfall and population density in earliest available year which is 1961.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Table 6. Annual Wheat Yields, HYV take up and Log Migrants at Partition

	Annual yields		Take up of HYV Variety	
	(1)	(2)	(3)	(4)
Log Migrants (1951)	0.0475933***	0.0199037**	0.0041836***	-0.0022936*
	(0.0116397)	(0.0098573)	(0.0013088)	(0.0013498)
Log Migrants (1951) X Green Revolution		0.0477704***		0.0128356***
		(0.0090729)		(0.0022229)
Mean Outcome	1.458703	1.458703	0.0707095	0.0707095
No of observations	12015	12015	10969	10969
Controls				
Soil Type Dummies	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: Column 1 of this table shows the regression of annual wheat yields for the period 1957 to 2009 on log number of Migrants in 1951. Column 2 adds the interaction between the post green revolution dummy (defined as taking a value 1 on or after the first year in which the proportion of HYV area for all major crops exceeds 5% and 0 otherwise) and the log number of Mgrants in 1951, the interaction of log of population in 1951 and post green revolution dummy and the post green revolution dummy to the regression in Column 1. Columns 3 and 4 replace annual wheat yields with annual rice yields as the dependent variable. The unit in which yields are measured is tons per hectare. All the above regressions include the log of population in 1951, state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type and population density in earliest available year which is 1961. * p < .10, ** p < .05, *** p < .01. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Table 7. Annual Use of Tractors and Log Migrants at Partition

	Tractors per 1000 hectares of land		Consumption of Nitrogen fertilizer per hectare		Consumption of Phosphorus fertilizer per hectare	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Migrants (1951)	0.4165486***	0.2774163***	1.533318**	0.1129429	0.9231105***	0.2599436
	(0.0861187)	(0.0615973)	(0.6624695)	(0.4795303)	(0.3040117)	(0.231008)
Log Migrants (1951) X Green Revolution		0.4078243***		2.555028***		1.216369***
		(0.1243062)		(0.6952638)		(0.365296)
Mean Outcome	2.012932	2.012932	33.14324	33.14324	33.14324	33.14324
No of observations	8308	8308	12231	12231	12231	12231
Controls						
Soil Type Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The annual use of tractors is measured as the number of tractors per thousand hecatares of the land that is planted in a given year. Tractor data is only avaialble in the IACD dataset, and hence available from 1958-1987. Data on fertilizers is in both the VDSA and IACD and is available from 1957-2009. The annual consumption of Nitrogen and Phosphorus fertilizers is measured in tons per hectare. All the above regressions include log of population in 1951, state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type and population density in earliest available year which is 1961.* p < .10, ** p < .05, *** p < .01. Clustered standard errors in parentheses.

Table 8: Migrants and pre partition/green revolution infrastructure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Area under canal irrigation	Area sown with canal irrigation	Hospitals per capita in 1961	Schools per capita in 1961	Post offices per capita in 1961	Banks per capita in 1961	Length of roads in district	Total geographic area of district	Growth in literate population 1911-1921
High_Median	-1.848	-0.0114	3.17e-06	-0.000147**	-1.34e-05	-5.47e-07	534.3*	-50.81	0.0319
	(2.325)	(0.180)	(4.64e-06)	(5.80e-05)	(3.10e-05)	(1.31e-06)	(308.3)	(93.50)	(0.0689)
Observations	197	197	61	62	45	60	65	67	67
R-squared	0.598	0.585	0.758	0.830	0.793	0.928	0.908	0.867	0.588

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Controls: State fixed effects and district area in all regressions (except column 8). Also included are controls for district level soil characteristics, latitude, longitude and altitude.

Appendix Table 1. Average Annual Revenue Per Hectare based on 1960 prices (1957 to 2009) and Log Migrants at Partition

	Average Annual Revenue Per Hectare Based on 1960 prices (1957 to 2009)		
	(1)	(2)	(3)
Log Migrants (1951)	18.82721**	17.64525**	17.2157*
	(7.77364)	(8.95059)	(8.951324)
Mean Outcome	497.1291	497.1291	498.1291
No of observations	264	264	264
Controls			
Soil Type Dummies	No	Yes	Yes
Population Density (1961)	No	No	Yes

Notes: This table shows regressions of annual revenue per hectare (based on 1960 prices) from all crops averaged over the period 1957 to 2009 on the log number of Migrants in 1951. The unit in which revenue is measured is rupees. All the above regressions include state fixed effects. Additionally, Column 2 includes 21 dummies which caputre soil type and Column 3 includes both the 21 soil type dummies and population density in earliest available year which is 1961.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 2. Average Annual Take-Up of HYVs of Wheat (1957 to 2009) and Log Migrants at Partition

	Average Annual Take-Up of HYVs of Wheat (1957 to 2009)		
	(1)	(2)	(3)
Log Migrants (1951)	0.0032553**	0.0040274***	0.0040405***
	(0.0013261)	(0.0013328)	(0.0013405)
Mean Outcome	0.0701439	1.0701439	2.0701439
No of observations	270	270	270
Controls			
Soil Type Dummies	No	Yes	Yes
Population Density (1961)	No	No	Yes

Notes: This table shows regressions of annual take-up of High Yielding Varieties (HYVs) of wheat averaged over the period 1957 to 2009 on the log number of Migrants in 1951. The annual take-up of HYV of wheat is measured as the proprotion of total area planted to all crops in a given year that is devoted to HYVs of wheat. All the above regressions include state fixed effects. Additionally, Column 2 includes 21 dummies which caputre soil type and Column 3 includes both the 21 soil type dummies and population density in earliest available year which is 1961.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 3. Average Annual Yields for Rice, Sugar and Maize (1957 to 2009) and Log Migrants at Partition

	Average Annual Yields (1957 to 2009)			
	Rice	Sugar	Maize	Wheat
Log Migrants (1951)	0.0237458	0.0912042***	0.0319746**	0.042932***
	(0.0145218)	(0.0311379)	(0.0124649)	(0.0121763)
Mean Outcome	1.330527	5.180279	1.334599	1.417064
No of observations	261	259	262	260
Controls				
Soil Type Dummies	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: This table shows regressions of annual yields for three major crops (rice, maize and sugar) averaged over the period 1957 to 2009 on the log number of Migrants in 1951. The unit in which yields are measured is tons per hectare. All the above regressions include state fixed effects, 21 dummies which caputre soil type and population density in earliest available year which is 1961.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 4a: Annual Yields and Proportion Migrants at Partition

	Annual Wheat Yields		Annual Rice Yields	
	(1)	(2)	(3)	(4)
Proportion Migrants (1951)	0.3513119	-0.2430415	0.4345647	-0.5232595
	(0.3081902)	(0.2507441)	(0.5423985)	(0.4647544)
Proportion Migrants (1951) X Green Revolution		0.8718263***		1.389071***
		(0.2539715)		(0.4329241)
Mean Outcome	1.458703	1.458703	1.31235	1.31235
No of observations	12015	12015	12365	12365
Controls				
Soil Type Dummies	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: Column 1 of this table shows the regression of annual wheat yields for the period 1957 to 2009 on Proportion Migrants in 1951. Column 2 adds the interaction between the green revolution dummy (defined as taking a value 1 on or after the first year in which the proportion of HYV area for all major crops exceeds 5% and 0 otherwise) and Proportion Migrants in 1951 and the green revolution dummy to the regression in Column 1. Columns 3 and 4 replace annual wheat yields with annual rice yields as the dependent variable. The unit in which yields are measured is tons per hectare. All the above regressions include state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type, annual rainfall and population density in earliest available year which is 1961. * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses.

Appendix Table 4b. Annual Wheat Yields, HYV take up and Log Migrants at Partition - accounting for overlapping data

	Annual yields		Take up of HYV Variety	
	(1)	(2)	(3)	(4)
Log Migrants (1951)	0.0479619***	0.0204548**	0.0042907***	-0.0021846
	(0.0116389)	(0.0099087)	(0.0013724)	(0.0014077)
Log Migrants (1951) X Green Revolution		0.0474003***		0.0122842***
		(0.0092344)		(0.0022082)
Mean Outcome	1.461622	1.461622	0.0739376	0.0739376
No of observations	11991	11991	10490	10490
Controls				
Soil Type Dummies	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: Column 1 of this table shows the regression of annual wheat yields for the period 1957 to 2009 on log number of Migrants in 1951. Column 2 adds the interaction between the post green revolution dummy (defined as taking a value 1 on or after the first year in which the proportion of HYV area for all major crops exceeds 5% and 0 otherwise) and the log number of Migrants in 1951, the interaction of log of population in 1951 and post green revolution dummy and the post green revolution dummy to the regression in Column 1. Columns 3 and 4 replace annual wheat yields with annual rice yields as the dependent variable. The unit in which yields are measured is tons per hectare. All the above regressions include the log of population in 1951, state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type, annual rainfall and population density in earliest available year which is 1961. * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009). Additionally for those years that overlap between the IACD and VDSA (1966 to 1987) it excludes observations where the dependent variable was either zero in IACD and non-zero in VDSA or non-zero in IACD and zero in VDSA.

Appendix Table 5a. Annual Wheat Yields, HYV take up and Log Migrants at Partition

	Annual wheat yields		Take up of HYV Variety	
	(1)	(2)	(3)	(4)
Log Migrants (1951)	0.0617377**	0.0374634	0.0068715**	-0.0042694
	(0.0296534)	(0.0270722)	(0.002891)	(0.0037026)
Log Migrants (1951) X Green Revolution		0.0377477**		0.0219015***
		(0.016656)		(0.0057222)
Mean Outcome	1.530777	1.530777	0.1210586	0.1210586
No of observations	3342	3342	2703	2703
Controls				
Soil type Dummies	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: The yields regressions in this table are based on the sample of 67 districts for which there was data available on wheat yields on a consistent basis for the pre-partition period 1911 to 1938. Column 1 shows the regression of annual wheat yields for the period 1957 to 2009 on the log number of Migrants in 1951. Column 2 adds the interaction between the green revolution dummy and log number of Migrants in 1951, the green revolution dummy and the interaction between the green revolution dummy and log of population in 1951 to the regression in Column 1. The unit in which yields are measured is tons per hectare. Columns 3 and 4 replace annual wheat yields with the take-up of High Yield Varieties (i.e. HYV) of Wheat as the dependent variable. The wheat HYV take-up regressions in this table are based on the sample of 68 districts for which there was data available on wheat yields on a consistent basis for the pre-partition period 1911 to 1938. The HYV take-up of Wheat is measured as the proprotion of total area planted to all crops in a given year that is devoted to HYVs of wheat. All the above regressions include log of population in 1951, state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type, annual rainfall and population density in earliest available year which is 1961. * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 5b. Annual Wheat Yields, HYV take up and Log Migrants at Partition (excluing Bihar, Bengal and Orissa)

	Annual wheat yields		Take up of HYV Variety	
	(1)	(2)	(3)	(4)
Log Migrants (1951)	0.0796982**	0.0432584	0.0099287***	-0.0061794
	(0.03456)	(0.0314304)	(0.0026516)	(0.0039061)
Log Migrants (1951) X Green Revolution		0.0577688***		0.0314265***
		(0.019459)		(0.0057455)
Mean Outcome	1.541978	1.541978	0.1286519	0.1286519
No of observations	2650	2650	2287	2287
Controls				
Soil type Dummies	Yes	Yes	Yes	Yes
Annual Rainfall	Yes	Yes	Yes	Yes
Population Density (1961)	Yes	Yes	Yes	Yes

Notes: The yields regressions in this table are based on the sample of 52 districts (excluding those of Bihar, Bengal and Orissa) for which there was data available on wheat yields on a consistent basis for the pre-partition period 1911 to 1938. Column 1 shows the regression of annual wheat yields for the period 1957 to 2009 on the log number of Migrants in 1951. Column 2 adds the interaction between the green revolution dummy and log number of Migrants in 1951, the green revolution dummy and the interaction between the green revolution dummy and log of population in 1951 to the regression in Column 1. The unit in which yields are measured is tons per hectare. Columns 3 and 4 replace annual wheat yields with the take-up of High Yield Varieties (i.e. HYV) of Wheat as the dependent variable. The wheat HYV take-up regressions in this table are based on the sample of 53 districts (excluding those of Bihar, Bengal and Orissa) for which there was data available on wheat yields on a consistent basis for the pre-partition period 1911 to 1938. The HYV take-up of Wheat is measured as the proprotion of total area planted to all crops in a given year that is devoted to HYVs of wheat. All the above models include log of population in 1951, state fixed effects, year fixed effects, state-specific time trends, 21 dummies which caputre soil type, annual rainfall and population density in earliest available year which is 1961. * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 6a. Pairwise correlations between Rural Male Literacy and Log Migrants

	Proportion of rural males who are literate			
	1961	1971	1981	1991
Log Migrants (1951)	0.1204*	0.1731***	0.1945***	0.3128***
	(0.0637)	(0.0087)	(0.0031)	(0.0001)
No of observations	238	229	229	151

Notes: This table shows the pairwise correlation coefficients between the proportion of rural mal literate and the log number of Migrants in 1951 for each census year in the period between 1961 a sample upon which the correlations are based excludes districts of the Rajasthan state due to miss the case of the census year 2001 quite a lof of the data on rural male literacy rates is missing and numbers for that year are not reliable. In the case of the census year 2001 quite a lof of the data c literacy rates is missing and therefore the numbers for that year are not reliable. * significant at the better, ** significant at the 5% level or better, *** significant at the 1% level or better. Standard er parentheses.

Appendix Table 6b. Pairwise correlations between Rural Male Literacy and HYV take-up year

	Proportion of rural males who are literate			
	1961	1971	1981	1991
Log Migrants (1951)	N/A	0.2691***	0.0843	0.2640***
	N/A	(0.0000)	(0.1996)	(0.0007)
No of observations	N/A	223	233	161

Notes: This table shows the pairwise correlation coefficients between the proportion of rural mal literate and the HYV take-up of all major crops for each census year in the period between 1961 a HYV adoption did not start to happen until after 1961 and therefore no numbers are reported fo sample upon which the correlations are based excludes districts of the Rajasthan state due to miss the case of the census year 2001 quite a lof of the data on rural male literacy rates is missing and numbers for that year are not reliable. * significant at the 10% level or better, ** significant at the better, *** significant at the 1% level or better. Standard errors in parentheses.

Appendix Table 7: Anecdotal evidence relating to litreacy of Migrants in districts from which they emigrated

District	Statement on literacy of Hindus & Sikhs	Source
Muzaffargarh	The district is still one of the most backward in the province. This backwardness is mainly among Muhammadans as the percentage of educated Muhammadans is only 1.12 against 15.47 of educated Hindus and 6.55 of educated Sikhs. No special measures are necessary in the case of Hindus and Sikhs as they are ready to take advantage of every opportunity; indeed, there are two private Hindu high schools in the district	Muzaffargarh District Gazetteer, 1929. Page 291.
Attock	Attock District is the most illiterate district in the Rawalpindi Division .... Literacy is highest among Hindus and Sikhs, among the non-Christian population	Attock District Gazetteer, 1907. Page 304.
Shahpur	The increase since the last census in the proportion of total males educated is largest for Sikhs, next for Hindus, and least of all for Musalmans; and now among Hindus one male out of every three is educated. Among the Sikhs almost half the male population have some education, but among Musalmans only about one if forty can read and write	Shahpur District Gazetteer, 1897. Page 91.
Jhelum	In the last Census, of the educated classes, 11,969 were Hindus and 9071 Musalmans. These figures are remarkable, when it is observed that the entire Hindu population amounts to 51,801, while the Musalmans number 526,725	Jhelum District Gazetteer, 1904. Page 259.
Lahore	The Sikhs show considerable progress in the education of males. The leading place taken by Hindus in the education among the three religions is remarkable	Lahore District Gazetteer, 1893-94. Page 84.
Multan	Among the Hindus the standard [of literacy] is very high, but the Muhammadans are very backward	Multan District Gazetteer, 1923-24. Page 260.
Gujrat	1.The majority of the girls are Hindu or Sikh. The Muhammadans are still slow to send their girls to school 2. He [the Hindu] has been quick to take advantage of the ducation in the schools, and his children have secured a large proprotion of Government appointments	Gujrat District Gazetteer, 1921. Page 156 and 50.
Jhang	Among Hindus and Sikhs apparently one in three males can read and write, while among Muhammadans about one in sixty only … Viewed in percentages the resut for females is less unsatisfactory, as among Hindus and Sikhs education has increased tenfold in the twenty and doubled among the Muhammadan population	Jhang District Gazetteer, 1908. Page 158.
Rawalpindi	As was expected, the Muhammadan portion of the population, in other words, the agricultural class, is grossly ignorant. Only two persons in a hundred can read and write, and only one is learning. Jains appear to be given a batter education than Sikhs, and Sikhs than Hindus. The district cannot be congratulated on the literacy acquirements of its inhabitants	Rawalpindi District Gazetteer, 1893-98.