Kent Academic Repository
Full text document (pdf)

Citation for published version

DOI

Link to record in KAR

Document Version

Author's Accepted Manuscript

KAR Kent Academic Repository

University of **Kent**

Displacement and Development: Long Term Impacts of Population Transfer in India [*]

Prashant Bharadwaj[†]   Rinchan Ali Mirza[‡]

January 2019

Abstract

The partition of British India in 1947 resulted in one of the largest and most rapid migrations and population transfers of the 20th century. Using refugee presence by 1951 as a measure for the intensity of the impact of the population transfer, and district level data on agricultural output between 1911-2009 from India, we find using difference in differences and event study approaches that areas that received more refugees have higher average yields, are more likely to take up high yielding varieties of seeds, and are more likely to use agricultural technologies. The increase in yields and use of agricultural technology coincide with the timing of the Green Revolution in India. Using pre-partition data, we show that refugee placement is uncorrelated with soil and water table characteristics, agricultural infrastructure, and agricultural yields *prior* to 1947; hence, the effects are not explained by selective movement into districts with a higher potential for agricultural development. We highlight refugee literacy and land reforms in areas with refugees as two of the many potential mechanisms that could be driving these effects.

JEL Codes: O13, O33, O15, Q16, N55

[†]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
[‡]Namur University, Belgium. Website: www.rinchanmirza.com

1 INTRODUCTION

The end of the British Empire in India in 1947 was marked with a mass migration and population transfer of nearly 17 million people. By many accounts this was a human rights disaster involving nearly a million deaths due to the riots that ensued between hindus and muslims on either side of the newly created India-Pakistan border. Historical events are known to shape modern day institutions and economic 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) and this paper seeks to examine the legacy of the migration and population transfer that took place at the time of partition on agricultural development, an important aspect of economic progress in India.[1]

We find that areas that received refugees did better in the long run in terms of agricultural development. Documenting this relationship is an important contribution as mass migrations, institutional upheaval, and partitions are a reality even today.[2] It is therefore crucial to understand how communities and areas develop long after such events take place. While almost by definition affected areas suffer in the short run, it is important to document whether the legacy of such events forever change the long run trajectory of these places.

We find that areas that received a high fraction of refugees by 1951 had significantly higher yields compared to areas with low fraction of refugees in the decades after India's independence. Between 1957 and 2009, districts that had a greater refugee presence saw average annual wheat yields increase by 9.4% compared to low refugee districts. We find similar results when examining annual revenue per hectare.[3] The take off in agriculture in

---

[1]Data limitations prevent us from examining these impacts on the Pakistani side. More details on this in the Appendix.

[2]The European refugee crisis (2015-2017) 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 the partitioning of the Southern Sudan Autonomous Region from the rest of the country was officially declared 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).

[3]This measure is used so as to not be reliant on any specific crop for our productivity measures. The 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.

refugee settled areas occurred at the same time as the green revolution started in India. 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 refugee presence is strongly correlated with the use of tractors (going from a low refugee district to a high refugee district increases the use of tractors by 56% between 1957-1987) and fertilizers (phosphorus and nitrogen).

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 concern with examining simple correlations of refugee presence and outcomes is that despite the uncertainty and chaos of partition, refugees 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 were not on differential trends until the start of the green revolution. While limited in our ability to examine *trends* along certain other variables, we use available data to examine at least in levels whether refugees 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, aquifer depth in districts (groundwater access has been shown to be a strong predictor of the green revolution, see D'Agostino (2017)), and refugee presence. We also find no correlation between refugee migration and the presence of other infrastructure variables such as banks, post offices, length of roads, and hospitals by 1961 (pre green revolution). 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 refugee presence might be generally related to district characteristics or trends that affect agricultural yields. Our results on refugee presence and yields are however *only* present for crops that were affected by the green revolution. Non green revolution crops, such as millets, chickpeas, rapeseed etc. do not show any changes in yields with refugee 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

3

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 our 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.[4]

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, increased religious homogeneity, and loss and restructuring of land, just to name a few. Hence, our interpretation of the results is that areas (districts in our case) that received refugees due to partition were more "affected" along various dimensions, like the ones we just mentioned, by partition than districts that did not receive any refugees.[5] While we use the refugee population as our metric for the intensity of the impact of partition, it would be incorrect to interpret our results as solely the effects of partition induced *migration*. For example, districts with more refugees could have received more government aid in the years after partition, and our effects should be interpreted as capturing the reduced form effect of both refugee presence and government assistance. 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 and communities. At its core, this paper

---

[4]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).

[5]Note that this also comes with an important interpretation when studying the impacts of migration or refugee flows that "more" refugees implies higher intensity of impact – this is only true if the composition of migrations in areas with "more" migrants is the same as the composition of migrants in areas with "less" migrants. If more educated (but fewer in number) refugees travel further then just using refugee *numbers* is not adequate to measure impact. While it is entirely plausible that different types of refugees ended up in Tamil Nadu relative to the Punjab, the fact that we use state fixed effects and examine variation in refugee settlements across districts but *within* a state mitigates this concern.

is a reduced form way of understanding how places that were affected by partition related population transfer fare in the long run.

Data limitations, the sheer magnitude of the event, and the two way nature of the refugee flows therefore makes it nearly impossible to make precise statements about any *one* leading factor. We do, however, provide some preliminary evidence that the composition of refugees 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. These findings are also consistent with the seminal work of Foster and Rosenzweig (1996).

We explore two other mechanisms, but provide more qualitative evidence towards these. It is understood that hindu refugees were more likely to have been involved in money lending and the data clearly show that refugees were involved in commercial activities. 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 refugees during the green revolution helped along this dimension. This mechanism is similar in spirit to the recent excellent work by Bazzi, Gaduh, Rothenberg, and Wong (2016), who examine the consequences of forced migration in Indonesia. Second, the way land was redistributed to refugees likely resulted in more land consolidation and redistribution in areas that had more refugees. If consolidated lands were more likely to see investments in technology during the green revolution, then this could be an important channel for the results. While we do not have data on farm size before and after partition at the district level, we provide qualitative evidence towards this mechanism. We hope future research in this area can shed empirical evidence towards this mechanism. Finally, since partition resulted in two way migration flows (muslims leaving India, replaced by hindus and sikhs arriving from Pakistan/East Pakistan) resulting in no major *net* population change at the district level in India (at least on the Western border), 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, or the local agglomeration effects in West Germany studied by Peters (2017). We are well aware that the mechanisms we explore are neither conclusive nor exhaustive. However, even identifying the reduced form impacts of refugee movements

poses significant challenges (both in terms of data and econometric identification) and we leave the deep exploration of individual mechanisms to future work.

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 refugees) impacted long run economic outcomes such as agricultural development.

2   BACKGROUND

2.1   PARTITION OF INDIA

Although the possibility that British India would be partitioned upon independence had been present for several years prior to the actual event, the partition when it finally came on August the 14th, 1947 was sudden, violent and chaotic for the millions who found themselves on the wrong side of the newly created India-Pakistan border. An estimated 14.5 million people were displaced within a short span of just four years after partition (Bharadwaj, Khwaja, and Mian, 2008, p. 39).

The decision to partition was formally laid out by the British in the shape of a plan on June the 3rd, 1947. The partition plan, as it was called, laid the foundations for redrawing the boundaries of the two most contested states of Punjab and Bengal. A British civil servant, Cyril Radcliffe, was appointed chairman of the Punjab and Bengal boundary

commissions. Radcliffe was both unfamiliar with boundary making and had no intimate knowledge of the people or the land he was about to carve up (Yong and Kudaisya, 2000, p. 84). His task was further complicated by the procedural difficulties involved in the partition plan itself. While all major political parties had agreed upon partition, they had vaguely laid down that boundaries would be demarcated by contiguous majority areas of muslims and non-muslims as well as by considering "other factors". Further adding to the chaos and confusion surrounding partition was the fact that Radcliffe's decision regarding the placement of the partition border was kept secret until the very last minute. Such secrecy heightened speculation regarding his methods of demarcating the border. It was 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. The reports of the boundary commissions were eventually made public on August 17th, 1947, two days after India had declared its independence. Immediately afterwards there were voices of dissent coming from all quarters. The border "zigzagged precariously across agricultural land, cut off communities from their sacred pilgrimage sites, paid no heed to railway lines or the integrity of forests, [and] divorced industrial plants from the agricultural hinterlands where raw materials, such as jute, were grown" (Khan, 2017, p. 126).

Once the partition line was revealed, 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 (Yong and Kudaisya, 2000, p. 98). Even before the declaration of independence, the violence in both Punjab and Bengal had started to take its toll with several incidents of rioting between the muslims on one side and hindus and sikhs on the other.

The above discussion points to important features of the partition that are relevant to our paper. First, the rapidity with which the partition unfolded meant that the majority of migratory flows took place under a relatively short span of time and without much preparation. Second, 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 Refugee Resettlement

This section discusses three specific aspects of the partition-related refugee resettlement: determinants of refugee resettlement, the process of refugee resettlement and the immedi-

ate consequences of refugee resettlement. Each of these aspects has an important bearing on our study that is worth explaining in more detail.

2.2.1 DETERMINANTS OF REFUGEE RESETTLEMENT

An understanding of the factors behind partition related refugee resettlement is important because it is the main source of variation we exploit for measuring the intensity of the partition-induced population transfer. Bharadwaj, Khwaja, and Mian (2008) identify a "replacement effect" according to which districts that experienced greater evacuee outflows also received greater refugee inflows. The replacement effect was found to be significant and very large for both India and Pakistan (Bharadwaj, Khwaja, and Mian, 2008). Although Bharadwaj, Khwaja, and Mian (2008) made an important contribution in identifying the replacement effect it did not tackle the issue of what caused the evacuee outflows in the first place. In another important study, Jha and Wilkinson (2012) identity violence as one of the key determinants of the variation in evacuee outflows. The study shows that districts that had a higher presence of demobilised soldiers with frontline combat experience suffered from greater levels of partition-related violence. The higher levels of partition-related violence in turn led to greater outflow of evacuees in such districts.

The distances that refugees had to travel in order to reach safety "imposed costs on them that entered into their location decisions" (Bharadwaj and Fenske, 2012, p. 6). As a consequence they were more likely to resettle in areas closer to the partition border. Bharadwaj, Khwaja, and Mian (2008) term such a phenomenon the "distance to the border effect" and consider it to be an important determinant of migratory flows. The distance to the border effect was found to be only statistically significant in India and generally of small size (Bharadwaj, Khwaja, and Mian, 2008).

Another important pattern the Bharadwaj, Khwaja, and Mian (2008) study uncovers is the tendency of migratory inflows to be greater in the major urban centres of India and Pakistan. Bharadwaj, Khwaja, and Mian (2008) term such a pattern the "big city effect" and attribute it to the greater economic opportunities that were on offer for the refugees in the major urban centres. The big city effect was found to hold strongly for India and was not statistically significant for Pakistan (Bharadwaj, Khwaja, and Mian, 2008).

2.2.2 REFUGEE RESETTLEMENT AND LAND

As highlighted previously, the amount of land that was available for permanent allotments to refugees in any given district depended crucially upon the extent of evacuee outflows

the district had experienced due to partition. Working under such a binding constraint, resettlement administrators devised a scheme to re-distribute evacuee land to refugees. The two main components of the scheme were the "standard acre" and the "system of graded cuts". The standard acre was different from the physical acre and represented a unit of value based on the productivity of the land rather than the physical area covered by the land. The diversity in local productivity conditions meant that the physical area behind a standard acre varied substantially across districts.

After valuing land in standard acre terms the resettlement administrators used a system of graded cuts to make permanent allotments to the refugees. According to the system of graded cuts, refugee claims were placed into different categories based on their size expressed in terms of standard acres. A different discount rate, known as the rate of cut, was then applied to the claims in each size category to reach the net allotment of land to refugees. The sliding scale used by the administrators under the graded cuts system is summarised in Figure 1. It is clear from Figure 1 that the graded cuts system involved a substantial re-distribution from large landholders to small and medium sized owner-cultivators amongst the refugees. For small claims of up to ten standard acres the rate of cut was 25 per cent. On the other hand, refugee claims upwards of 500 standard acres were subjected to a 95 per cent rate of cut. Alongside the re-distribution of land the scheme devised by resettlement administrators also led to a substantial consolidation of fragmented land holdings. It was the "immense scale of property transfers" involved in re-distributing evacuee land to refugees that provided the impetus behind land consolidation (Yong and Kudaisya, 2000, p. 83).

FIGURE 1: Sliding scale used for net allotment of refugees

Grade (Standard Acres)	Rate of Cut (per cent)	Net Allotment (Standard Acres)
Up to 10	25	7.5
10 to 30	30	21.5
0 to 40	40	27.5
40 to 60	55	36.5
60 to 100	70	48.5
100 to 150	75	61
150 to 200	80	71
200 to 250	85	78.5
250 to 500	90	103.5
500 to 1000	95	128.5

Source: Randhawa, Out of the Ashes, p. 99 (year).

The system of refugee land compensation devised by resettlement administrators described above only applies to the northwestern border of India. In contrast, the process of resettling refugees on land along the eastern border differed in several respects.

The process of land compensation for refugees along the northwestern border was both organized and swift, whereas the same process along the eastern border was both disorganized and protracted over a long period of time. Two main reasons explain the failure of the Indian government to resettle refugees along the eastern border. First, the amount of evacuee land that became available along the eastern border after Partition was not adequate for resettling refugees. This was because evacuee outflows were far less than refugee inflows along the eastern border (Yong and Kudaisya, 2000, p. 89). Second, both the Indian and Pakistani governments had a deliberate policy[6] to avoid a population exchange along the eastern border and therefore did not set up a system of land compensation for refugees.

2.2.3 REFUGEE RESETTLEMENT AND HUMAN CAPITAL

The resettlement of refugees increased literacy and caused a shift towards non-agricultural professions. There were two reasons behind the increase in literacy. First, the refugees were a highly literate community that contributed to increases in literacy of areas. Bharadwaj, Khwaja, and Mian (2015) find that for districts along the northwestern border a 10 percent increase in refugee inflows "resulted in an increase in literacy of nearly 36 percent" (Bharadwaj, Khwaja, and Mian, 2015, p. 91). Second, the out-migrating evacuees whom the refugees replaced were a largely illiterate community whose out-migration tended to increase the literacy of areas. Bharadwaj, Khwaja, and Mian (2015) find a small positive impact of evacuee outflows on the literacy of districts.

In addition to increasing literacy the resettlement of refugees also caused a shift towards non-agricultural professions. There were two major reasons behind such a shift. First, the refugees were overwhelmingly concentrated in non-agricultural professions such as trade and commerce. Second, the complexities and uncertainties involved in the "land-to-land transfer of occupation" for agricultural refugees meant that some refugee farmers had to move into non-agricultural professions (Bharadwaj, Khwaja, and Mian, 2015, p. 96). Such a movement further reinforced the shift towards non-agricultural occupations in areas where refugees settled in greater numbers.

---

[6]The Nehru-Liaqat pact of 1950 attempted to create a sense of security amongst minorities in both East and West Bengal to avoid a population exchange (Yong and Kudaisya, 2000, p. 87).

The discussion in Sections 2.2.1 to 2.2.3 above highlights several features of partition-related refugee resettlement that are important to the context of our study. First, it identifies three important factors—evacuee outflows, distance to the border and the presence of a big city—that were behind the variation in refugee resettlement across districts. Second, it shows that the assessments resettlement administrators made about the productivity of the land that was reflected in the standard acre valuations also contributed to varying levels of refugee resettlement. Refugee land allotments were larger in size in districts where a greater number of physical acres were required to make up the standard acre. Third, it points out that the graded cuts system caused the re-distribution and consolidation of agricultural land. We will explore both changes to the land regime as potential mechanisms through which the partition-induced population transfer could have impacted long-run agricultural development. Fourth, it argues that the resettlement of refugees led to an increase in literacy and a shift toward non-agricultural occupations. Both literacy and non-agricultural occupations are important for agricultural development and will be explored as potential mechanisms in this paper. Finally, the discussion also points out differences in refugee resettlement between the eastern and the northwestern borders of post-partition India. Such differences mean that our analysis is more relevant to the population transfer along the northwestern border. However, despite the lack of relevance our results also hold for states along the eastern border.

2.3 Green Revolution in India

2.3.1 Development of High Yielding Varieties

The Green Revolution originates from the cross-breeding 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, 2003, p. 758). The development and diffusion 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, 2003). As late as the 1980s only a few HYVs of crops like sorghum and millet had been developed (Evenson and Gollin, 2003, 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 in households producing wheat and rice.

2.3.2 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 large scale release of such locally adapted varieties in the late 1960s that marked the start of the Green Revolution in India. 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[7]. Due to the rapid adoption of the wheat varieties, 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 the 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 the continuing research efforts, a new generation of wheat varieties were developed that were able to penetrate into 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 less 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 limited success of the early rice varieties Indian agricultural scientists concentrated their research efforts on developing varieties that were suited to local conditions in rice areas (Munshi, 2004, p. 190) and also incorporated resistance to pests and diseases (Evenson and Gollin, 2003, p. 759).

---

[7]The states of Punjab, Haryana, (western) Uttar Pradesh, Delhi and Rajasthan make up the North-west 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.

The greater success of wheat HYVs relative to the rice varieties means that we focus exclusively on wheat yields in this paper; although we do show results for all crops in an aggregate revenue per acre measure as well.

2.3.3 ROLE OF CANALS AND AQUIFERS

In addition to the differences in seed technology discussed above, there were other factors that were important in the adoption of rice and wheat HYVs in India. One of these was the timely and controlled provision of water (Rud, 2012, p. 353). The uninterrupted supply of water at specific periods of growth, development and flowering was crucial to the successful performance of the HYVs. That is why pre-existing patterns of irrigation and climate were one of the main drivers behind their diffusion (Gollin, Hansen, and Wingender, 2016, p. 5). The importance of irrigation can be gauged by the fact that states like Punjab, Haryana and Tamil Nadu that had a well developed irrigation infrastructure dating back to the colonial period rapidly adopted HYVs shortly after the start of the Green Revolution (Evenson and Gollin, 2003b, p. 91). On the other hand, at around the same time states like Gujarat, Maharashtra, Orissa, West Bengal, Bihar, Kerala and Rajasthan, that lacked an extensive irrigation network, were lagging behind considerably in terms of HYV adoption (NCAI-I, 1976, p. 284).

Beginning with the introduction of HYVs in the late 1960s, more minor irrigation projects were undertaken to rapidly expand irrigation beyond the historically canal-irrigated states. The minor irrigation projects used electrified tube-wells to access groundwater instead of using canals to access river water (Rud, 2012, p. 353). They offered greater control in terms of flow and timing of water supplies compared with canal irrigation (NCAI-V, 1976, p. 20). Due to their cost effectiveness, the minor irrigation projects were also particularly attractive for small and medium farmers in areas without canal irrigation. Recognizing the importance of minor irrigation projects the Indian state financed the extension of the electricity network across rural India and provided credit to farmers for purchasing electrified tube-wells (NCAI-V, 1976, p. 20). As a consequence, the use of electrified tube-wells for groundwater irrigation accelerated remarkably after the start of the Green Revolution in the late 1960s. It is important to note that groundwater irrigation depends crucially on the depth of aquifers (D'Agostino, 2017). Aquifers are underground layers of water bearing permeable rock, rock fractures or other unconsolidated materials from which groundwater can be extracted. The closer an aquifer is to the surface (i.e. of lower depth) the more likely it is to be used for groundwater irriga-

13

tion, and hence, a strong correlation between aquifer depth and HYV adoption emerges (Dasgupta, 2018). Hence, in this paper we crucially account for the presence of historic canals as well as aquifer depth in our estimations.

3 DATA AND EMPIRICAL FRAMEWORK

3.1 FULL SAMPLE ANALYSIS

For our full sample analysis that relates to the post-partition period 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). Details on each of these sources are provided in Appendix C. The 1951 census data was used to construct a measure of refugee presence that was then related to measures of agricultural development from 1957 to 2009 that were constructed from data in the IACD and VDSA datasets.[8] An important task in relating the two measures was to make district boundaries comparable between 1951, the year in which data on partition refugees was recorded, and the first year for which data is available in the combined IACD—VDSA panel dataset (i.e.1957).[9] 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.

---

[8]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. Details on the procedure we used to combine the IACD and VDSA datasets are provided in Appendix C.

[9]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).

3.2 Sub Sample Analysis

For our sub-sample analysis we combine post-partition data from the joint IACD—VDSA dataset with pre-partition data from the Agricultural Statistics Reports of British India for a select group of districts for which data on crop yields is available on a consistent basis throughout the pre and post partition periods. The four major crops we focus on are wheat, rice, sugar and maize. We have already described the data contained in the joint IACD—VDSA dataset above. The Agricultural Statistics Reports of British India, produced on an annual basis by the Department of Revenue and Agriculture, contained information on yields for all major crops and most other crops for a select group of districts. Although the reports came out on an annual basis, the yield numbers were only revised intermittently after gaps of several years. Therefore, the panel we construct for our sub sample difference in differences analysis contains information on yields for only four pre-partition years between 1910 and 1940.[10] 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 panel for the sub sample analysis was determined by the substandard quality of data prior to that date.

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 refugees relative to areas that did not receive refugees. 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:

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

---

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

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 refugee presence measure which is also the main variable of interest. The refugee presence measure that proxies for the intensity of the effect of partition, D_{is}^{51}, is defined as a dummy variable that takes a value of 1 if the fraction of the population that is composed of refugees in 1951 in district i, in state s, is above the median based on the full sample of districts. $D_{is}^{51} \times Post_t$ is the interaction of the refugee presence measure with time (either via a single "Post" dummy that indicates either the post-partition or the post-green revolution period, or simply year dummies in a more flexible specification). 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 21 different soil types at the district level. We control for each of these soil types as soil quality plays an important role in both the adoption of agricultural technology and in agricultural productivity. We also 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. Finally, given the panel nature of the data we cluster the standard errors at the district level.

3.3.2 FULL SAMPLE PANEL REGRESSIONS

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

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

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 computes the total revenue generated from all crops that are produced in a district using a single calendar year price (in our case 1960)[11], and then divides the resulting total by the area under cultivation in that district. The unit of our revenue based measure is "revenue per acre". In additional specifications, we use measures of technology adoption other than HYV seeds (tractors per acre and fertilizers per acre) to

---

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

16

further examine the role of refugees in the overall advancement of agricultural technology. 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 (see Appendix Tables 1 and 2). The main advantage of the panel form is in our examination of the effect of refugees 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 differences specification used for the sub sample analysis.

It is important to reiterate that our estimates on the refugee presence measure, D_{is}^{51}, in the above estimating equation represents a reduced form or "net" effect of refugee settlement and associated changes due to settlement 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

4.1 SUB SAMPLE ANALYSIS

We first present results using the sub sample of districts for whom we have data for years prior to the partition. The results from estimating equation 1 is presented in Table 2. Aside from our preferred refugee presence measure—a dummy variable that takes on a value of 1 if a district is above the median in terms of the fraction of its population in 1951 that is composed of refugees—we also use the log number of refugees to capture the effects of partition. Table 2 shows that districts with a greater refugee presence 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 or equal to 10%. In particular Column 8 of Table 2 shows a stark result when we examine the effects by each 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 the high refugee districts. Since it is easier to interpret the magnitudes of the coefficients on a dummy variable, we choose the high refugee dummy as our preferred measure for the intensity of the effect of partition.

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

Table 3 examines whether other crops not directly associated with improvements in seed technology during the green revolution responded similarly to refugee presence. Our main argument is that refugee 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 refugees were somehow better at farming *all* crops. Table 3 shows that this is broadly not the case across eight other crops for which we have consistent data. Figure 3 follows the same methodology as Figure 2 and shows 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 refugee areas.

Table 4 uses alternative measures of partition affectedness. Bharadwaj, Khwaja, and Mian (2009) established two crucial facts about the migratory flows in India after the partition: refugee inflows of hindus and sikhs were correlated with refugee outflows of muslims, and areas closer to the partition border experienced both greater refugee inflows as well as refugee outflows. Hence, we can employ two alternatives measures to capture the effects of partition without actually using refugee presence in 1951. While magnitudes are not easily interpretable in this table, we show consistency of sign and significance. In columns 1 and 2 of Table 4, we show that areas that had more muslims in 1931 see similar patterns regarding wheat yields as in Table 2. Columns 3-6 employ a more nuanced proxy for partition impacts: while areas with more muslims pre-partition were

18

FIGURE 2: Refugee Presence and Wheat Yields
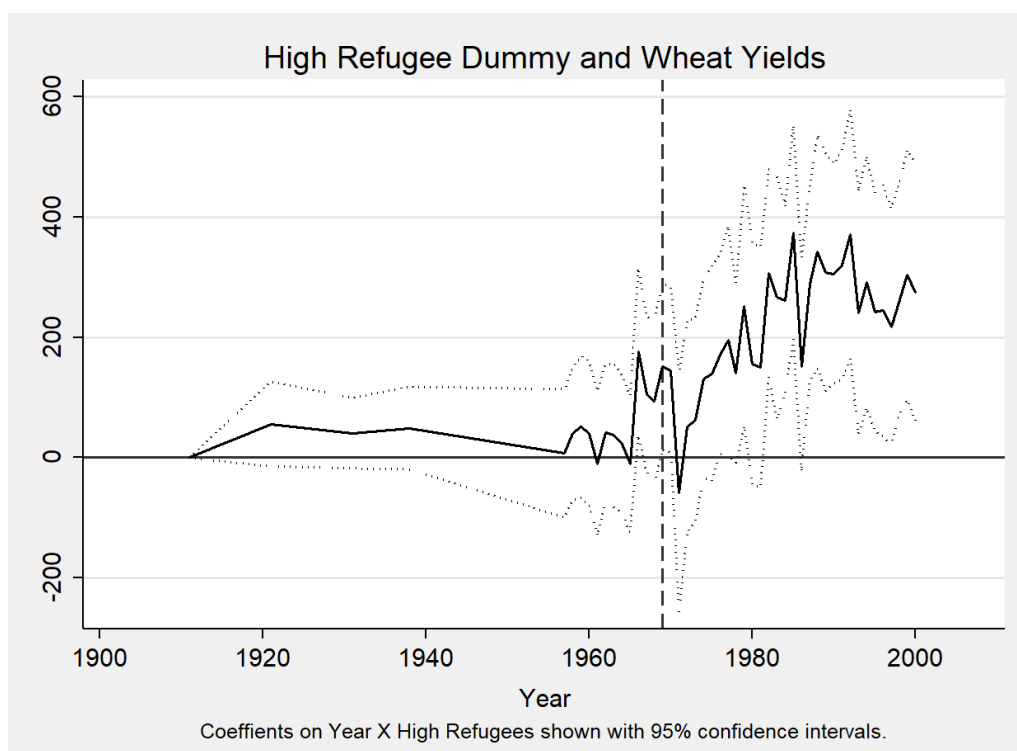
Notes: Each point on the graph is the interaction coefficient from a regression where year dummies are interacted with a dummy for high refugee presence at the district level. The regression controls for the main effects, state fixed effects and state specific quadratic trends along with controls for 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. The vertical line in the figure is at 1969.

FIGURE 3: Refugee Presence and Other Crops
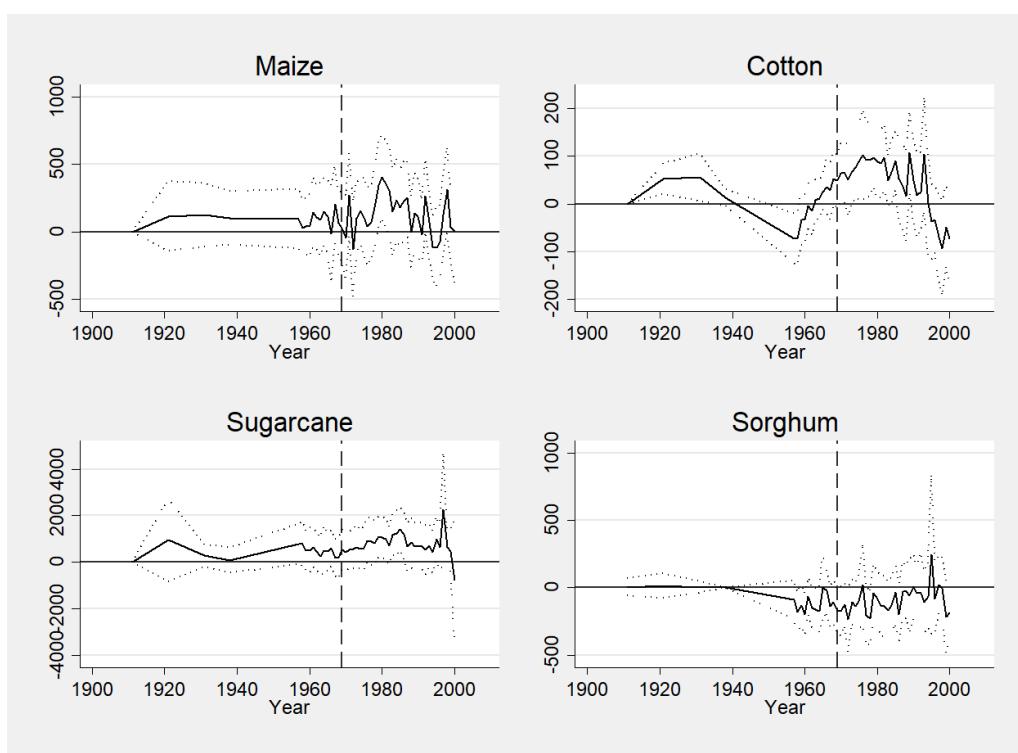
Notes: See Figure 2 and text for details.

more likely to see refugees in 1951, the effect differed by whether these muslim areas were close to the border. Hence, in columns 3-6 we use the interaction of fraction muslims in 1931 and distance to the border as a measure for places that were affected by partition related refugee flows. The triple interaction term in column 4 is positive and significant; however, the triple interaction term in Column 6 is negative and insignificant as this column controls for all double interactions, whereas column 4 only controls for the post green revolution dummy and proportion muslim in 1931 and inverse distance to the border as main effects. However, note that in column 6, the interaction term of post green revolution dummy and inverse distance to the border is positive (and indeed as are all the double interaction terms, although not statistically significant), suggesting that areas closer to the border had greater gains after the green revolution.

4.1.1 ROBUSTNESS OF SUBSAMPLE RESULTS

As discussed in Section 2, it is crucial that we account for characteristics of districts that might have, for reasons unrelated to refugee flows, made them more suitable for green revolution technologies. In Table 5, we show robustness to an aforementioned important characteristic – aquifer depth. Table 5 essentially replicates the specification in Table 2, but also controls for aquifer depth at the district level. The aquifer depth controls enter flexibly in separate specifications through their interaction with either a post-partition dummy, a post green revolution dummy or a series of decade dummies. Coefficients in Table 5 look very similar to Table 2, and hence, our takeaway from this is that while aquifer depth may have been a crucial input into the take up of the green revolution, refugee presence is uncorrelated with aquifer depth.

Table 6 deals more directly with the influence of public infrastructure in the relationship between refugee presence and agricultural development by examining a host of outcomes: historic canal irrigation, pre-green revolution presence of banks, post offices, hospitals, schools, roads, and aquifer depth. The broad take away is that across a host of pre-green revolution district characteristics and across all measures of refugee presence we have used to proxy for the effects of partition in this paper, there does not appear to be a systematic correlation between factors that predict refugee flows and factors that may predict the take up of the green revolution.

In Table 7 we show that the results are robust to the inclusion of controls from Banerjee and Iyer (2005). As mentioned earlier, Banerjee and Iyer (2005) show important long term effects of colonization in areas that were under different forms of British governance.

21

In the event that refugee places correlated with these historical forms of land taxation, it is important to control for these factors. We obtained the data used in Banerjee and Iyer (2005) and Table 7 shows that controlling for these historical features does not alter our results.

Appendix Table 1 shows robustness to excluding districts from the sub-sample that were classified as "Intensive Agriculture District Programs" or "IADP" districts. According to Mohan and Evenson (1975), the IADP districts were essentially pilot districts that the Government of India used to figure out its plan to boost food production in the 1960s. Specifically, Mohan and Evenson (1975) list four main criteria that were used to select districts into the pilot program: 1) assured water supply, 2) minimum natural hazards, 3) well developed cooperatives and panchayats, and 4) high potential for rapid agricultural growth. While it is clear that refugee presence from 1951 is not one of the factors, we show that excluding these pilot IADP districts from our sub sample does not affect our results (Appendix Table 1).

As discussed in Section 2.2, the rehabilitation of partition refugees was based on a scheme of permanent land allotment that used 'standard acre valuations' for dispersing refugees across areas with variable agricultural conditions. Column 2 of Appendix Table 3 shows that there is indeed a strong positive association between the availability of standard acres and refugee presence at the district level for Punjab[12]. Moreover, it was also the case that the number of ordinary acres per standard acre that a district had determined its mean agricultural conditions (Kudaisya, 1995, p. 81-82). If it turns out that the number of ordinary acres per standard acre were somehow systematically correlated with trends in crop yields prior to partition then this could introduce a bias into our main results. In order to test whether such a bias exists we regress the growth in wheat yields from 1911-1938 on the number of ordinary acres per standard acre for Punjab in column 4 of Appendix Table 3.[13] As the result shows, the coefficient estimate in column 4 is insignificant and close to zero. This implies a lack of evidence for there being a relationship between the mean agricultural conditions (i.e. number of ordinary acres per standard acre) and trends in pre-partition wheat yields for Punjab.

---

[12]Additionally, Appendix Table 2, which lists the area in standard acres separately for High and Low refugee districts of Punjab, also shows a positive relationships between the availability of standard acres and refugee presence.

[13]Ideally, we would have liked to base the regression on the full sample of districts for which we have data on pre-partition wheat yields. However, data on the number of ordinary acres per standard acre is only available for districts of Punjab.

In Appendix Table 4, we show that our results are robust to restricting our sample to the two most impacted northwestern states of Punjab and United Provinces. As discussed in Section 2.2, the main focus of our study is on the northwestern border where the post-independence Indian state made concerted efforts for resettling refugees. It was, therefore, important for us to show that our results also hold when restricting the sample to only the northwestern states of United Provinces and Punjab.

Despite our paper being focused on the northwestern states we would still like to determine whether the results change substantively if we focus on other regions of post-independence India. Accordingly, in Appendix Table 5 we re-estimate our main specification after dropping the most impacted northwestern state of Punjab. As is clear from Appendix Table 5, even though the magnitude of the main impacts decline slightly[14], both the sign and significance of the results remains the same.

Finally, in Appendix Table 6 we re-estimate a variant of our main specification replacing the state fixed effects with district fixed effects. District fixed effects account for all time-invariant characteristics such as latitude, longitude, altitude, soil quality and aquifer depth that vary across districts. As is clear from Appendix Table 6, replacing state fixed effects by district fixed effects aside from reducing the magnitudes of the coefficient estimates does not change the main results in terms of sign or significance.

4.2 Full sample analysis

Having established in the prior section that refugee settlement is not correlated with factors that eventually predict agricultural success, in this section, we turn to analysis of all districts in India using data just from the post partition period. Analogous to Figure 2, we show fully flexible graphs with the full sample of districts where we interact the high refugee dummy with calendar year dummies along with the main effects, state fixed effects and state-by-year fixed effects included as controls. Hence, each point on the graphs in Figure 4 is the differential effect of higher refugee areas compared to lower refugee areas. The graphs show that high and low refugee areas within the same state were quite similar until the mid-late 1960s, after which the high refugee districts experienced 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). The

---

[14]Dropping Punjab from the sample reduces the coefficient on the interaction (post partition X high refugee) by 1.8 percent (column 6) and the coefficient on the interaction (post green revolution X high refugee) by 0.2 percent (column 7).

divergence in revenue and acreage under HYV seeds between high and low refugee districts does not appear to be as striking as that of wheat yields or tractor use. However, Tables 9 and 10 further on show strong evidence for the high refugee districts pulling ahead in terms of both revenue and acreage under HYV seeds after the start of the green revolution. Our main take away here is that the graphs in Figure 4 are broadly consistent with the main story in Figure 2.

FIGURE 4: Refugee presence and agricultural outcomes using full sample

Coeffients on Year X High refugee dummy shown with 95% confidence intervals. Vertical line is at 1969.
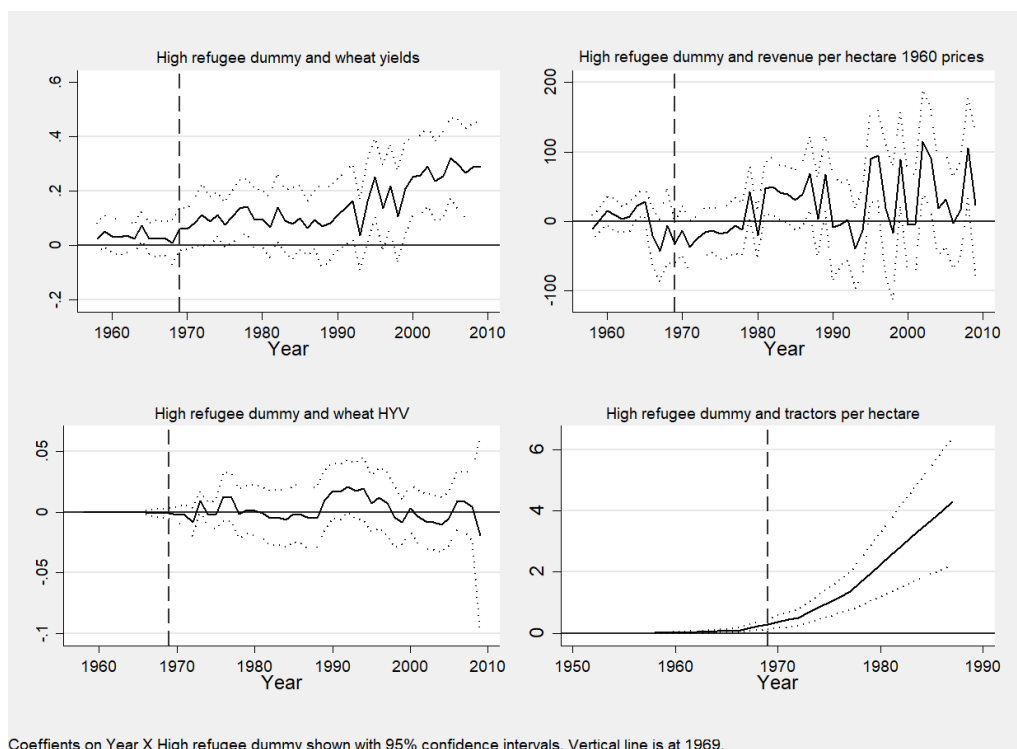
Notes: Each point on the above graphs is the interaction coefficient from a regression where year dummies are interacted with a dummy for high refugee presence at the district level. The regression controls for the main effects, state fixed effects and state-by-year fixed effects. The regressions are based on the full sample of districts for which annual agricultural data was available in the post partition period as described in the text. The vertical line represents the year 1969 which is the start of the Green Revolution in India. The vertical line in each of the plots is at 1969.

Turning to analysis in the tables, the outcome variable in Table 8 is revenue per acre based on 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 sample period is used is presented in Appendix Table 7). In column 1 of Table 8, we estimate equation 2 with no controls for soil conditions,

population density and rainfall, but including controls for state and year fixed effects, as well as state specific time trends. Column 1 shows that high refugee areas saw an increase in annual revenue per acre of nearly 51 rupees. Given the average revenue per hectare of approximately 486 Rupees, this is a meaningful increase of 10%.

In Columns 2 and 3, we sequentially add the controls for soil quality, population density and rainfall. We do this primarily to asses whether refugee selection into districts was systematically correlated with these variables, which might also affect the outcome of interest. Adding soil quality, population density and rainfall keeps the results largely stable, suggesting that refugee selection on the basis of soil quality and suitability for agriculture is not a concern in our case.

Table 9 examines whether the effect of refugee presence in a district is greater after the start of the green revolution. We define the start of the green revolution as the calendar year after which 5% or more of acreage (in the state of the district) is under HYV seeds.[15] Note that we do not interpret the timing of the green revolution as exogenous. In fact, as we show in Table 10, refugee presence at the district level was correlated with the take up of HYV seeds.

Table 10 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 refugee presence, and the effects are only larger after the start of the green revolution (the cross sectional results for take up of HYV are presented in Appendix Table 8). Visually, this is confirmed in Figure 4. Column 1 of Table 10 suggests that, compared to low refugee districts, high refugee areas saw yields increase by 9.4%. As expected, this effect is stronger after the green revolution occurs in a given district. Column 4 of Table 10 suggests that high refugee districts saw an increase in HYV use of 40% compared to low refugee districts after the start of the green revolution. Table 11 confirms the graphical result seen for tractor and fertilizer use in regression form - tractor use per acre is nearly 100% higher in high refuges areas compared to low refugee areas, and is even more so after the green revolution; nitrogen and phosphorus fertilizer use also increases quite dramatically after the start of the green revolution.

---

[15]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.

4.2.1 Robustness of full sample results

Our full sample results are broadly similar when we specify the right hand side variable in terms of proportion refugees (rather than the high refugee dummy) as shown in Appendix Table 9. While Columns 1 and 3 are not statistically significant, the refugee proportion interacted with the green revolution dummy is statistically significant. Appendix Table 10 shows that our full sample results are robust to exclusion of mismatching data across the overlapping years in the VDSA and IACD data sets. Appendix Table 11 restricts the full sample results to the smaller set of districts that we use in the sub sample analysis and the results are more or less consistent (although the results for wheat yields become insignificant).

5 Mechanisms

Our empirical analysis has shown a positive relationship between the refugee presence who arrived at partition and long-run agricultural development after the advent of the green revolution in India. Why was it the case that agricultural development after the start of the green revolution differed between areas with a greater or lesser extent of refugee presence? And why did the difference persist for an extended period that stretched beyond the intial phases of the green revolution? In this section we explore two possible channels that provide answers to such questions.

The two channels we focus on are (1) human capital and (2) the land regime. Firstly, the refugees who arrived in India differed in terms of their educational achievement and occupational choices from both the natives of the districts in which they settled and the evacuees who had left for Pakistan. Hence we argue for human capital as an important pathway through which refugees shaped persistent differences in agricultural development in the period after the start of the green revolution. Secondly, the post-independence government of India in its efforts to resettle the refugees on agricultural land aggressively implemented programs of land redistribution and consolidation in districts in which the refugees settled in large numbers. Accordingly, we argue that land redistribution and consolidation is another key channel through which high refugee areas were able to pull ahead of low refugee areas in terms of their agricultural development once the green revolution had started.

5.1 Refugees and Human Capital

5.1.1 Literacy

According to Bharadwaj, Khwaja, and Mian (2009) Indian districts that received refugees at Partition experienced a net increase in their literacy rates. Simple correlations in Appendix Table 12a show that refugee presence is indeed correlated with increased literacy of Indian districts in the years after partition. The correlation coefficient between the high refugee dummy 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.

The are two reasons why refugees were, at least in part, associated with net increases in literacy. First, the refugees formed a highly literate community in the areas of Pakistan from where they had emigrated. This is clear from Figure 5 that compares the pre-partition literacy of hindus and sikhs with that of the muslims in districts that became Pakistan. The hindus and sikhs who formed the refugees that came to post-independence India vastly out performed the muslims in terms of literacy throughout the four pre-partition census years of 1901, 1911, 1921 and 1931.

FIGURE 5: Hindu, Sikh and Muslim literacy prior to partition in districts that went to 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 Pakistan.

In addition to higher literacy of the in-coming refugees the lower literacy of the out-going evacuees also contributed to net increases in literacy. This is clear from Figure 6 that compares the pre-partition literacy of muslims with that of hindus in districts that became part of post-independence India. The muslims who formed the evacuees that

left for Pakistan under performed the hindus in terms of literacy throughout the four pre-partition census years of 1901, 1911, 1921 and 1931.

FIGURE 6: Hindu and Muslim literacy prior to partition in districts that went to India.
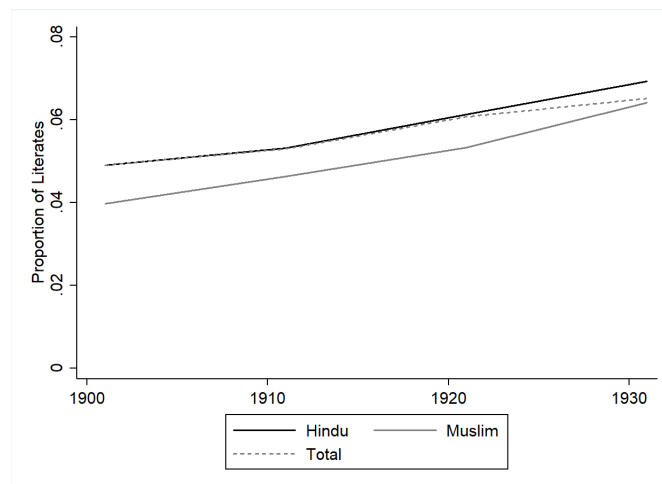
Notes: The figure is based on the eight colonial regions of Eastern Punjab, Uttar Pradesh, Western Bengal, Central Provinces & Berar, Bihar and Orissa, Madras, Bombay Presidency and Assam, all of which became part of post-independence India.

Official colonial documents also acknowledge the superior position held by the refugees in terms of education in the Pakistani districts from which they came. For instance, literacy was highest "among hindus and sikhs, among the non-christian population" of the Attock district[16]. 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[17]. 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[18]. 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 Pakistan is given in Appendix Table A6.

Other sources, outside of the official colonial publications, also point to the contribution the refugees had made to education. Raychaudhuri, Habib, and Kumar (1983) when discussing the aftermath of partition in Pakistan observe that the event led to the sudden

---

[16]Gazetteers, Punjab District. Gazetteer of the Attock District, 1907. Page 304

[17]Gazetteers, Punjab District. Gazetteer of the Lahore District, 1893-94. Page 84

[18]Gazetteers, Punjab District. Gazetteer of the Muzaffargarh District, 1929. Page 291

departure of teachers and instructors who mainly came from the hindu and sikh communities[19]. 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[20]. 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 went to 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"[21].

Another reason why the refugees could have led to the persistent differences in literacy documented in Appendix Table 12a is the inter-generational transmission of human capital. The correlation in human capital outcomes across generations is well documented in the literature examining persistent human capital differences. Rocha, Ferraz, and Soares (2017) find that the education of first generation immigrants into Brazil in the late nineteenth century, attracted through a state-sponsored open immigration policy, led to persistent differences in literacy outcomes throughout the twentieth century. Wantchekon, Klašnja, and Novta (2014) in their study on the US show that educated parents lead to better living standards, a shift away from farming and greater political activism amongst the future generations. Bleakley and Ferrie (2016) argue that it is through the cognitive ability and social aptitudes of parents that the inter-generational transfer of human capital takes place.

We postulate that the persistent change in literacy outcomes brought by the refugees in the districts in which they settled translated into the adoption of high yielding varieties of seeds during the green revolution. Several papers show a positive relationship between education and agricultural technology adoption. 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 found that the education of farmers had 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

---

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

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

[21]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

Revolution. Ram (1976) in yet another study on India shows that the contribution of farm operators to production was positively related to their education. Schultz (1964), Gerhart (1974) and Jamison, Lau, et al. (1982) all show similar relationships.

The argument usually put forward in such studies 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). Feder, Just, and Zilberman (1985) provide a comprehensive review of the broader literature connecting human capital to agricultural technology adoption. In our context, we show using simple correlations in Appendix Table 12b, 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 start of the green revolution. The correlation coefficient between take up of high yielding variety of all major crops and rural male literacy is 0.2691 in 1971, 0.0843 in 1981 and 0.2640 in 1991.

There is also some suggestive qualitative evidence that the higher literacy of the refugees had already led to them engaging in superior farming practices even before partition. The evidence relates to the minority hindu and sikh communities in those districts of colonial India that later became Pakistan. For instance, the hindu Jats[22] of Lyallpur district were considered by the colonial administrators as being the "most useful class of peasants"[23]. The hindu and sikh Jats of Sialkot district were deemed to be far superior cultivators than their muslim counterparts[24]. The gazetteer of the Lahore district notes that the hindu and sikh Jats were "good husbandmen"[25]. The sikh Virakhs[26] of the Montgomery district were considered first-rate cultivators[27]. 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"[28].

---

[22]An agricultural caste of the Punjab
[23]Gazetteers, Punjab District. Gazetteer of the Chenab Colony, 1904. Vol.A. Page 51
[24]Gazetteers, Punjab District. Gazetteer of the Sialkot District, 1893-94. Page 75
[25]Gazetteers, Punjab District. Gazetteer of the Lahore District, 1883-84. Page 65
[26]An agricultural caste of the Punjab
[27]Gazetteers, Punjab District. Gazetteer of the Montgomery District, 1898-99. Page 86
[28]Report on the Census of the Panjab Taken on the 17th of February 1881. Page 229

5.1.2 Money-Lending

Literacy is not the only dimension of human capital along which the refugees were different from the native population. From the qualitative evidence we have gathered we know that a significant number of the refugees 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. A substantial proportion amongst them belonged to the three great hindu and sikh mercantile castes of India—Khatris, Aroras and Baniahs—that dominated commercial activity. Figure 7 provides a snapshot of the advantage that the refugees had over the native population in commercial occupations at the time of partition. It compares the proportion of refugees 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 7: Refugees in the commercial sector
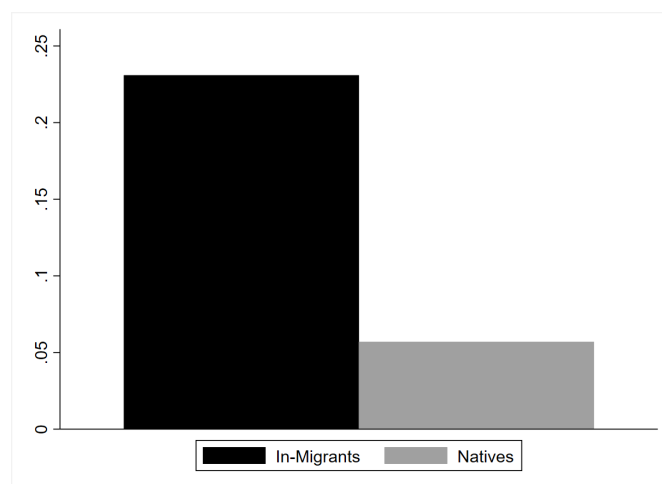
Notes: The black bar is the proportion of the displaced persons in 1951 that were previously engaged in commerce. This data is given in Appendix II of Table IV of the 1951 census of India. The grey bar is the proportion of the non-displaced persons (natives) that were previously engaged in commerce. This data is also available in the 1951 census of India.

As was the case with their educational superiority the higher concentration of the refugees in money-lending in the areas where they came from was also noted in official publications dating from the colonial period. The notes pertain to the hindu and sikh minority com-

munities in areas that later became Pakistan. For instance, the (1881) census of Punjab states that the hindus and sikhs were mostly traders.[29] Hindus from the Arora caste controlled "almost the whole of the trade, moneylending, and banking" in the Muzaffargarh district.[30] The hindu Aroras were also considered as being the "chief moneylenders and capitalists" and the "chief creditors of the agriculturists" in the Jhang district.[31] The hindus and sikhs from the Arora caste were identified as being the main moneylenders in the Montgomery district[32]. 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[33].

The higher concentration of the refugees in money-lending is likely to have had important consequences for agricultural technology adoption. It is plausible that refugee presence could have led to an expansion of credit to farmers that in turn led to faster adoption of high yielding varieties during the green revolution. The argument that farmers who are credit constrained find it difficult to adopt new agricultural technologies is well documented in the literature (Bhalla, 1979; Pitt and Sumodiningrat, 1991; Lipton, 1976). Often, introducing a high yielding variety of seeds or purchasing a tractor requires having access to loans because farmers simply do not have adequate savings that are required to make such investments on their own. Access to credit then acts as a supplement to savings that can be used to invest in agricultural 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. We, therefore, postulate that money-lending expertise was another potential pathway through which the refugees could have influenced post green revolution agricultural development in India.

5.2   Refugees and Land Reforms

5.2.1   Land Redistribution

As discussed in Section 1, the arrival of refugees was associated with land reforms that were crucial to the adoption of high yielding varieties during the green revolution. The first of these reforms was the substantial redistribution of land from large landholders to

---

[29]Report on the Census of the Panjab Taken on the 17th of February 1881. Page 125-138.

[30]Gazetteers, Punjab District. Gazetteer of the Muzaffargarh District, 1929. Page 78.

[31]Gazetteers, Punjab District. Gazetteer of the Jhang District, 1883-84. (1884). Page 68.

[32]Gazetteers, Punjab District. Gazetteer of the Montgomery District, 1883-84. (1884). Page 69-70

[33]Gazetteers, Punjab District. Gazetteer of the Attock District, 1930. Page 115.

small and medium sized owner-cultivators that took place amongst the refugee population. One extreme example of redistribution occurred when one of the large landowners amongst the refugees submitted a claim for 11,582 acres based on her previous land holdings in Pakistan. In return, she received an allotment of only 835 acres, which represented a 93 percent reduction in her previous land holdings (Kudaisya, 1995, p. 82). In addition to dramatically reducing large landholdings the redistribution also increased small and medium sized landholdings amongst the refugee population. It did so by allotting some of the poorest refugee farmers small parcels of land, known as 'plough units', for cultivation regardless of their holdings in Pakistan (Randhawa, 1954, p. 67).

The redistribution of land amongst the refugee population had important repercussions for the long term agricultural development in the high-refugee districts. To begin with, land redistribution increased the size of those landowning groups (i.e. small and medium owner-cultivators) that were most likely to undertake the lumpy and/or risky investments required for improving agricultural productivity.[34] According to Kudaisya (1995) it was the superior risk taking ability of the refugee farmers that "enabled them to make significant changes in their methods of irrigation and farming" that were required for adopting HYVs during the green revolution.

The redistribution also affected farming incentives within the refugee community in a way that was beneficial to agricultural development. By forcing large landowners to take up cultivation themselves rather than live as absentee landlords and by giving land to landless labourers it increased the incidence of owner-cultivation amongst the refugee population. Since owner-cultivators have superior incentives relative to tenant farmers and landless labourers,[35] the redistribution of land contributed to agricultural technology adoption and productivity through through the channel of incentives. Indeed, Kudaisya (1995) credits the process through which evacuee land was redistributed to refugees with creating a system of peasant-proprietorships through which agriculturists in general were encouraged to 'work hard and stand on their own feet'.

---

[34]The link between small and medium sized owner cultivators and intensity of agricultural technology adoption is well established in the literature. For example Muthiah (1971), Schluter (1971) and Sharma (1973) all find that the proportion of acreage devoted to high yielding varieties on small and medium sized farms in India is higher than the corresponding number for large farms.

[35]The argument that cultivating other people's land rather than your own leads to incentive problems that reduce investment and productivity is well established in the literature (See Banerjee and Iyer (2005) for reference).

5.2.2 Land Consolidation

Another important change in the land regime associated with the arrival of refugees was the consolidation of fragmented land. The need to consolidate fragmented land for improving agricultural productivity had been officially recognised since the enactment of the Punjab Consolidation Act in 1936 during the colonial period. However, progress in the field of land consolidation had been slow up until Partition. The large number of property transfers involved in redistributing evacuee land to refugees provided the required impetus behind accelerating the process of land consolidation shortly after Partition. Consequently, efforts towards land consolidation picked up pace through the enactment of the East Punjab Consolidation and Fragmentation Act in 1948. The legislation imposed 'restrictions on the creation of fragments as a preventive measure of consolidation' (Sarkar, 1989, p. 84). For instance, one of the clauses of the Act prohibited any transfer or partitioning of land that could result in the creation of a fragment (Sarkar, 1989, p. 84). Another clause barred the practice of fragmenting undivided agricultural land for distribution amongst inheritors. In addition to the Act of 1948 a series of further laws were passed for promoting the consolidation of landholdings (Gill, 1989, p. 80). A separate department was even opened by the East Punjab government for the sole purpose of enforcing land consolidation (Gill, 1989, p. 80). As a result of legislation and other proactive measures taken by the government, the size of consolidated land holdings went up from 700,000 acres in 1948 to over 22.084 million acres in 1966 in East Punjab as a whole (Kudaisya, 1995, p. 83). Most of consolidation took place in districts where the problem of land fragmentation had become especially acute as a result of refugees settling in large numbers.

The increase in consolidated land in high refugee districts could have influenced the adoption of HYVs during the Green Revolution in several ways. First, large areas of wasteland, such as embankments and field boundaries, under the previous fragmented land regime were brought under cultivation through consolidation. The increase in plot size that accompanied such consolidation could have led to farmers being more willing to experiment with planting the new, more risky, HYVs on a fraction of their cultivated land. Second, farmers with consolidated landholdings could have found it economically viable to sink tube-wells in their farms to gain access to sub-soil water for cultivation. It was largely through land consolidation that the number of tube-wells in East Punjab rose from close to zero in 1950 to 570,000 in 1978 (Kudaisya, 1995, p. 83). The increased access to sub-soil water that farmers were able to secure through consolidation was crucial to the

adoption of high yielding varieties that depended on the timely and controlled provision of water. Finally, land consolidation could also have led to improved access to canal irrigation since supplying canal-irrigated water to tiny scattered plots under the previous fragmented land regime had been practically impossible (Kudaisya, 1995, p. 84). Just as was the case with sub-soil water, the access to canal irrigated water proved vital to the adoption of high yielding varieties during the green revolution.

6 CONCLUSION

In this study, we examine the impact of the partition of India on the take up of agricultural technology and agricultural productivity post-partition. Using refugee presence as a proxy for the intensity of the impact of partition, we find that areas with more refugees 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. In particular, we show that more affected districts diverge from the less affected districts after the start of the Green Revolution in the late 1960s. We further show, using pre-partition agricultural data, that the effects are not solely explained by selective refugee movement into districts with a higher potential for agricultural development. While limited in our ability to explore empirically the potential pathways through which the partition impacted long run agricultural development, we nonetheless argue that the greater levels of education of the refugees and their higher concentration in money-lending contributed to agricultural development. Additionally, we also provide qualitative evidence for the role of land reforms soon after partition, particularly in areas with refugees, as a factor in explaining higher agricultural productivity.

While our study makes progress on an important question about the legacy of India's partition, there are several aspects that we hope future work can address. Prime among these is the effects of the partition in the long run on the Pakistani side. While data limitations prevent us from doing a similar analysis for agriculture, perhaps other relevant outcomes such as literacy and urbanization could be analyzed. Future work on land reforms after the partition would lend fascinating insights into why places affected by refugee settlement evolved differently, as would work on other mechanisms (political, social capital, religious homogeneity) that may explain the patterns seen in this paper.

A Appendix

Figure A.1: IACD to VDSA Wheat Yields Comparison (1966-87).
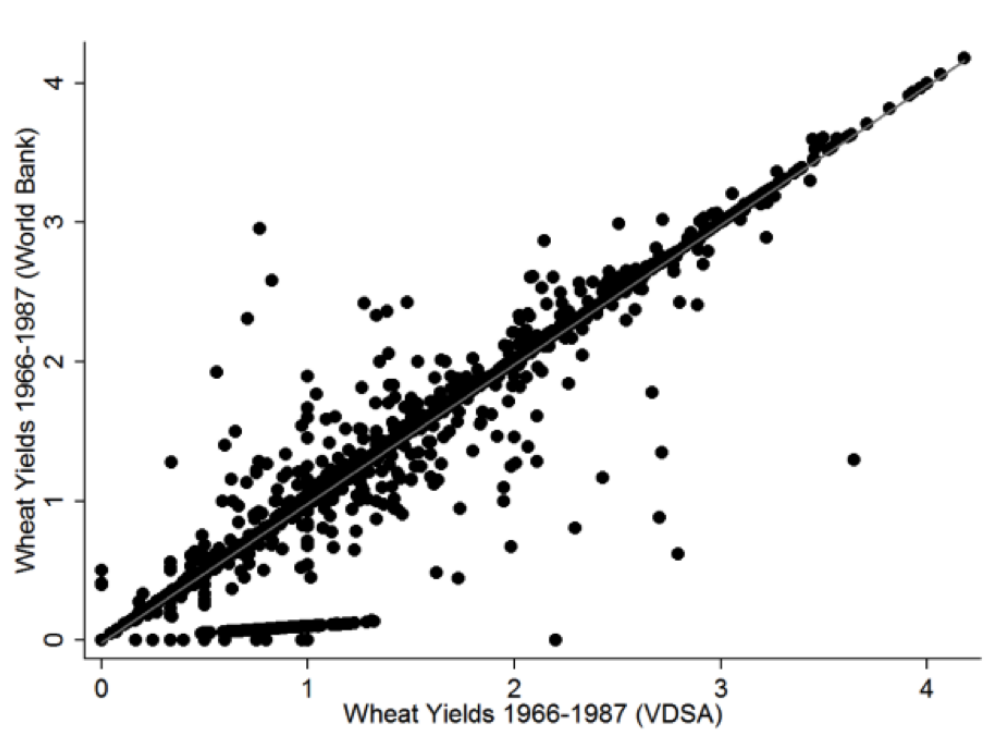
Notes: There were 19 cases in which for the same year and district the annual wheat yield was zero in the World Bank dataset but was non-zero in the VDSA dataset. 68% of these cases came from the Andhra Pradesh state and 32% came from the Karnataka state. On the other hand there were 5 cases in which 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).
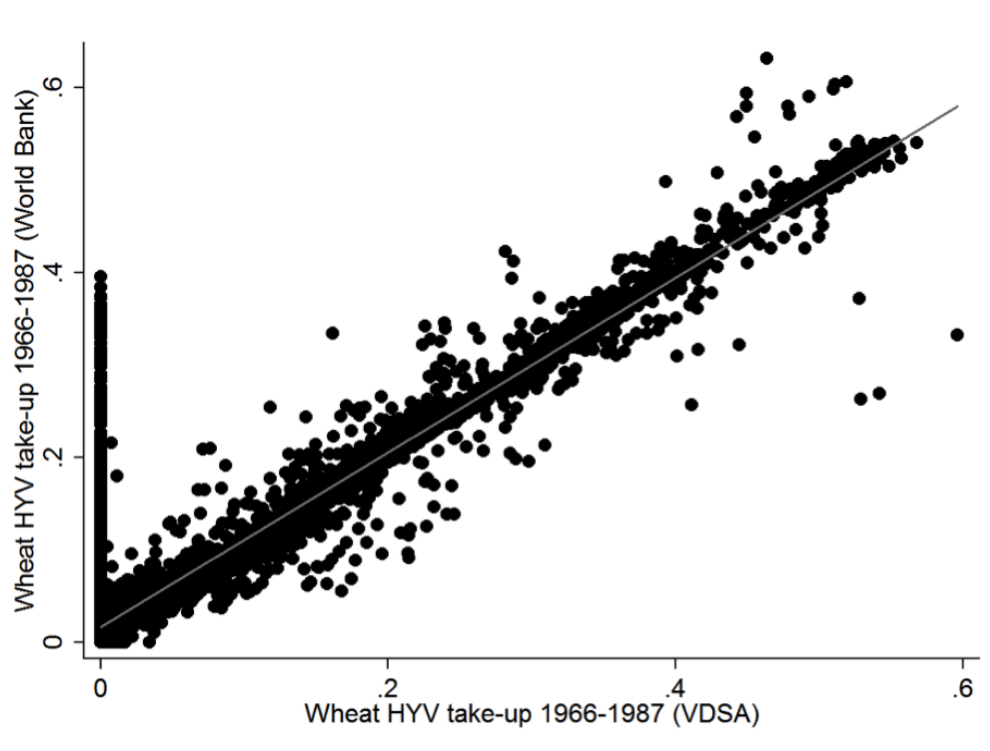
Notes: There were 22 cases in which for the same year and district the annual fraction of HYV of wheat was zero in the World Bank dataset but was non-zero in the VDSA dataset. 82% of these cases came from the Maharashtra state, 9% came from the Gujarat state, 4.5% from Rajasthan state and 4.5% from Tamil 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.

B Appendix

West Bengal Refugee Resettlement: Rather than coming in one big rush around the time of Partition, as was the case with the East Punjabi refugees, the West Bengali refugees instead came in waves over a prolonged period. The prolonged nature of the migration from East Pakistan to West Bengal was due to the gradual deterioration in the security situation of the minorities in East Pakistan after Partition. Initially, the hindu minorities felt safe enough to remain behind in East Pakistan. However, after 1950 when tensions between the muslim and hindu communities began to rise, many of the hindus who had originally decided to remain behind in East Pakistan at the time of Partition decided to leave for West Bengal in India.

Unlike its East Punjabi counterpart, the government of West Bengal failed to resettle refugees on agricultural land. Part of the reason behind this failure was the deliberate policy of both the Indian and Pakistani governments to avoid a population exchange between West Bengal and East Pakistan. A clear reflection of the policy was the Nehru-Liaqat pact of 1950. The Pact aimed at creating a sense of security amongst minorities in both East and West Bengal to avoid population exchange across the Partition border (Kudaisya, 1995, p. 87). In accordance with its policy of avoiding population exchange, the West Bengal government did not make any detailed plans regarding the distribution of land to refugees. Consequently, many of the refugees ended up becoming landless and congregated in areas around Calcutta city, forming a vast urban underclass (Kudaisya, 1995, p. 90). Another factor behind the West Bengal government's failure to resettle refugees was the amount of evacuee land available after Partition. The number of muslim evacuees who left for East Pakistan was far less than the number of hindu refugees who replaced them in West Bengal (Kudaisya, 1995, p. 89). Moreover, the muslim evacuees who left for East Pakistan were mostly either landless peasants or small-scale farmers with meagre land holdings. Such a situation meant that the evacuee land available for resettling refugees was insufficient.

Up until 1950, most of the refugees who arrived in West Bengal "came from the middle classes and relied on their own resources to begin life afresh" in their new homeland (Kudaisya, 1995, p. 89). However, when fresh riots broke out in East Bengal in early 1950, many hindu agriculturists were forced to migrate to West Bengal. Such a change in the economic profile of the incoming refugees caused a fundamental shift in the nature of the rehabilitation problem faced by the West Bengal government. The post-1950

refugees came with "few possessions, capital or skills" and urgently sought resettlement on agricultural land (Kudaisya, 1995, p. 87). The government, however, was slow in responding to the refugee resettlement problem. As late as 1958, almost eleven years after Partition, the West Bengal government had "been able to acquire only 61,000 acres of agricultural land for redistribution among the refugees, which was hardly adequate considering the numbers involved" (Kudaisya, 1995, p. 89).

Eventually, the official policy on refugees shifted from ignoring the land resettlement needs of the refugees to actively resettling refugees on agricultural land. The policy shift manifested itself in the form of the Dandakaranya scheme. The scheme represented the West Bengal government's attempt at finding 'a long-term solution for the resettlement of the East Bengali refugees' (Kudaisya, 1995, p. 91). According to the scheme, refugees were allotted land in the Dandakaranya region that comprised the districts of Bastar in Madhya Pradesh State and Kalahandi and Koraput in Orissa State. Initially, the scheme looked promising with around 12,000 refugee families being allotted land in the Dandakaranya region by 1967. However, a major drawback of the scheme was that the agricultural conditions in Dandakaranya were unsuitable for paddy – a crop the refugees specialized in cultivating. Dandakaranya was a forested plateau where the soil was hard, red, and porous and, as a result, the potential for surface irrigation was limited (Kudaisya, 1995, p. 91). All these factors severely limit the potential for paddy cultivation. Once the refugees realized that the land they had been allotted was severely limited in terms of its potential for large-scale paddy cultivation, they abandoned their holdings and left. Between 1972 and 1978, some 10,923 refugee families deserted their land holdings in the region. The Dandakaranya scheme had completely failed by the late 1970s.

The failure of the Dandakaranya scheme and the lack of state support for resettlement led to some refugees creating their own settlements by clearing forests and colonising agricultural lands in regions of Eastern India like the Sunderbans, Assam valley, and Tripura. Despite such independent efforts, the bulk of the refugee community remained landless throughout the post-partition period (Kudaisya, 1995, p. 86).

C Appendix

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 (in rupees). The quantity and price of the various inputs used in agriculture such as bullocks, tractors, and fertilizer 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 decadal population census reports from 1951 to 1981 is also available on the number of persons, literacy, number of cultivators and the number of agriculture labourers. Finally, there is a set of 21 indicator variables, each specifying a different soil quality type in the dataset.

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.

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 data from the decadal population census reports from 1961 to 2001 on the number of persons, literacy, number of cultivators and the number of agricultural labourers.

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 persons were enumerated, of whom 4.7 and 2.55 million had come from West and East Pakistan, respectively, and 0.05 million did not specify their place of origin (Visaria, 1969). Information on refugee inflows was disaggregated by gender, age, occupation and region of origin. In the case of sex, separate inflows were recorded for both males and females[36]. 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 refugee was identified as being either West or East Pakistan. In addition to demographic characteristics, there was also data on the occupation of refugees. Appendix II of Table IV in the census provides a detailed occupational classification of the refugees[37].

The 1951 census provides the best estimate to date of the spatial distribution of the partition related refugees who moved from Pakistan to India. 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 a refugee came from. Secondly, substantial changes in the administrative machinery and the relatively unsettled conditions in those districts that received refugees 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 refugees (Visaria, 1969). Finally, the high mortality rate amongst the refugees who arrived between

---

[36]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.

[37]Bharadwaj, Khwaja, and Mian (2009) find that the refugees tended to engage more in non-agricultural professions relative to the resident population.

1947 and 1951 meant that the true scale of partition related displacement could not be established (Visaria, 1969).

Combining IACD and VDSA datasets: In constructing our post-partition panel for the full sample of districts 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 3b 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.

Lack of Data in Pakistan: In the case of Pakistan, post-partition data on agricultural yields for the major crops is only available annually from the 1980s onwards and does not cover the crucial initial phases of the Green Revolution. Moreover, no data on yields is available on the Baluchistan province, the Bahawalpur Division of the Punjab province and the Khairpur Division of the Sind province for the pre-partition period. All these regions were former princely states for whom no records exist in the Agricultural Statistics Reports of British India.

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.

BAZZI, S., A. GADUH, A. D. ROTHENBERG, AND M. WONG (2016): "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia," *American Economic Review*, 106(9), 2658–98.

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. KHWAJA, AND A. MIAN (2008): "The big march: migratory flows after the partition of India," *Economic and Political Weekly*, pp. 39–49.

BHARADWAJ, P., A. I. KHWAJA, AND A. MIAN (2015): "Population Exchange and its Impact on Literacy, Occupation and Gender–Evidence from the Partition of India," *International Migration*, 53(4), 90–106.

BHARADWAJ, P., A. I. KHWAJA, AND A. R. MIAN (2009): "The partition of India: demographic consequences," .

BLEAKLEY, H., AND J. FERRIE (2016): "Shocking behavior: Random wealth in antebellum Georgia and human capital across generations," *The quarterly journal of economics*, 131(3), 1455–1495.

BYERLEE, D., AND P. MOYA (1993): *Impacts of international wheat breeding research in the developing world, 1966-1990*. CIMMYT.

CHAKRAVARTI, A. (1973): "Green revolution in India," *Annals of the Association of American Geographers*, 63(3), 319–330.

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.

D'AGOSTINO, A. (2017): "Technical Change and Gender Wage Inequality: Long-Run Effects of India?s Green Revolution," *Working Paper.*

DASGUPTA, A. (2018): "Technological Change and Political Turnover: The Democratizing Effects of the Green Revolution in India," *American Political Science Review*, 112(4), 918–938.

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.

EVENSON, R. E., AND D. GOLLIN (2003): "Assessing the impact of the Green Revolution, 1960 to 2000," *Science*, 300(5620), 758–762.

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.*

GILL, S. S. (1989): "Changing land relations in Punjab and implications for land reforms," *Economic and Political Weekly*, pp. A79–A85.

GOLLIN, D., C. W. HANSEN, AND A. M. WINGENDER (2016): "Two Blades of Grass: The Impact of the Green Revolution," *Centre for Economic Policy Research Discussion Paper DP11611.*

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. 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.

KHAN, Y. (2017): *The great partition: The making of India and Pakistan.* Yale University Press.

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

KUDAISYA, G. (1995): "The demographic upheaval of partition: Refugees and agricultural resettlement in India, 1947–67," *South Asia: Journal of South Asian Studies*, 18(s1), 73–94.

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.

MOHAN, R., AND R. E. EVENSON (1975): "The intensive agricultural districts programme in India: A new evaluation," *The Journal of Development Studies*, 11(3), 135–154.

MUNSHI, K. (2004): "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of development Economics*, 73(1), 185–213.

MUTHIAH, C. (1971): "The Green Revolution-Participation by Small and Large Farmers," *Indian Journal of Agricultural Economics*, 26(1), 53.

NCAI-I (1976): "Report of the National Commission on Agriculture, 1976," .

NCAI-V (1976): "Report of the National Commission on Agriculture, 1976," .

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.

PETERS, M. (2017): "Refugees and Local Agglomeration: Evidence from Germany's Post-War Population Expulsions," *Yale University Working Paper*.

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.

RANDHAWA, M. S. (1954): *Out of the ashes: An account of the rehabilitation of refugees from West Pakistan in rural areas of East Punjab*. Public Relations Dept., Punjab.

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

ROCHA, R., C. FERRAZ, AND R. R. SOARES (2017): "Human capital persistence and development," *American Economic Journal: Applied Economics*, 9(4), 105–36.

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

RUD, J. P. (2012): "Electricity provision and industrial development: Evidence from India," *Journal of development Economics*, 97(2), 352–367.

SARKAR, B. (1989): *Land Reforms in India, Theory and Practice: A Study of Legal Aspects of Land Reforms Measures in West Bengal.* APH Publishing.

SCHLUTER, M. (1971): *Differential Rates of Adoption of the New Seed Varieties in India: The Problem of the Small Farm.* Department of Agricultural Economics, Cornell University.

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

SHARMA, A. (1973): "Influence of certain economic and technological factors on the distribution of cropped area under various crops in the Ludhiana district," *Journal of Research (Punjab Agricultural University)*, 10, 243–49.

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.

WANTCHEKON, L., M. KLAŠNJA, AND N. NOVTA (2014): "Education and human capital externalities: evidence from colonial Benin," *The Quarterly Journal of Economics*, 130(2), 703–757.

YONG, T. T., AND G. KUDAISYA (2000): *The Aftermath of Partition in South Asia.* London, Routledge.

Table 1A. Summary Statistics

Variable	Mean	Median	Standard Deviation	Number of Observations
Annual Wheat Yields (1957-2009) (tons per hectare)	1.489	1.292	0.893	12763
Annual Take-Up of HYVs of Wheat (1957-2009) (proportion)	0.070	0.005	0.128	11573
Annual Revenue per hectare based on 1960 prices (rupees per hectare)	493.391	386.945	356.939	11500
Annual consumption of Nitrogen Fertilizer (tons)	40.241	16.262	118.801	13001
Annual consumption of Phosphorus Fertilizer (tons)	15.812	5.906	50.427	13001
Annual consumption of Potassium Fertilizer (tons)	10.257	1.258	99.389	13001
Annual number of tractors per 1000 hectares (1957-1987)	1.811	0.395	4.896	8370
Log Refugees (1951)	7.443	7.549	2.576	270
Log of Population (1951)	13.792	13.846	0.639	270
High Refugee Dummy (1951)	0.485	…	…	270
Proportion Refugees (1951)	0.021	0.002	0.055	269
Annual Rainfall (1957-2009) (millimeters)	1039.807	945.000	573.994	13971
Population Desnity (1961)	1.789	1.251	2.117	270

Notes: Data on annual rainfall does not change between 1957 and 1965. This is because the Indian Agriculture and Climate Dataset from which the data between 1957 and 1965 is taken only provides estimated average rainfall for each district for each month across all the years between 1957 and 1987. Data for the annual number of tractors from 1988 to 2009 does not exist. This because the Village Dynamics in South Asia Dataset from which the data from between 1957 to 2009 was supposed to be taken does not include information on tractors.

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

Variable	Full Sample (1957-2009)					Restricted Sample (1911-2009)					Difference[a]	p-value of difference
	Actual	Mean	Median	Std. Dev	Observations	Actual	Mean	Median	Std. Dev	Observations		
Number of Refugees (1951)	5862319	…	…	…	265	3329178 **(57%)**	…	…	…	67	…	…
Population (1951)	302980160	…	…	…	265	94666560 **(31%)**	…	…	…	67	…	…
Annual Wheat Yields (lbs per acre)	…	1180.53	1024.25	710.97	12215	…	1351.18	1116	843.45	3785	171	0.000
Annual Rice Yields (lbs per acre)	…	1076.82	888.98	659.87	12769	…	880.52	366	1469.35	3371	196	0.000
Annual Maize Yields (lbs per acre)	…	1067.24	873.31	730.65	12144	…	1182.10	1057	660.28	3441	115	0.000
Annual Sorghum Yields (lbs per acre)	…	568.13	528.48	351.176	11445	…	653.76	634.00	386.61	2918	86	0.000
Annual Pearl Millet Yields (lbs per acre)	…	577.88	503.54	370.69	8074	…	661.02	587.00	418.13	2367	83	0.000
Altitude (meters)	…	353.35	357.00	140.24	13577	…	411.08	396.00	154.80	3785	58	0.000
Latitude (degrees)	…	22.75	23.36	4.82	13577	…	24.58	25.00	3.93	3785	2	0.000
Longitude (degrees)	…	78.80	78.10	4.38	13577	…	79.37	78.00	4.51	3785	1	0.000

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.

[a]Difference represents the absolute difference between the full and restricted sample means.

Table 2: Wheat Yields and Refugees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Wheat Yields				
Post Partition X Refugee Measure	32.3	32.24			150.1**	150.0**		
	(21.82)	(23.33)			(60.99)	(65.21)		
Post GR X Refugee Measure			53.10**				196.3***	
			(25.69)				(72.67)	
1957-1967 X Refugee Measure				-6.926				2.094
				(15.78)				(44.99)
1967-1977 X Refugee Measure				8.937				67.05
				(22.20)				(57.10)
1977-1987 X Refugee Measure				51.11				183.5**
				(31.00)				(81.00)
1987-1997 X Refugee Measure				63.16*				251.4***
				(33.19)				(93.12)
1997-2009 X Refugee Measure				42.26				224.1**
				(34.76)				(106.0)
Refugee Measure:		Log Refugees				High Refugee Dummy		
Mean outcome	1351	1351	1351	1351	1351	1351	1351	1351
Observations	3785	3785	3785	3785	3785	3785	3785	3785
R-squared	0.893	0.925	0.926	0.926	0.914	0.946	0.948	0.927
Controls:								
Main effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State polynomial trends	Yes	No	No	No	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Table 3: Refugee presence and other crops

	(1) Sugarcane	(2) Maize	(3) Sorghum	(4) Pearl millets	(5) Other millets	(6) Chickpea	(7) Rapeseed	(8) Linseed
Post Partition X High Refugee Dummy	360.8**	48.19	-93.74	20.05	-122.2	53.25	170.9	1.538
	(164.4)	(121.6)	(85.68)	(63.38)	(103.7)	(50.65)	(104.0)	(18.34)
Mean outcome	4446	1182	654	661	664	646	561	329
Observations	3,656	3,441	2,918	2,367	1,130	3,108	2,948	1,786
R-squared	0.530	0.492	0.329	0.418	0.188	0.055	0.524	0.480
	Sugarcane	Maize	Sorghum	Pearl millets	Other millets	Chickpea	Rapeseed	Linseed
Post GR X High Refugee Dummy	361.7	-33.51	23.71	8.086	-48.13	33.88	50.98	59.31**
	(243.1)	(69.47)	(52.79)	(47.11)	(98.07)	(42.90)	(37.12)	(27.88)
Mean outcome	4446	1182	654	661	664	646	561	329
Observations	3,656	3,441	2,918	2,367	1,130	3,108	2,948	1,786
R-squared	0.533	0.497	0.325	0.412	0.187	0.055	0.516	0.482
Controls:								
Main effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State linear trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Table 4: Wheat Yields and Other proxies

	(1)	(2)	(3)	(4)	(5)	(6)
			Wheat Yields			
Post Partition X Other Proxy	382.45	….	356.73	….	1222	….
	(241.55)	….	(1577.24)	….	(5296.65)	….
Post GR X Other Proxy	….	660.04***	….	1881.01***	….	-4149.872
	….	(161.19)	….	(516.35)	….	(3153.58)
Post Partition X Inverse Distance to the border	….	….	….	….	-1647.804	….
	….	….	….	….	(2078.64)	….
Post Partition X Proportion Muslim in 1931	….	….	….	….	244.78	….
	….	….	….	….	(1127.84)	….
Post GR X Inverse Distance to the border	….	….	….	….	….	2038.69
	….	….	….	….	….	(1342.30)
Post GR X Proportion Muslim 1931	….	….	….	….	….	1377.41**
	….	….	….	….	….	(678.20)
Proportion Muslim 1931 X Inverse Distance to the border	….	….	….	….	-1438.225	2440.71
	….	….	….	….	(4092.30)	(3658.05)
Other Proxy:	Proportion Muslim 1931		Proportion Muslim 1931 X Inverse Distance to the border			
Mean outcome	1351	1351	1351	1351	1351	1351
Observations	3,785	3,785	3,785	3,785	3,785	3,785
R-squared	0.882	0.872	0.883	0.873	0.883	0.874
Controls:						
Main effects	Yes	Yes	Yes	Yes	Yes	Yes
All double interactions	N/A	N/A	No	No	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
State linear trends	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Table 5: Wheat Yields and Refugees accounting for soil conditions and aquifer depth after partition

	(1)	(2)	(3)	(4)
		Wheat Yields		
Post Partition X High Refugee Dummy	154.1**	154.0**		
	(59.84)	(63.99)		
Post GR X High Refugee Dummy			197.0***	
			(70.74)	
1957-1967 X High Refugee Dummy				-0.578
				(45.52)
1967-1977 X High Refugee Dummy				57.81
				(56.42)
1977-1987 X High Refugee Dummy				184.1**
				(77.11)
1987-1997 X High Refugee Dummy				236.7**
				(91.03)
1997-2009 X High Refugee Dummy				200.3*
				(102.6)
Mean outcome	1351	1351	1351	1351
Observations	3,785	3,785	3,785	3,785
R-squared	0.896	0.928	0.930	0.929
Main effects	Yes	Yes	Yes	Yes
State FE	Yes	No	No	No
State polynomial trends	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes
Aquifer depth	Yes	Yes	Yes	Yes
Aquifer depth X Post Partition	Yes	Yes	No	No
Aquifer depth X Post GR	No	No	Yes	No
Aquifer depth X Decade FE	No	No	No	Yes
Geographic controls	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Table 6: Proxies and pre partition/green revolution infrastructure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Area sown with canal irrigation in 1938	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 in 1966	Total geographic area of district	Growth in literate population 1911-1921	Aquifer Depth 1	Aquifer Depth 2	Aquifer Depth 3
High Refugee Dummy	0.025	3.82E-06	-0.0001482***	-3.88E-06	-4.54E-07	536.142*	-56.179	0.044	-0.041	0.161	0.011
	(0.190)	(0.00000431)	(0.0000556)	(0.0000246)	(0.00000118)	(288.654)	(89.957)	(0.065)	(0.082)	(0.097)	(0.046)
Proportion of Muslims in 1931	0.374	-0.0000357	-0.0002056	-0.0001149	1.06E-06	5951.346***	197.667	-0.279	0.018	-0.236	0.276
	(0.954)	(0.0000278)	(0.0002409)	(0.0001137)	(0.00000454)	(1799.139)	(339.894)	(0.283)	(0.365)	(0.537)	(0.371)
Inverse distance to the border	8.656	0.00003	-0.0006708	0.0000561	-8.07E-06	2973.371	652.567	-0.375	-2.022	-2.566	2.905
	(5.607)	(0.0001394)	(0.0005547)	(0.0003821)	(0.000031)	(3276.293)	(849.875)	(0.818)	(1.387)	(2.415)	(1.854)
Prop Muslim X Inverse distance to the border	8.130	-0.0000777	-0.001274*	-0.0006656	-3.10E-06	16932.324**	917.754	-1.196	-0.860	-2.492	2.949
	(5.180)	(0.0001846)	(0.0007525)	(0.000459)	(0.000033)	(8037.000)	(1071.041)	(0.903)	(0.903)	(2.322)	(1.886)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	197	61	62	45	60	65	67	67	67	67	67

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.

Table 7: Wheat Yields and Refugees - accounting for Banerjee & Iyer (2005) districts

	(1)	(2)
	Wheat Yields	
Post Partition X High Refugee Dummy	120.8**	120.8*
	(58.01)	(62.14)
Banerjee-Iyer District X Post	188.2***	188.2***
	(65.44)	(70.11)
Mean outcome	1361	1361
Observations	3,690	3,690
R-squared	0.896	0.928
Controls:		
Main effects	Yes	Yes
State polynomial trends	Yes	No
State X Year FE	No	Yes
Geographic controls	Yes	Yes

Robust standard errors in parentheses, clustered at the district level. Tables notes are the same as in Table 2, with the addition of districts level data from Banerjee and Iyer (2005), which classify districts based on various types of colonial tax regimes.

*** p<0.01, ** p<0.05, * p<0.1

	Annual revenue (in rupees) per hectare based on 1960 prices		
	(1)	(2)	(3)
High Refugee Dummy	51.34**	62.72***	61.65***
	(22.50)	(20.93)	(20.03)
Mean outcome	485.64	484.68	478.28
Observations	11,500	11,469	11,312
R-squared	0.606	0.675	0.691
Controls			
Soil type dummies	No	Yes	Yes
Population density (1961)	No	No	Yes
Annual rainfall	No	No	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes

Notes: * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Table 9: Revenue per hectare based on 1960 prices, Refugees and the Green Revolution

	Annual revenue (in rupees) per hectare based on 1960 prices		
	(1)	(2)	(3)
High Refugee Dummy	31.58**	42.49***	43.01***
	(15.60)	(16.16)	(15.92)
Post GR X High Refugee Dummy	33.72**	34.14**	31.84**
	(16.86)	(15.31)	(14.87)
Mean outcome	485.64	484.68	478.28
Observations	11,500	11,469	11,312
R-squared	0.607	0.676	0.691
Controls			
Soil type dummies	No	Yes	Yes
Population density (1961)	No	No	Yes
Annual rainfall	No	No	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes

Notes: All Columns also include 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% at the state level and 0 otherwise). * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Table 10: Wheat yields/HYV take up, Refugees and the Green Revolution

	Annual wheat yields		Take up of wheat HYV	
	(1)	(2)	(3)	(4)
High Refugee Dummy	0.138***	0.0212	0.00714*	-0.00768*
	(0.0407)	(0.0288)	(0.00381)	(0.00408)
Post GR X High Refugee Dummy		0.198***		0.0287***
		(0.0424)		(0.00807)
Mean outcome	1.46	1.46	0.07	0.07
Observations	12215	12215	11295	11295
R-squared	0.788	0.791	0.806	0.819
Controls				
Soil type dummies	Yes	Yes	Yes	Yes
Population density (1961)	Yes	Yes	Yes	Yes
Annual rainfall	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes	Yes

Notes: The unit for yields is tons per hectare. Columns 2 and 4 also include 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% at the state level and 0 otherwise). * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

	Tractors per 1000 hectare		Consumption of nitrogen fertilizer per 1000 hectares		Consumption of phosphorus fertilizer per 1000 hectares	
	(1)	(2)	(3)	(4)	(5)	(6)
High Refugee Dummy	1.128***	0.477***	16.42	0.959	8.304*	1.424
	(0.298)	(0.176)	(10.70)	(3.339)	(4.998)	(1.589)
Post GR X High Refugee Dummy		2.009***		27.89		12.39*
		(0.598)		(18.68)		(6.968)
Mean outcome	2.01	2.01	33.14	33.14	13.09	13.09
Observations	8339	8339	12562	12562	12562	12562
R-squared	0.597	0.604	0.333	0.336	0.284	0.287
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
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The unit for yields is tons per hectare. Columns 2, 4 and 6 also include 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% at the state level and 0 otherwise). The use of tractors is measured as the number of tractors per thousand hecatares of cultivated land. Tractor data is only avaiable in the IACD, and hence available from 1958-1987. Data on fertilizers is in both the VDSA and IACD and is available from 1957-2009. The consumption of Nitrogen and Phosphorus fertilizers is measured in tons per thousand hectares of cultivated land. * p < .10, ** p < .05, *** p < .01. Clustered standard errors in parentheses.

Appendix Table 1: Wheat Yields and Refugees - Excluding IADP districts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Wheat Yields				
Post Partition X Refugee Measure	28.27	28.21			133.5**	133.4**		
	(19.98)	(21.46)			(59.25)	(63.65)		
Post GR X Refugee Measure			48.31**				186.4**	
			(24.09)				(70.95)	
1957-1967 X Refugee Measure				-7.457				-8.465
				(15.60)				(45.00)
1967-1977 X Refugee Measure				5.300				45.79
				(20.75)				(56.44)
1977-1987 X Refugee Measure				44.17				148.4*
				(26.55)				(75.15)
1987-1997 X Refugee Measure				59.50*				239.3**
				(31.91)				(93.88)
1997-2009 X Refugee Measure				37.18				217.2**
				(33.44)				(103.7)
Refugee Measure:		Log Refugees				High Refugee Dummy		
Mean outcome	1307	1307	1307	1307	1307	1307	1307	1307
Observations	3,557	3,557	3,557	3,557	3,557	3,557	3,557	3,557
R-squared	0.892	0.927	0.928	0.928	0.891	0.926	0.928	0.928
Controls:								
Main effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State polynomial trends	Yes	No	No	No	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Appendix Table 2: No. of oridinary acres per standard acre and refugee presence

District	Area in Ordinary Acres	Area in Standard Acres	Ordinary Acres per Standard Acre
High Refugees			
Jullundur	363131	225471	1.61
Ludhiana	230261	117883	1.95
Hoshiarpur	372144	158546	2.35
Ferozepur	855490	512239	1.67
Amritsar	257857	181377	1.42
Gurdadpur	277309	189585	1.46
Ambala	256813	128212	2.00
Karnal	443775	250751	1.77
Kapurthala	182939	109151	1.68
Patiala	139375	77095	1.81
Simla	123	38	3.24
Mean	*307202*	*177304*	*1.91*
Low Refugees			
Hissar	770642	248993	3.10
Rohtak	150089	95013	1.58
Gurgaon	178380	76221	2.34
Barnala	60352	41884	1.44
Kohistan	28964	12832	2.26
Mohindergarh	41740	15225	2.74
Fatehgarh	75890	49346	1.54
Kangra	16524	4286	3.86
Bhatinda	159710	91792	1.74
Sangrur	153141	82042	1.87
Mean	*163543*	*71763*	*2.25*

Notes: The data comes from Appendix XII of the Punjab Land Resettlement Manual written by Sardar Tarlok Singh in 1952.

Appendix Table 3: Refugees, Land and Pre-Partition Yields (Punjab only)

	(1) Area in Ordinary Acres	(2) Area in Standard Acres	(3) Ordinary Acres per Standard Acre	(4) Growth in Wheat Yields (1911-1938)
High Refugee Dummy	118200.391	99527.472**	-0.394	…..
	(1.208)	(2.098)	(-1.389)	…..
Log Refugees	109184.215***	62548.149***	-0.259***	…..
	(3.911)	(4.766)	(-2.806)	…..
Ordinary Acres per Standard Acre	…..	…..	…..	0.028
	…..	…..	…..	(0.83)
Observations	21	21	21	13

Notes: t-statistic in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For columns 1 to 3, data is available for 21 Punjab districts: Amritsar, Karnal, Ambala, Jullundhur, Ludhiana, Rohtak, Simla, Gurgaon, Hissar, Gurdaspur, Ferozepore, Hoshiarpur, Kangra, Patiala, Kohistan, Barnala, Mohindergarh, Fatehgarh, Sangrur, Bhatinda and Kapurthala. For column 4, data on pre-partition yields is only available for the 13 Punjab districts: Amritsar, Karnal, Ambala, Jullundhur, Ludhiana, Rohtak, Simla, Gurgaon, Hissar, Gurdaspur, Ferozepore, Hoshiarpur and Kangra.

Appendix Table 4: Wheat Yields and Refugees (restricting sample to Punjab and UP only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Wheat Yields				
Post Partition X Refugee Measure	109.6**	109.6**			359.6***	359.6***		
	(40.30)	(42.76)			(110.5)	(117.2)		
Post GR X Refugee Measure			111.2**				315.8**	
			(43.57)				(126.7)	
1957-1967 X Refugee Measure				21.59				120.7
				(31.08)				(77.25)
1967-1977 X Refugee Measure				61.55				230.4**
				(40.75)				(99.39)
1977-1987 X Refugee Measure				119.9**				379.1**
				(53.19)				(138.0)
1987-1997 X Refugee Measure				181.1***				556.0***
				(55.10)				(153.6)
1997-2009 X Refugee Measure				151.5**				476.6**
				(61.57)				(182.5)
Refugee Measure:		Log Refugees				High Refugee Dummy		
Mean outcome	1875	1875	1875	1875	1875	1875	1875	1875
Observations	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R-squared	0.926	0.950	0.952	0.954	0.927	0.951	0.953	0.955
Controls:								
Main effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State polynomial trends	Yes	No	No	No	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at the district level

*** p<0.01, ** p<0.05, * p<0.1

The sample for the regressions in columns (1) to (8) has been restricted to the western India states of Punjab and Uttar Pradesh. The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Appendix Table 5: Wheat Yields and Refugees (dropping Punjab from the sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Wheat Yields				
Post Partition X Refugee Measure	30.34	30.27			147.3**	147.3**		
	(21.59)	(22.94)			(61.08)	(64.92)		
Post GR X Refugee Measure			51.43*				195.9***	
			(25.78)				(72.24)	
1957-1967 X Refugee Measure				-6.898				-0.791
				(15.58)				(44.95)
1967-1977 X Refugee Measure				7.081				64.20
				(21.70)				(56.87)
1977-1987 X Refugee Measure				48.55				180.6**
				(30.60)				(80.62)
1987-1997 X Refugee Measure				60.89*				248.6***
				(32.85)				(92.76)
1997-2009 X Refugee Measure				39.35				221.7**
				(34.63)				(105.2)
Refugee Measure:		Log Refugees				High Refugee Dummy		
Mean outcome								
Observations	3,215	3,215	3,215	3,215	3,215	3,215	3,215	3,215
R-squared	0.810	0.863	0.865	0.865	0.811	0.864	0.867	0.868
Controls:								
Main effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State polynomial trends	Yes	No	No	No	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

The sample for the regressions in columns (1) to (8) drop Punjab from the sample. The unit for yields is lbs per acre. Geographic controls include 21 dummies for district soil type, latitude, longitude and altitude.

Appendix Table 6: Wheat Yields and Refugees (including district fixed effects)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Wheat Yields				
Post Partition X Refugee Measure	29.68	29.67			144.6**	144.6**		
	(22.02)	(23.56)			(62.45)	(66.82)		
Post GR X Refugee Measure			53.40**				196.7***	
			(25.82)				(73.16)	
1957-1967 X Refugee Measure				-10.23				-4.657
				(16.12)				(45.80)
1967-1977 X Refugee Measure				5.653				60.34
				(22.37)				(58.44)
1977-1987 X Refugee Measure				47.83				176.7**
				(31.27)				(82.61)
1987-1997 X Refugee Measure				60.41*				245.5**
				(33.29)				(94.77)
1997-2009 X Refugee Measure				41.43				221.9**
				(34.93)				(107.0)
Refugee Measure:		Log Refugees				High Refugee Dummy		
Mean outcome	1351	1351	1351	1351	1351	1351	1351	1351
Observations	3,785	3,785	3,785	3,785	3,785	3,785	3,785	3,785
R-squared	0.914	0.946	0.947	0.947	0.914	0.946	0.948	0.948
Controls:								
Main effects [a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State polynomial trends	Yes	No	No	No	Yes	No	No	No
State X Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes

[a]Main effects do not include Log Refugees or High Refugee Dummy as these are time-inavariant district level variables that became

redundant after the inclusion of district FE. Robust standard errors in parentheses, clustered at the district level.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The unit for yields is lbs per acre.

Appendix Table 7: Average revenue per hectare based on 1960 prices (1957 to 2009) and Refugees

	Average Annual Revenue Per Hectare Based on 1960 prices (1957 to 2009)		
	(1)	(2)	(3)
High Refugee Dummy	84.51***	90.32***	89.08***
	(24.40)	(24.98)	(24.81)
Mean outcome	497.13	497.13	497.13
Observations	265	265	265
R-squared	0.5607	0.6554	0.6561
Controls			
State fixed effects	Yes	Yes	Yes
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 high refugee dummy. The unit in which revenue is measured is rupees. * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 8: Average Take-Up of HYVs of wheat (1957 to 2009) and Refugees

	Average Annual Take-Up of HYVs of Wheat (1957 to 2009)		
	(1)	(2)	(3)
High Refugee Dummy	0.00729	0.00829*	0.00821*
	(0.00538)	(0.00461)	(0.00462)
Mean outcome	0.07	1.07	2.07
Observations	271	271	271
R-squared	0.852	0.896	0.897
Controls			
State fixed effects	Yes	Yes	Yes
Soil type dummies	No	Yes	Yes
Population density (1961)	No	No	Yes

Notes: The unit for yields is tons per hectare. This table shows regressions of annual take-up of High Yielding Varieties (HYVs) of wheat averaged over the period 1957 to 2009 on the high refugee dummy. 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.* $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors in parentheses.

Appendix Table 9: Yields and Proportion Refugees

	Annual wheat yields		Annual rice yields	
	(1)	(2)	(3)	(4)
Proportion Refugees	0.351	-0.243	0.435	-0.523
	(0.308)	(0.251)	(0.542)	(0.465)
Post GR X Proportion Refugees		0.872***		1.389***
		(0.254)		(0.433)
Mean outcome	1.46	1.46	1.31	1.31
Observations	12,015	12,015	12,410	12,410
R-squared	0.784	0.785	0.702	0.703
Controls				
Soil type dummies	Yes	Yes	Yes	Yes
Population density (1961)	Yes	Yes	Yes	Yes
Annual rainfall	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes	Yes

Notes: The unit for yields is tons per hectare. Columns 2 and 4 also include 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% at the state level and 0 otherwise). * $p < .10$, ** $p < .05$, *** $p < .01$. Clustered standard errors (at the district level) in parentheses. Data used is a combination of IACD (1957 to 1965) and VDSA (1966-2009).

Appendix Table 10: Wheat yields/HYV take up, Refugees and the Green Revolution - accounting for overlapping data

	Annual wheat yields		Take up of wheat HYV	
	(1)	(2)	(3)	(4)
High Refugee Dummy	0.139***	0.0227	0.007	-0.010**
	(0.0408)	(0.0287)	(0.00407)	(0.004)
Post GR X High Refugee Dummy		0.196***		0.030***
		(0.0425)		(0.008)
Mean outcome	1.46	1.46	0.07	0.07
Observations	12191	12191	10815	10815
R-squared	0.788	0.791	0.838	0.845
Controls				
Soil type dummies	Yes	Yes	Yes	Yes
Population density (1961)	Yes	Yes	Yes	Yes
Annual rainfall	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes	Yes

Notes: The unit for yields is tons per hectare. Data used for the regressions is a combination of IACD (1957 to 1965) and VDSA (1966-2009). 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. Columns 2 and 4 also include 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% at the state level and 0 otherwise). * 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).

Appendix Table 11: Wheat yields/HYV take up, Refugees and the GR (pre-partition consistent wheat sample)

	Annual wheat yields		Take up of wheat HYV	
	(1)	(2)	(3)	(4)
High Refugee Dummy	0.0620	0.00942	0.0152***	-0.00325
	(0.0579)	(0.0498)	(0.00510)	(0.00986)
Post GR X High Refugee Dummy		0.0817		0.0290*
		(0.0804)		(0.0153)
Mean outcome	1.53	1.53	0.12	0.12
Observations	3409	3409	2747	2747
R-squared	0.894	0.895	0.864	0.865
Controls				
Soil type dummies	Yes	Yes	Yes	Yes
Population density (1961)	Yes	Yes	Yes	Yes
Annual rainfall	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
State-specific linear time trends	Yes	Yes	Yes	Yes

Notes: The unit for yields is tons per hectare. The 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. Columns 2 and 4 also include 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% at the state level and 0 otherwise). * 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).

Apppendix Table 12a: Pairwise correlations between Rural Male Literacy and Refugees

	Proportion of rural males who are literate				
	1961	1971	1981	1991	2001
High Refugee Dummy	0.1104*	0.1289**	0.1407**	0.1477**	-0.0475
	(0.0552)	(0.0232)	(0.0147)	(0.0092)	(0.4045)
Observations	302	310	300	310	310

Notes: This table shows the pairwise correlation coefficients between the proportion of rural males who are literate and the high refugee dummy for each census year in the period between 1961 and 2001. The 2001 census year numbers for rural male literacy are not reliable. * significant at the 10% level or better, ** significant at the 5% level or better, *** significant at the 1% level or better. Significance level in parentheses.

Appendix Table 12b: Pairwise correlations between Rural Male Literacy and HYV take-up in the same year

	Proportion of rural males who are literate				
	1961	1971	1981	1991	2001
HYV Adoption	N/A	0.2822***	0.1994***	0.3466***	0.0654
	N/A	(0.0000)	(0.0009)	(0.0000)	(0.5012)
Observations	N/A	267	272	189	108

Notes: This table shows the pairwise correlation coefficients between the proportion of rural males who are literate and the HYV take-up of all major crops for each census year in the period between 1961 and 2001. HYV adoption did not start to happen until after 1961 and therefore no numbers are reported for that year. The 2001 census year numbers for rural male literacy are not reliable. * significant at the 10% level or better, ** significant at the 5% level or better, *** significant at the 1% level or better. Significance level in parentheses.