The Big March: Migratory Flows after the Partition of India

PRASHANT BHARADWAJ, ASIM KHWAJA, ATIF MIAN

The Partition of India in 1947 along ostensibly religious lines into India, Pakistan, and what eventually became Bangladesh resulted in one of the largest and most rapid migrations in human history. In this paper district level census data from archives are compiled to quantify the scale of migratory flows across the subcontinent. We estimate total migratory inflows of 14.5 million and outflows of 17.9 million, implying 3.4 million "missing" people. The paper also uncovers a substantial degree of regional variability. Flows were much larger along the western border, higher in cities and areas close to the border, and dependent heavily on the size of the "minority" religious group. The migratory flows also display a "relative replacement effect" with in-migrants moving to places that saw greater outmigration.

Prashant Bharadwaj (*prashant.bharadwaj@yale.edu*) is with the Economic Department at Yale University; Asim Khwaja (*akhwaja@ksg.harvard.edu*) teaches public policy at Harvard University's Kennedy School of Government; Atif Mian (*amian@chicagogsb.edu*) specialises in financial markets at the University of Chicago, Graduate School of Business.

Involuntary migrations continue to play an important role in today's world, with wars and political strife forcing hundreds of thousands to leave their country. Whether it is Rwanda, Bosnia-Yugoslavia, or Israel, people are constantly faced with situations where they have no choice but to flee. The US Committee for Refugees and Immigrants estimates a total of 12 million refugees and an additional 21 million internally displaced people in the world [World Refugee Survey 2006]. Yet despite the large-scale and costly ramifications of these flows, our empirical understanding of even the very basic questions – such as the size and variability of these flows – remains limited. How many people moved? From where and to where? How did the flows differ across regions? Too many of these questions often remain unanswered.

1 Introduction

Unlike voluntary migrations – where individuals move by choice and not due to safety concerns – involuntary movements are harder to study because they are almost invariably driven and accompanied by extraordinary events such as wars, partition and ethnic/religious strife. They also often involve the movement of a large number of people in a very short span of time. These events make it all the more hard to gather basic demographic information, and even in their aftermath such data are hard to recall.

The Partition of India in August 1947 is one such example. Despite being one of the largest and most rapid migrations in human history with an estimated 14.5 million people migrating within four years, there is little analytic work that examines the nature or consequences of this rapid movement.[1]

However, at least in this case, the lack of quantitative data is not an issue. There is extensive and detailed data available both for the periods before Partition (during the British period), as well as after Partition. This therefore offers a unique opportunity for a more quantitative analysis.

A contribution of this work is to compile historical data sources in a manner that is amenable for empirical analysis and at a disaggregated enough level – the district. Doing so provides a more detailed picture of the migratory flows and allows for comparisons across time. This is a challenging task, particularly since administrative boundaries underwent substantial change after Partition. While this data will form the basis of a series of studies that ultimately examines the socio-economic consequences of the large flows, in this paper we will focus on the size and nature of the flows. In a related follow-up paper (BKM), we examine the demographic consequences of the flows.

Using the 1931 and 1951 population census data we find that by 1951, within four years after Partition, 14.5 million people had migrated into India, Pakistan, and what later became Bangladesh. While outflows are not directly reported, we use region specific population projections to estimate total outflows of 17.9 million people during the same period. This suggests there were 3.4 million people "missing" or unaccounted for during Partition.

While these numbers underscore how large and sudden involuntary flows can be, they hide substantial variation. Although both the western (between India and Pakistan) and eastern border (between India and Bangladesh) regions had large populations, migratory flows along the western border were almost three times as large. The flows on the western border were also substantial relative to the population: Pakistani Punjab saw 20.92 per cent of its population leave while by 1951, 25.51 per cent of its population was from across the border;[2] in Indian Punjab,[3] 29.78 per cent of the population left and 16.02 per cent of the population was migrant. In comparison, West Bengal (on the Indian side) saw only 6.31 per cent of its population leave to be replaced by migrants who constituted only 8.47 per cent of the population. On the Bangladeshi side, 6.5 per cent of the population left and 1.66 per cent of the population was migrant by 1951. Thus, while Partition was ostensibly along religious lines, those along the western border were much more likely to move, presumably due to greater perceived threats.

Variation in Flows

In addition to variation between the two borders, the disaggregated results show high variation in flows even across nearby districts both in absolute numbers and as a fraction of the district's population. For example, the districts of Nadia and Murshidabad in West Bengal (India) are right on the border, yet Nadia received almost 4,27,000 migrants while Murshidabad received only around 58,700. As we mention below, this is because more Muslims moved out of Nadia than from Murshidabad and migrants typically moved more to places where people had left from. The point though is that these differences suggest that migratory flows can be highly localised and even areas in close proximity can be faced with very different numbers.

Using district level variation also allows us to ask where migrants moved to and where they left from. This allows us to document that even involuntary migrations have a degree of predictability. Not surprisingly, distance to the border plays a significant role with migrants both more likely to leave from and migrate to closer places. Similarly, larger cities are more likely to be destinations for migrants. However, these are by no means the primary factors.

Given that Partition was ostensibly along religious grounds it is not surprising that the dominant factor determining out-migration, especially along the western border, was religion. Indian districts with greater numbers of Muslims and Pakistani/Bangladeshi districts with greater number of Hindus/Sikhs saw greater outflows.[4] Along the western border this religious "minority" exit is quite stark: The percentage of Muslims fell from 30 per cent in 1931 to 1.75 per cent by 1951 in districts that were eventually in Indian Punjab. Similarly, in the districts that became part of Pakistani Punjab, the percentage of Hindus/Sikhs fell from 21.7 per cent to 0.16 per cent!

What is perhaps most surprising is that there is a strong "relative replacement" pattern in determining where migrants went. In-migrants moved into the same areas/districts that saw greater outflows. For example, Delhi had 0.45 million people moving out to be replaced by 0.5 million people from across the border (about 28 per cent of the population of 1951). This replacement effect is all the more remarkable given that it is over and above any distance effect, i e, when comparing close by districts we find that those with greater outflows are precisely the ones with greater inflows. For example, Ajmer district, approximately same distance from the border as Delhi, had about 72,500 people move out and 71,300 people move in (only about 10 per cent of the population). Whether these in-migrants were allotted the property of those leaving is a much harder question to answer given the available data, yet our results do hint at this. More broadly they suggest that despite all the chaos that accompanies involuntary migrations, they can display a surprising degree of predictability.

In the subsequent sections we detail the construction of the data and variables of interest and then present the results. The data and methodology is of particular interest, and by making it available to a wider group we hope that it can form the basis of further work that can start examining both the short- and long-term consequences of the Indian Partition and more generally of involuntary migrations.[5]

2 Data and Variables

The primary sources of data used to compare pre- and post-Partition movements are the 1931 Census of British India and the 1951 Censuses of India and Pakistan. Since there is some controversy regarding the quality and coverage of the 1941 Census, with most demographers not considering it to be reliable, we use the 1931 Census instead to obtain pre-Partition demographics.[6] An important issue in using the two censuses, however, is identifying comparable enumeration areas. We describe how we address this issue and the construction of primary measures below. British India was divided into states which in turn were subdivided into districts.[7] In order to be able to present a detailed analysis, an important consideration for this study was to compile data at the lowest feasible geographical unit – the district. The district is the lowest administrative unit at which we are consistently able to find demographic data. Moreover, identifying the same geographical units over time becomes nearly impossible if one were to try and use lower administrative units such as tehsils.

Mapping districts pre- and post-Partition is a challenging task. Not surprisingly, the boundary creation as a result of Partition was accompanied by substantial reorganisation of state and district boundaries, not just for those regions that were split across the two countries but even within these countries. This was particularly true in areas where there were a lot of princely states since these states were by and large integrated into the provinces and districts of the new countries. At times a district

was split into two, or smaller districts merged into one for administrative or political reasons. Thus, district names need not match up between the two censuses, and even if they do there is no guarantee that they represent the same geographical area. An important contribution of our work has been constructing district level mappings between the two censuses. We do so by using detailed administrative maps from the two census periods to identify comparable areas and then comparing census data on reported land areas to ensure that our visual match was accurate. In several cases, the only feasible comparison entailed combining (typically adjacent) districts in 1931 and/or 1951. The matching process is described in more detail in the Appendix (p 48). Only a few districts could not be mapped. We were able to map 462 of the 472 districts and princely states of British India in 1931 and 363 of the 373 districts in India and Pakistan in 1951. Since some districts had to be merged we obtain a total of 287 comparable "districts" between the two census years.

There are two main variables used in our analysis: Inflows, the number of people moving into an area due to Partition, and outflows, the number of people moving out. We describe how both are obtained.

2.1 Inflows

An important variable we use in our analysis is the number of people who migrated into a district due to Partition – the inflows of migrants into the district. These numbers are obtained directly from the census since both the 1951 Censuses of India and Pakistan explicitly asked census respondents whether they had migrated during Partition. In the Indian census the term used for such migrants was "displaced persons," while the Pakistani census uses the term 'muhajir'. Displaced and muhajir specifically measure people that moved from India/Pakistan due to Partition. Internal migration is not measured by this variable and therefore it provides a good measure of the number of people who moved into both countries due to Partition.[8]

2.2 Outflows

Equally important is a measure of the number of people who left a district due to Partition – outflows. Unfortunately, the census data provides no direct way of estimating this number.[9] However, the fact that the migratory flows were essentially entirely along religious lines provides us with a methodology to estimate such outflows. While the methodology is admittedly rough, it does provide us with a sense of the magnitude and variability of outflows.

The methodology we use exploits the fact that the migratory flows were almost entirely along religious lines. Outflows are therefore considered to be Muslims leaving India (for Pakistan/Bangladesh) and Hindus/Sikhs leaving Pakistan and Bangladesh (for India). To simplify terminology we "abuse" notation slightly, by henceforth referring to these groups as "minorities". Hindus/Sikhs are minorities in Pakistan and Muslims are minorities in

India. The remaining groups in both countries will be referred to as the "majority". Note that we do not include other religious groups such as Christians, Buddhists, etc, as minorities since these groups were not thought to have been as affected in either country.[10] Consistent with this assumption we find that the percentage of Christians in India and Pakistan stayed relatively constant in 1931 and 1951. In order to compute outflows we need to estimate how many minorities left a district. The main issue in arriving at this number is to estimate the counterfactual of how many minorities would there have been in a district had Partition not occurred. Once this counterfactual, expected minorities, is estimated, outflows can be computed by subtracting the actual number of minorities in a district in 1951 from the expected minorities estimated for that district. So the main challenge is estimating the expected minorities in a district.

**Figure 1: Inflows in Terms of Absolute Numbers and Percentage of District Population**
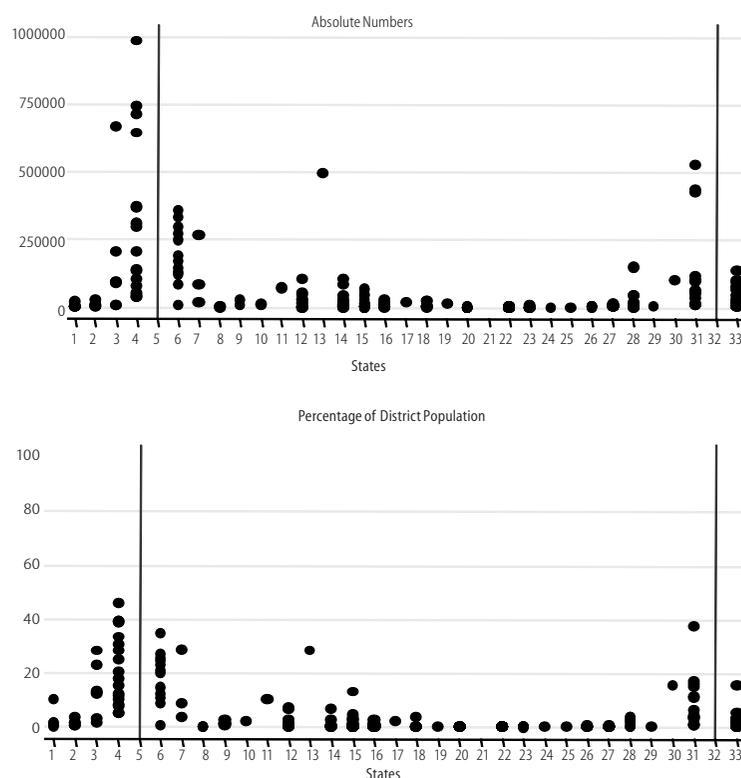
States Key (Used in subsequent figures as well).
Pakistan: 1=Baluchistan, 2=NWFP, 3=Sind, 4=Punjab, 5=Western border.
India: 6=Punjab,7=Pepsu, 8=Himachal Pradesh, 9=Saurashtra, 10=Kutch, 11=Ajmer, 12=Rajasthan, 13=Delhi, 14=Bombay, 15=Uttar Pradesh, 16=Madhya Pradesh, 17=Bhopal, 18=Madhya Bharat, 19=Vindhya Pradesh, 20=Hyderabad, 21=Andhra, 22=Madras, 23=Mysore, 24=Travancore Cochin, 25=Coorg, 26=Orissa, 27=Bihar, 28=Assam, 29=Manipur, 30=Tripura, 31=West Bengal, 32=Eastern border.
Bangladesh: 33=East Bengal.

An example will illustrate. Suppose that an Indian district had 1,00,000 Muslims in the 1931 Census. The 1951 Census shows that this district had 50,000 Muslims. Suppose the expected growth rate for Muslims in the 20-year period between 1931 and 1951 was a doubling of the population. Given the 1931 numbers, the expected number of Muslims in 1951 to have been 2,00,000. This gives total outflows in the district as 1,50,000 (i e, 2,00,000-50,000).

The accuracy of this calculation primarily relies on two assumptions. First, that flows due to Partition from a district

Figure 2: Outflows in Terms of Absolute Numbers and Percentage of District Population
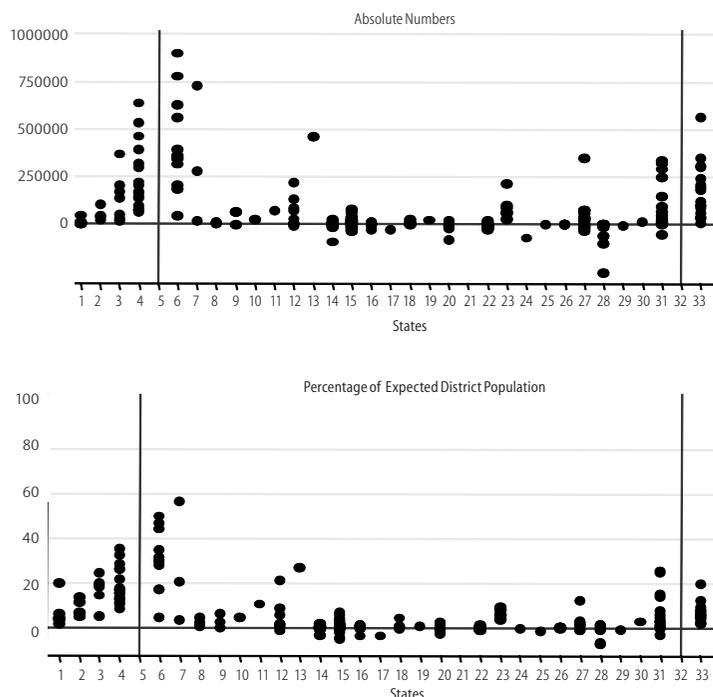
Expected population is the population of that district had Partition not happened.
States Key as in Figure 1.

were indeed religion specific (i e, Muslims were unlikely to migrate to India). Second, that we have correctly imputed the minority growth rate. Both the anecdotal evidence and data suggests that the first is likely to be true with the exception of maybe a few districts, particularly in Bengal (i e, along the eastern border).[11] However, estimating the counterfactual minority growth rate is a harder task and of particular concern as even small differences in growth rates can lead to large differences in absolute numbers.

To compute the minority growth rate from 1931 to 1951 we clearly cannot directly use 1951 minority numbers as these numbers changed due to Partition. Nevertheless, since the 1951 Census reports majority numbers separately for residents and migrants, we can calculate the growth rate for the majority group that is not directly affected by Partition flows. This is not enough, however, since imposing the majority growth rate on the minority population is likely to be problematic since the minority and majority groups typically had different growth rates prior to Partition. To address this problem we use a "scaling factor" which is the ratio of minority to majority growth rates from the previous 20-year period, 1901 to 1921. The 1931-51 majority growth rate is then rescaled by this factor to obtain the desired 1931-51 minority growth rate.

An alternate and seemingly simpler method would have been to directly use the 1931-41 (or 1901-21) growth rate of minorities to determine the 1931-51 minority growth rates. While we construct this measure and also present outflow estimates using it in the Appendix, we prefer not to use it since it makes a much stronger assumption in the data – that population growth rates did not significantly change over time.[12] While this is truer for states along the western border (where the two methods in fact give similar

outflow numbers), the high mortality due to the Bengal famine in 1943-44 meant that this was not true for the eastern states. We discuss these issues in more detail in the Appendix.

3 Results

This section presents the results of our analysis.

3.1 Overall Flows

Inflows: The total inflows into all three countries combined, measured in 1951, was 14.49 million or about 3.3 per cent of the total population at the time. However, this percentage hides substantial differences in the relative importance of flows. The absolute number of migrants into India was 7.3 million, into Pakistan 6.5 million, and into Bangladesh around 0.7 million. As a percentage of their populations these numbers are 2.04 per cent, 20.9 per cent and 1.66 per cent respectively. Migrants into Pakistan were clearly a very substantial presence.[13]

Outflows: While necessarily more tentative given the assumptions needed to construct them, we estimate that there were total outflows of 17.9 million from all three countries combined.[14] The outflow numbers for the three countries are 9.6 million out of India, about 5.4 million out of Pakistan, and 2.9 million out of Bangladesh. For numbers obtained using different methods of computing outflows, see the Appendix.

Interestingly, while outflows relative to the total counterfactual population are in similar proportions as inflows were in India and Pakistan (2.68 per cent and 18.01 per cent respectively) with Pakistan experiencing relatively large outflows (and inflows), in Bangladesh outflows were much larger than inflows both in absolute and relative terms. As a percentage of Bangladesh's population, outflows were a sizeable 6.51 per cent (as compared to 1.66 per cent for inflows).

Total Population: Did Partition result in a net increase or decrease in population for any country? While the inflows and outflows are large, the overall increase or decrease in population is smaller as these flows mitigate each other. For India, the net effect is a decrease of around 2.3 million people. In Pakistan, however, while 6.5 million entered, 5.4 million left Pakistan, hence a net increase of about 1.1 million people. In Bangladesh, there is more of a disjoint between inflows and outflows leading to a decrease of about 2.1 million.

Missing Persons: Since outflows represent people who left and inflows those who eventually arrived, by subtracting total inflows from total outflows we can obtain an estimate of the total number of "missing" people. We estimate a total of 3.7 million missing due to Partition. To the extent that the outflow measures are estimated accurately, this missing number includes people who died during Partition and those who migrated to another country

(apart from India, Pakistan or Bangladesh). While precise numbers are not available for the latter it is likely that it was not that significant, suggesting that, to the extent that the outflow calculations are accurate, the greater part of the missing number is likely to reflect mortality during Partition.
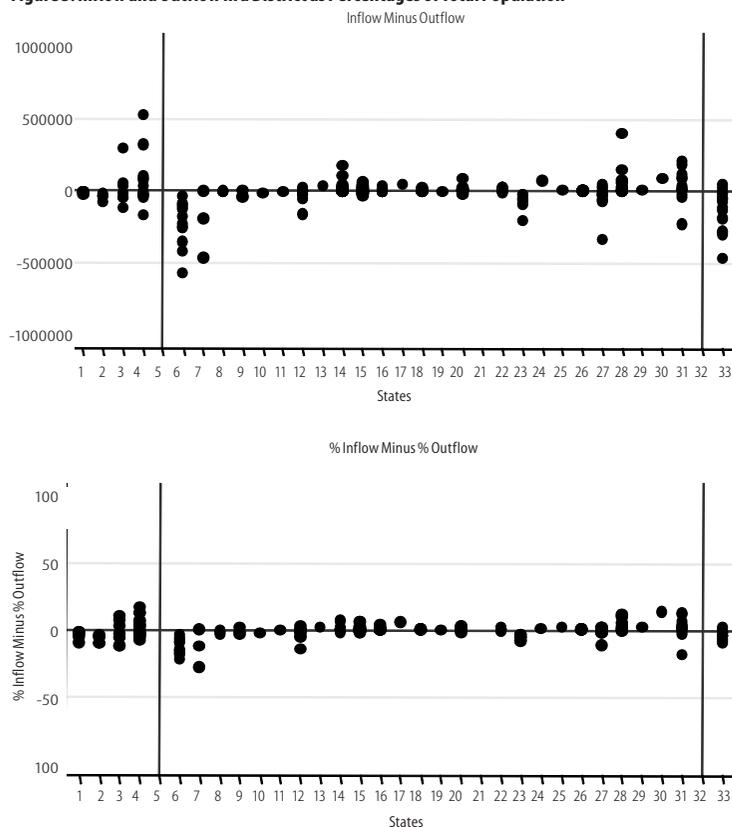
These estimates are fairly large but consistent with accounts in the literature. Lawrence James notes that, "Sir Francis Mudie, the governor of West Punjab, estimated that 5,00,000 Muslims died trying to enter his province, while the British high commissioner in Karachi put the full total at 8,00,000…This makes nonsense of the claim by Mountbatten and his partisans that only 2,00,000 were killed" [James 1998: 636]. Our estimate for the number of missing Muslims who left western India[15] but did not arrive in Pakistan is 1.26 million, reasonably close to the number cited by James. The corresponding missing Hindus/Sikhs along the western border is 0.84 million. This puts the total missing people due to Partition-related migration along the Punjab border at around 2.2 million. As percentages of the population that was on the move, the mortality rate along the Punjab border is similar for both religious groups. While approximately 16.1 per cent of all migrating Muslims went missing, 15.6 per cent of all migrating Hindus/Sikhs went missing. In another demographic study of the Partition, Hill et al (2006) estimate the number missing along the Punjab border at between 2.2 and 2.9 million. While our methods

Figure 3: Inflow and Outflow in a District as Percentages of Total Population

Inflow Minus Outflow

% Inflow Minus % Outflow

States Key as in Figure 1.

differ,[16] the fact that our numbers are similar for the Punjab border is a step forward in correctly thinking about mortality due to Partition.

Along the eastern border, our estimates are 1.1 million missing Muslims in Bangladesh (those Muslims who left India but were

not accounted for by arrivals in Bangladesh) and 0.24 million missing Hindus/Sikhs in the eastern Indian states, giving a total of 1.34 million missing along the Bengal border. Hill et al (2005) put the number missing in Bengal at around nine million. However, their methodology is unable to distinguish between mortality due to migration and mortality due to the Bengal famine. Our methodology is able to exclude mortality due to the Bengal famine, so long as we can assume that Muslims and Hindus/Sikhs suffered the same mortality rate due to the famine. In addition, while we do our best to account for internal migration from Bengal and the famine, we would exercise caution in attributing all the missing people in Bengal due to Partition-related mortality. This is party because anecdotal accounts suggest a lesser degree of violence along the Bengal border as opposed to the Punjab border.

3.2 Differences in Flows across Regions

Not surprisingly, migratory flows vary significantly across states with those closer to the borders both sending and receiving greater flows. However, what is somewhat surprising is that there is a substantial variation in these flows across districts within the same states, suggesting that distance was not the only factor. While we will try to determine the factors that influenced migration, in this section we simply illustrate the differences in migratory flows across districts.

Inflows: Figure 1 (p 41) shows inflows into each district in terms of absolute numbers and as a percentage of the district population. Since we will make use of such figures subsequently, it is important to explain this figure more carefully. Each point on the figure represents inflows into a particular district. The x-axis of this graph labels the state these districts belong to (thus all districts in a given state are plotted along the same vertical line). States are roughly organised from west to east within each country so the graph is akin to converting a map of the region into a single line "map".[17] The western and eastern borders are plotted as vertical lines for reference. Note that the distance between states in the figure does not reflect actual distance between them. We will provide graphs in subsequent figures that show actual distance.

The graphs illustrate several patterns. First, the migratory inflows that took place in the aftermath of Partition were primarily centred around Punjab (Indian and Pakistani), West Bengal and Bangladesh. The separation of over 2,000 km between Punjab and Bengal therefore made for two centres of Partition in India with other states playing a minor role in receiving displaced persons.

Second, the western and eastern borders experienced different dynamics of Partition. Till 1951 the flows on the western border were almost three times the size of the flows on the eastern border. The west in general received about 10.7 million people while the east received about 3.2 million. Moreover, while there was greater movement out of India than into it along the western border (Pakistani Punjab

received about twice the number of migrants as compared to Indian Punjab), it was the opposite along the eastern border – West Bengal received about twice the number of migrants as compared to Bangladesh.

Third, despite these large differences across states, there are significant differences in flows across districts. For example, in Indian Punjab, the district of Amritsar received about 3,32,000 people, while Gurgaon received only 84,000. In Pakistani Punjab, Lyallpur (now Faisalabad) received nearly a million migrants while Rawalpindi received about 1,06,000. Moreover, as the lower panel in Figure 1 makes clear, these differences are not only due to districts in a given state having different populations but also hold if we consider inflows as a percentage of a district's population.

Figure 4: Where Did the Migrants Go?

Y axis is inflows as % of total inflows.
Distance from border denotes distance from the closest border. Minority in India are Muslims and in Pakistan and Bangladesh, Hindus and Sikhs.

Outflows: The picture for outflows is similar to inflows and the same three patterns emerge (Figure 2, p 42). First, people moved out from the same two centres that saw the most inflows – Punjab and Bengal – both in absolute terms and relative to the population of these states. The rest of the states saw substantially lower outflows.

Second, as before the western border saw more people moving out than the eastern border although the outflows out of Bangladesh were fairly sizeable.

Third, there was a lot of difference in outflows across districts within the states that had large outflows. For example, in Indian Punjab district outflows vary from 17,000 (Kohistan district) to almost 9,00,000 (Amritsar district). In Bangladesh, Bogra saw an outflow of only 33,500 while 5,60,000 people were estimated to have left Dacca.

Total Population: Figure 3 (p 43) examines the net impact of Partition on the population of each district (i e, inflows less outflows). While the net effect of the migratory flows at the district level is also mitigated by outflows being compensated with inflows, as Figure 3 illustrates, there are nevertheless districts which experienced significant net population changes. For example, districts in Indian Punjab typically experienced reasonably large net population decreases while districts in Pakistani Punjab experienced net increases. On the eastern border, the effects were somewhat more muted, with districts in Bangladesh generally showing net

decrease with East Bengal districts typically showing net population increases.

3.3 Where Did the Migrants Go?

What determined where the migrants moved to during Partition? Our analysis reveals three important factors. First, migrants moved to places closer to the border – a distance effect. Second, they moved to the places vacated by those who were migrating out – a "replacement" effect. Third, large cities were more likely to attract migrants.

While we will employ multivariate regression analysis to establish these findings, Figure 4 illustrates their importance for the three countries. The y-axis is the inflows into a district as a percentage of total inflows into the country. We also display "fitted lines" in the figures which depict the bivariate relationship between percentage inflows and our factors of interest.

The fact that most of the movement took place around the border regions is clear from Figures 4.1-4.3. Moreover, we see in the data that districts within a 20 mile radius of the borders received about 12 per cent of the total inflows. Districts within a 50 mile radius received almost 50 per cent of the total inflows. This is a rather small radius for India and Pakistan, where the furthest district was 1,225 and 425 miles respectively from the closest border. In Bangladesh, this radius is relatively large since the furthest district was only 75 miles from the border. In terms of districts, this 50 mile radius captures 7.8 per cent of the total districts in India, 20 per cent and 64 per cent in Pakistan

and Bangladesh respectively. However, these figures also show that distance does not explain a lot of the variation across districts.

The Figures 4.4-4.6 show that at least for Indian and Pakistan, the replacement pattern is very significant. The fitted line for Pakistan, and to some extent for India as well, is almost a 45 degree line, implying a strong replacement effect.[18] This close relationship between people moving out and moving in is suggestive of reallocation of evacuee property to those migrating in. Interestingly, in Bangladesh this replacement effect is less important. As Kudaisya and Tan (2000) note, "while in Punjab the Indian government had facilitated an 'exchange of population', in Bengal it wanted to prevent precisely such an exchange, and took a number of initiatives to this end." Table 1 examines these effects in multivariate regressions. The dependent variable is the same as that on the Y-axis in Figure 2, inflows as a percentage of total inflows in the country. We run separate regressions for the three countries. All regressions also include the district's population as a percentage of country population to take into account whether migrants may simply have moved to larger districts. We also include a "big city" dummy variable which captures whether the district included a large city in 1931.[19] In addition we also include state level fixed effects to ensure our results are not just driven by comparing different states. Note when we include state fixed effects it is probable that most of the distance effects are unlikely to matter as much since distance does not vary as much across districts within a state. So our primary interest in looking at the regressions with state fixed effects is how robust the results

Table 1: Where Did Incoming Migrants Go?

(Dependent Variable: Inflows in district as % of total inflows in country)

	(1)	(2)	(3)	(4)	(5)
	India		Pakistan		Bangladesh
Distance	-2.522	0.012	9.003	22.011	262.624
(in 1,000 miles)	[0.731]***	[1.161]	[22.581]	[31.749]	[724.296]
Distance squared					
(/1000000)	2.133	-0.158	-21.847	-62.016	-7636.558
	[0.846]**	[1.114]	[90.307]	[106.091]	[11613.805]
Outflows as %					
of total out flows	0.539	0.535	1.145	1.169	0.954
in country	[0.036]***	[0.049]***	[0.290]***	[0.351]***	[0.507]*
City dummy	0.496	0.191	0.112	-0.501	-1.001
	[0.167]***	[0.153]	[1.691]	[2.019]	[5.495]
District population					
as % total population	0.06	0.155	0.221	0.38	-0.87
in country	[0.129]	[0.131]	[0.470]	[0.543]	[0.725]
Constant	0.539	0.546	-1.54	-3.57	4.59
	[0.128]***	[0.603]	[1.479]	[2.879]	[7.904]
State fixed effects	No	Yes	No	Yes	NA
Observations	233	233	35	35	17
R-squared	0.63	0.76	0.81	0.81	0.33

Std errors in brackets. * Significant at 10%; ** significant at 5%; ***significant at 1%.
The table examines the impact of distance from border, % outflow from a given district and the existence of a large city in the district on the inflows into that district. Columns 2 and 4 include state fixed effects for India and Pakistan. Bangladesh does not have state fixed effects since it has only one state. Computation of outflow is discussed in the Appendix. Inflows are people moving into a given district due to Partition, outflows are those moving out. Distance is measured as the straight line to the nearest India-Pakistan border from the centre of a district. City dummy was created from the 24 largest cities (in terms of population) from 1931. This data was obtained from the Historical Atlas of south Asia [Schwartzberg 1978]. There are 25 states in India and four states in Pakistan.

are for the variables which in fact do vary within a state, like outflows from a district.

Intuitively, distance would have a mostly negative effect on where people move; however, this result is only statistically significant in India and generally of small size. Inflows fall with distance to the border though at a decreasing rate (the distance squared term is positive albeit small) and for the first 100 miles inflows drop by around 0.23 per cent. However, beyond 600 miles from the border (the maximum distance in our data is 940 miles) the additional effect of distance is slightly higher inflows.

In contrast, the replacement effect is very large and holds for all three countries though it holds with less statistical significance in Bangladesh. Pakistan shows this replacement effect to be very important since the regression coefficient is close to one. It also matters in India, but not as much. Since these regressions include the district's relative population and state fixed effects, we can be assured that the replacement effect is indeed capturing a pattern of in-migrants going to places with high numbers of outmigrants and not simply that larger districts saw more in-migration or that certain states were more important. In fact, the insignificant coefficient on district population suggests that this was not an important factor at all.

Finally the results show that large cities attract more migrants. The effect

Figure 5: Where Did the Migrants Come From?
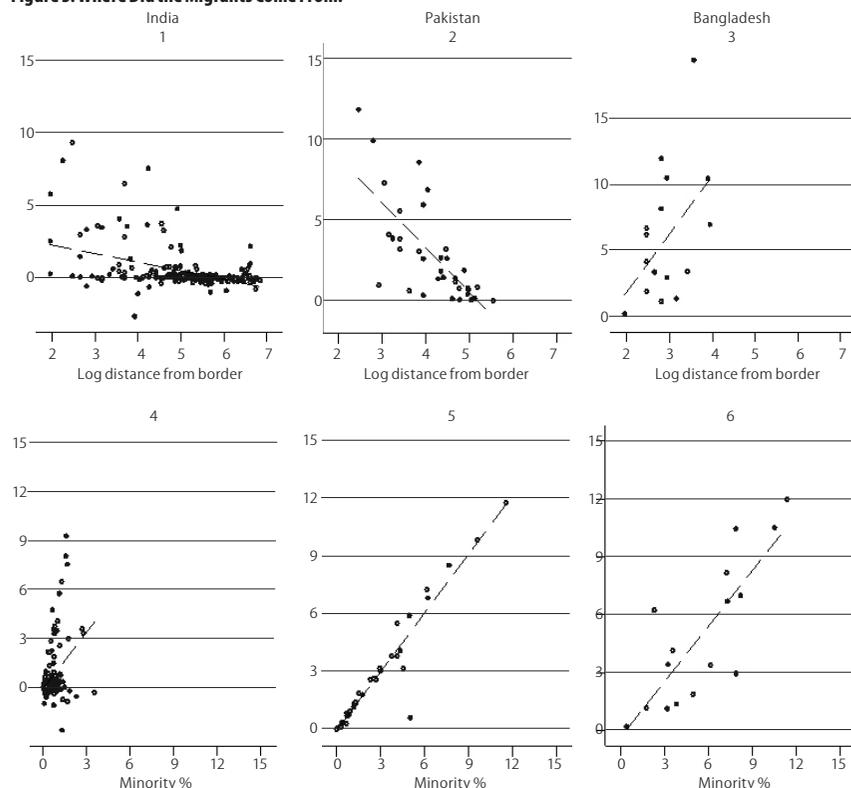
Y axis is district outflow as % of total outflow.
Distance from border denotes distance from the closest border. Minority in India are Muslims and in Pakistan and Bangladesh, Hindus and Sikhs.

holds strongly for India – having a large city in a district leads to 0.5 per cent more migrants. This variable is not statistically significant for Pakistan and Bangladesh. Part of the problem, however, is that there were very few big cities in these two countries (for example, Dhaka was the only "big-city" in our dataset for Bangladesh). Examining large cities in these countries does suggest that they mattered as well. For example, Karachi in Pakistan received more migrants than all of the districts in Sind put together. In fact, large cities often overcame distance barriers. The Indian city of Madras (Chennai), a very distant 800 kms from the closest border, still received about 4,000 migrants as compared to roughly the same number for the rest of the entire state of Madras (which includes 13 other districts and an area almost 1,000 times that of Madras city).

Table 2: Where Did Incoming Migrants Come From?
(Dependent Variable: Outflows in district as % of total inflows in country)

	(1)	(2)	(3)	(4)	(5)
	India		Pakistan		Bangladesh
Distance (in 1,000 miles)	−5.034	−5.083	−3.364	−18.023	−151.274
	[1.227]***	[1.519]***	[10.605]	[11.845]	[398.155]
Distance squared	5.084	3.977	28.694	55.191	3126.296
(/1000000)	[1.413]***	[1.472]***	[40.275]	[39.550]	[6326.028]
Minorities in district					
as % of total minorities	1.192	0.785	0.764	0.675	0.758
in country	[0.191]***	[0.144]***	[0.130]***	[0.123]***	[0.504]
City dummy	0.639	0.5	0.091	0.846	3.492
	[0.280]**	[0.202]**	[0.746]	[0.752]	[2.865]
District population					
as % total population	−0.915	−0.383	0.558	0.486	0.445
in country	[0.257]***	[0.205]*	[0.174]***	[0.183]**	[0.490]
Constant	0.977	1.208	−0.833	0.742	−0.299
	[0.215]***	[0.805]	[0.721]	[1.110]	[4.295]
State fixed effects	No	Yes	No	Yes	No
Observations	233	233	35	35	17
R-squared	0.31	0.72	0.94	0.95	0.73

Std errors in brackets. * Significant at 10%; ** Significant at 5%; ***Significant at 1%.
The table examines the impact of distance from border, % minorities in a given district, population size of the district and the existence of a large city in the district on the outflows from that district. Columns 2 and 4 include state fixed effects for India and Pakistan. Bangladesh does not have state fixed effects since it has only one state. Computation of outflow is discussed in the Appendix. Outflows are those moving out of a district due to Partition. Distance is measured as the straight line to the border from the centre of a district. Minorities in India are Muslims. In Pakistan and Bangladesh minorities are Hindus and Sikhs. City dummy was created from the 24 largest cities (in terms of population) from 1931. This data was obtained from the Historical Atlas of south Asia [Schwartzberg 1978]. There are 25 states in India and four states in Pakistan.

3.4 Where Did the Migrants Come From?

As in the decision of where to go, we find that while migrants typically came from places closer to the border, there was an important "unwelcome" effect analogous to the replacement effect: Outflows were far more likely to come from areas which had a greater proportion of minorities to begin with. This is not surprising since these minorities were likely to feel threatened in the newly created countries.

Figure 5 illustrates these relationships. The y-axis is the outflows from a district as a percentage of total outflows from the country and provides a measure of where the migrants came from.

With the exception of Bangladesh, we can see in Figures 5.1-5.3 that distance had a negative effect on outflows. Migrants were more likely to come from places closer to the border. Nearly 34 per cent of India's outflows were from regions that were within a 20 mile radius of the border, while the analogous number for

Pakistan is about 22 per cent. However, as before, the graphs also show that distance is not the only factor.

The Figures 5.4-5.6 show that outflows from Pakistan and India were determined in large part by the relative importance of minorities. Places with greater minorities saw greater outflows. The relationship is dramatic in Pakistan, where the exchange was almost one to one suggesting an complete exodus of Hindus/Sikhs. A striking feature of the migration on the western border was an almost complete "switching" of populations from Indian Punjab to Pakistani Punjab and vice versa. In Indian Punjab, the number of Muslims in 1931 was around 3.5 million and this had been reduced to 0.2 million in 1951. In terms of percentages of populations we see a drop from 30 per cent in 1931 to 1.8 per cent in 1951. In Pakistani Punjab, the numbers are even more drastic – the percentage of Hindus/Sikhs in the population drops from 22 per cent to a mere 0.16 per cent. At the district level the numbers reveal the same movement, in a more dramatic fashion. Amritsar in India had more than half a million Muslims in 1931, and in 1951 only 4,000 Muslims remained. Gujrat district in Pakistani Punjab had over 1,30,000 Hindus/Sikhs, but after Partition only 100 remained in 1951.

Table 2 presents the multivariate regression results for these factors and confirms the above relationships.

Distance matters as before, and is significant for India. Outflows fall with distance to the border, though at a decreasing rate (the distance squared term is positive albeit small) and shows that in India those areas next to the border had 0.36 per cent higher outflows than those regions 100 miles from the border. While this distance effect is still relatively small, it is larger for outflows than inflows. Thus distance mattered somewhat more while leaving as compared to migrating in.

The relative proportion of minorities in the district matters strongly in both India and Pakistan even after controlling for the relative population of the district. In India, for every 1 per cent increase in the minority ratio in 1931 we see a 0.78 per cent increase in the outflows. In Pakistan the analogous number is 0.67 per cent for every 1 per cent increase in the minority ratio.

While Figure 5.6 for Bangladesh suggested similar population changes on the eastern side, in fact it is misleading since the same districts with large minority fractions were also large. Once we take this into account in the regression analysis in Table 2 we see no dramatic population exchanges. In fact, the proportion of Hindus/Sikhs went from about 30 per cent in 1931 to 22 per cent in 1951 in Bangladesh. In West Bengal the numbers are similar – Muslims accounted for about 30 per cent of the population in 1931 and fell to 19 per cent in 1951. The fact that neither distance nor percentage of minorities seems to matter much in Bangladesh implies that, as suggested by Kudaisya and Tan (2000: 144-61) and others, the decision to migrate in Bangladesh was fairly different from that along the western border.

4 Conclusions

This paper examines the nature of migratory flows four years after Partition. While migration continued even after 1951, it is safe to say that the numbers presented in this paper capture the bulk of the migration.[20]

While we estimate the number of outflows and the number of people missing due to Partition, we do urge the reader to regard these numbers with caution. The assumptions used in computing these figures are detailed in the Appendix. However, demographers that have specifically studied mortality due to Partition in Punjab and Bengal [Hill et al 2006] also produce similar estimates.

This paper serves to answer the more basic question of how many moved and where. In Bharadwaj, Khwaja and Mian (2008) we study the effects of the migratory flows on overall gender ratios, literacy levels and occupation structures of India and Pakistan. In subsequent work we hope to examine the impact of these flows on outcomes like agricultural productivity, health, etc. We hope that quantifying this event of human history will encourage more empirical research related to Partition as well as forced migrations in general. To facilitate this, we provide the data used in this paper on the web. Combined with our other work in this area, we hope to provide a glimpse into the long-term consequences of these movements, specifically not just how well migrants are able to adapt but how their settlement impacts the trajectory of the places they moved to. Given the current events and importance of south Asia, we hope that such a historical empirical analysis may prove to be of value.

NOTES

1 For a literature review please see Bharadwaj, Khwaja and Mian, 2008, hereon BKM.

2 Per cent inflows are calculated relative to the current population in 1951, while per cent outflows are calculated relative to the district's projected population in 1951 had Partition not occurred.

3 We include princely states (Pepsu) and Himachal Pradesh – these states were created from or merged into Indian Punjab districts.

4 Throughout this paper we refer to the religion of migrants coming to India as "Hindus/Sikhs". It should be noted that while Hindus and Sikhs formed the majority of the in-migrants in Indian Punjab, very few Sikhs were part of the in-migrating population in West Bengal.

5 We hope to make all the basic census data collected available on the web site hosted by the South Asia Initiative at Harvard University.

6 The introduction to the 1941 Census itself raises concerns about quality and coverage of the census, with the census commissioner admitting that "There was a tendency in the more communal quarters to look on the census enumerators as the ready tools of faction" (p 9) and that "The main point [about completion of enumeration] which emerges at once is that the great population regions of the Indus and Ganges systems in which nearly half the total population of India lies have only a limited representation in the census figure" (p 11). More details are in the Appendix.

7 For more on the British spatial system, see Kant (1988).

8 These numbers could be inaccurate if individuals misreported their migrant status. However, we have little reason to suspect that there were significant incentives to do so.

9 Unlike migrants into a district that can be directly ascertained by asking a person's status in 1951, there is no simple way to ask how many people left. The direct way would have been to ask the migrants in 1951 which district they migrated from, but to our knowledge no such information was solicited in the census.

10 See http://www.ishr.org/activities/religious-freedom/pakistan-india-bangladesh.htm Downloaded September 2006.

11 Unfortunately since the census does not ask the religion of migrants there is no direct way to test this in the data.

12 One major reason to not use this growth rate is that the Bengal famine occurred in 1943-44 and reportedly killed about four million people. Hence our estimates for expected minorities would then include people that died due to the famine, which makes mortality estimates due to Partition more difficult to calculate. By using 1931-51 growth rates of majorities, we do assume that majority and minority groups had an equal probability of dying during the famine.

13 To put the number for India in perspective, we calculate from Srivastava and Sasikumar (2003) that internal migration rate in India was around 11 per cent in 1992. Hence an impact of 2 per cent in migration in 1951 is a potentially large effect.

14 These numbers are estimated in terms of 1951 population levels, i e, given our construction they also include any children born between 1947 and 1951 for the outmigrating families. We do so because the numbers for inflows are also in 1951 and therefore include children born to the in-migrants. One could convert these numbers into 1947 numbers by discounting the numbers by the population growth rate between 1947 and 1951. However, we prefer not to do so both because such accurate birth and mortality rate data is not available and also because these flows did not only occur in 1947 but continued for a few years.

15 Western India is defined by districts whose closest border is the Punjab border. This calculation also assumes that outflows from these districts were headed to Pakistan, and not Bangladesh.

16 Hill et al use the intercensal survival technique to compute losses due to Partition. At the district level they also use a technique based on the demographic balancing equation. They make extensive use of age specific information available in the censuses, and also use the 1941 Census. While their methods are likely more suited for computing mortality, they do potentially suffer more from biases generated due to non-partition related mortality such as that during the Bengal famine.

17 The west to east sequence is not always preserved. For example, Assam is to the east of East Bengal but since the former is in India and the latter in Bangladesh we distort this single-line "map" slightly in order to keep all states in a country together, by putting Assam before East Bengal.

18 This finding is not inconsistent with the finding of net population effects. To uncover a replacement effect, we examine inflows and outflows as percentages of total inflows and outflows respectively. Hence, the replacement effect states that if 10 per cent of all inflows went to a district, it is likely that a similar percentage of the outflows left from this district.

19 Large cities are the 24 largest cities (in terms of population) from 1931. This data was obtained from the Historical Atlas of South Asia [Schwartzberg 1978].

20 An estimated 1.7 million people migrated between 1950 and 1952 (Pakistan Constituent Assembly Debates 1952).

REFERENCES

Bharadwaj, P, A Khwaja and A Mian (2008): 'The Partition of India: Demographic Consequences', Working Paper.

Davis, K (1951): *The Population of India and Pakistan*, Princeton University Press, Princeton.

Government of India (1933): *Census of India 1931*, Central Publication Branch, India.

– (1941-45): *Census of India 1941*, The Manager of Publications, Delhi.

– (1952): *Census of India 1951*, The Manager of Publications, Delhi.

Government of Pakistan (1954-56): *Census of Pakistan 1951*, The Manager of Publications, Karachi.

Hill, K et al (2005): 'The Demographic Impact of Partition: Bengal in 1947', down-loaded from iussp2005. princeton.edu/download.aspx?submissionId=52236.

– (2006): 'The Demographic Impact of Partition: The Punjab in 1947', Working Paper 06-08, Weatherhead Centre for International Affairs, Harvard University.

James, L Raj (1998): *The Making and Unmaking of British India*, St Martin's Press, New York.

Kant, Surya (1988): *Administrative Geography of India*, Rawat Publications, Jaipur.

Kudaisya, G and T Tan Yong (2000): *The Aftermath of Partition in South Asia*, Routledge, London.

Schwartzberg, Joseph E (ed) (1978): *A Historical Atlas of South Asia*, University of Chicago Press, Chicago.

Sherwani, L A (1986): *The Partition of India and Mount-batten*, Council for Pakistan Studies, Karachi.

Srivastava, R and S K Sasikumar (2003): 'An Overview of Migration in India, Its Impacts and Key Issues', www.livelihoods.org.

US Committee for Refugees and Immigrants (2006): 'World Refugee Survey', http://www.refugees.org.

Appendix

The Census of 1941

As we noted in the main text, our decision not to use the 1941 Census of India is based on various significant concerns regarding the quality and coverage of this census.

This is perhaps best illustrated by a series of statements by the Census Commissioner of the 1941 Census, M W M Yeats, in his introductory remarks to the 1941 Census.

Yeats starts off by noting that:

> The war has laid its hand on the Indian census as on every other activity of the India government and people.…It was considered however that financial conditions did not permit the completion of the tables and as I write this brief introduction I am no longer, and have not been for a year, a whole-time Census Commissioner (p 2)

and goes on to lament that

> One of the last things to be desired in a census is uncertainty; yet that pursued us to the end. It was till February 1940 that the government of India decided to have a census at all. A still greater difficulty was caused by the delay in deciding how far to go with the tabulation (p 2).

In addition, Yeats talks about lack of tabulation facilities, buildings and officers. He talks about problems with some provincial tables that had to remain unresolved because provincial census officers were removed from their jobs as soon as the tables went to press, and hence were not available for further clarifications. He states that, "The main point [about completion of enumeration] which emerges at once is that the great population regions of the Indus and Ganges systems in which nearly half the total population of India lies have only a limited representation in the census figures" (ibid: 11) and also points out his concern regarding biased estimates and mismeasurement since, "There was a tendency in the more communal quarters to look on the census enumerators as the ready tools of faction…" (p 9) and "At that time Mr Gandhi's civil disobedience campaign was in full swing and all over north India the census, as a governmental activity, incurred hostility" (p 24). The 1951 Pakistan census also starts (p 1) by noting that the 1941 Census "had not been tabulated in full owing to the war, and their accuracy has been prejudiced by the efforts of different communities to inflate their figures for political purposes".

District Mapping Over Time

Unlike later censuses, the 1951 Census does not provide a comprehensive mapping of the districts in 1951 to those in previous census years.

As such, our approach is to use detailed maps in 1951 and 1931 and start by visually identifying mappings between districts in the two time periods. Once the visual exercise reveals potential matches between the two census years, we use census data for land areas of these regions and only consider a mapping to be permissible if the land areas of the two units are within 10 per cent of each other. We also perform robustness tests with lower thresholds. If two areas do not meet these criteria we attempt to map them at higher levels of aggregation (for example, by combining adjacent districts). In the majority of cases we are able to map regions over time and only a few districts could not be mapped. Thus for the 472 districts and princely states of British India in 1931 we are able to map 462. The equivalent number for the 1951 districts is 373 mapped out of a total of 363. Since some districts had to be merged this gives us a total of 287 comparable "districts" between the two census years.

Districts Not in Dataset

These districts are not in our data set because of lack of information in a certain year or merging issues.

NWFP Frontier Areas (only British areas were censused in 1931): (1) Chitral, (2) Malakand, (3) Swat, (4) Dir, (5) North and South Waziristan, (6) Khurran, (7) Khyber.

Baluchistan (one area was not censused in 1951): (1) Dera Ghazi Khan.

Gilgit Agency (not censused): (1) Yasin, (2) Kuh Ghizar, (3) Punial, (4) Tangir & Darel, (5) Ishkuman, (6) Gilgit, (7) Chilas, (8) Astor, (9) Hunza and Nagir.

Assam Hill/Tribal Areas (not censused in 1951): (1) Sadiya Frontier Tract, (2) Khasi and Jaintia Hills.

Jammu and Kashmir (not censused in 1951): (1) Baramula, (2) Anantnag, (3) Riasi, (4) Udhampur, (5) Chamba, (6) Kathua, (7) Jammu, (8) Punch, (9) Mirpur, (10) Muzaffarabad.

Andaman and Nicobar Islands (have missing information in the 1931 Census). Sikkim (its status was uncertain in 1951 and was only inducted into state of India in 1975).

Computing Outflows

Outflows – the Main Measure: Our method of computing outflows determines expected
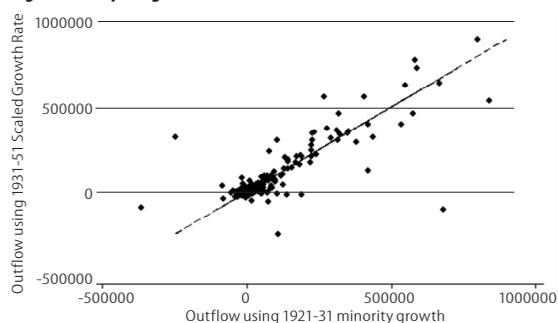
minority growth rates by re-scaling the growth rates of the majority population during the relevant period (1931-51). Note that "minorities" in Pakistan are non-Muslims, while minorities in India are Muslims. "Majority" in India are non-Muslims, while majority in Pakistan are Muslims. We define the resident majority growth rate as $M_g = M_r^{1951}/M_r^{1931}$, where M_r denotes resident majority. The resident majority population in 1951 is calculated as the total population of the majority group in 1951 less the population of incoming migrants (incoming migrants belonged to the majority). Majority is defined as the population minus minority populations. In our notation, upper case M always refers to the majority, while lower case m refers to minorities.

Next we construct the scaling factor to adjust the majority growth rate to reflect minority growth rate from 1931-51. We need a scale because, as is clear in Table A1, Muslims tended to grow faster than non-Muslims in British India.

We use a 20-year scale because our majority growth rate is measured over 20 years as well. It is obvious that we cannot use 1931-51 growth rates of minorities as a scale, since minorities were on the move by 1951. We need to look to previous years for a scale. We did not use the 1941 Census because its quality is suspect. Our next choice was using 1911-31 growth rates to compute the scale. However, these growth rates are likely to be very different from those in 1931-51 due to large internal migrations that took place in the 1920s. These migrations were primarily located in the east, with people moving from Bengal into Assam to work on the tea estates [Davis 1951]. In comparison we are aware of no significant criticism of 1901-21 Censuses as far as religious enumeration is concerned. To avoid problems of countering massive internal migrations and census accuracies, we therefore use the 1901-21 growth rates to compute our scale.

Now we can impute the minority growth rate between 1931 and 1951 as:

$G_m^{1931-51} = G_M^{1931-51} \times S$, where G_m and G_M refer to minority and majority growth rates between the relevant period.

Finally we can compute the expected number of minorities in 1951.

$E(m_{1951}) = m_{1931} \times G_m^{1931-51}$

Outflows is the number of expected minorities less the actual number of minorities in a given district:

Outflow = $E(m_{1951}) - m_{1951}$

The above analysis is computed at the district level with one exception. We do not have 1901 Census figures at the district level. Hence, we just use the countrywide scale on the 1931-51 majority growth rate at the district level.

Secondary Measure of Outflows: The departure in this method of computing outflows

is in the way we compute minority growth rates:

$G_m^{1921-31} = m_r^{1931}/m_r^{1921}$

In other words, rather than re-scaling the 1931-51 majority growth rates we instead use the minority growth rate from 1921-31. Therefore, the counterfactual number of minorities in 1951 is:

$E(m_{1951}) = m_{1931} \times G_m^{1931-51}$
Outflow = $E(m_{1931}) - m_{1951}$

The problem with this measure is that the Bengal famine occurred in 1943-44 and its effect is hard to separate at the country level – i e, we do not have information on how many Muslims or Hindus died as a result of it. Hence, once we compute expected minorities in 1951, we need to subtract the deaths due to famine to get at the number missing due to Partition. Given that this measure would be heavily dependent on estimates of numbers of people that died due to the famine, it is likely to be less accurate than the first method.

There are additional problems in using this measure along the eastern border. Bengal saw large out migration of people moving into Assam until the 1930s. As a result, 1921-31 growth rates are in fact lower than the actual growth rates in 1931-51 (when there was no longer this migration into Assam) and this in turn would lead to underestimates of outflows from Bangladesh. In fact, for exactly the same reason we would predict that estimates of outflows from the eastern part of India would be overestimated if we use the 1921-31 growth rates of minorities (Muslims) in India. Examining these estimates shows that this is indeed the case.

This method also suffers from the fact that growth rates of religious populations are far from stable over decades. A glance at Table A 1 confirms this.

Comparing aggregates we find that while the results for India and Pakistan are about the same, Bangladesh's outflows are severely underestimated using the second measure (Table A2).

For reasons stated above we believe that outflow 1 captures the out migration more accurately. We can also compare outflows obtained at the district level using these two methods. We will see that they are essentially the same, except for the eastern region estimates. Most points are on or near the 45 degree line. The outliers are, not surprisingly are districts in Assam and Bengal, where we suspected problems with over and underestimations of growth rates.

Table A1: Growth Rates in British India

Years	Non-Muslims Growth Rate	Muslim Growth Rate
1901-11	1.0572	1.0920
1911-21	0.9945	1.0467
1921-31	1.0963	1.1169
1931-41	1.1376	1.1909

Table A2: Comparing Outflows (in million)

	Measure 1	Measure 2
India	8.4	8.7
Pakistan	5.4	5.5
Bangladesh	2.9	1.7

Figure 1: Comparing Different Measures of Outflow