Migration, externalities, and the diffusion of COVID-19 in South Asia☆

Jean N. Lee [a], Mahreen Mahmud [b], Jonathan Morduch [c,*], Saravana Ravindran [d], Abu S. Shonchoy [e]

[a] World Bank, United States
[b] University of Exeter, United Kingdom
[c] New York University, United States
[d] National University of Singapore, Singapore
[e] Florida International University, United States

ABSTRACT

The initial spread of COVID-19 halted economic activity as countries around the world restricted the mobility of their citizens. As a result, many migrant workers returned home, spreading the virus across borders. We investigate the relationship between migrant movements and the spread of COVID-19 using district-day-level data from Bangladesh, India, and Pakistan (the 1st, 6th, and 7th largest sources of international migrant workers). We find that during the initial stage of the pandemic, a 1 SD increase in prior international out-migration relative to the district-wise average in India and Pakistan predicts a 48% increase in the number of cases per capita. In Bangladesh, however, the estimates are not statistically distinguishable from zero. Domestic out-migration predicts COVID-19 diffusion in India, but not in Bangladesh and Pakistan. In all three countries, the association of COVID-19 cases per capita and measures of international out-migration increases over time. The results show how migration data can be used to predict coronavirus hotspots. More broadly, the results are consistent with large cross-border negative externalities created by policies aimed at containing the spread of COVID-19 in migrant-receiving countries.

© 2020 Published by Elsevier B.V.

1. Introduction

As COVID-19 spread globally, governments faced one of the hardest choices in public economics: how to choose between saving lives or saving livelihoods (Viscusi, 2020)? Some argued that placing priority on public health is a moral imperative and would also be the best economic policy in the long run. In Spring 2020,

most European countries accordingly sharply restricted mobility and ordered non-essential businesses closed, even though it carried steep short-run economic costs (Thomsen, 2020). Other countries (especially low- and middle-income economies) were pressed to put greater weight on immediate challenges, including poverty and hunger, caused by lost jobs and disappearing businesses (Abi-Habib and Yasir, 2020; Dahir, 2020; Sen et al., 2020). As a result, lockdowns in Spring 2020 were shorter and restrictions on mobility and commerce were generally less stringent—even if that risked raising rates of COVID-19 infection (International Monetary Fund, 2020; Malik et al., 2020).

The strict lockdowns in richer countries closed workplaces and curtailed infection by limiting negative externalities from personal interactions (Bethune and Korinek, 2020). Simultaneously, however, the policies created negative externalities elsewhere by pushing migrant workers from poorer countries to travel home (World Bank, 2020; Mitra et al., 2020). We quantify the spread of COVID-19 as migrants returned to Bangladesh, India, and Pakistan, bringing the risk of contagion from their former workplaces. We then document how patterns of prior international migration predict potential hotspots (see also Ahsan et al., 2020).

Globally in 2019, India was the leading country of origin for international migrants, Bangladesh was sixth, and Pakistan was

seventh (United Nations DESA, 2019).[1] The countries were also the three most affected by the initial spread of the coronavirus in South Asia. In South Asia, the movement of migrants heading home in anticipation of lockdowns, both internationally and within countries, was the largest mass migration since the 1947 partition of India, Pakistan, and what is now Bangladesh (Ellis-Petersen and Chaurasia, 2020).[2] The migration brought fear of COVID-19 contagion and the shunning of migrants as they spread through the subcontinent (Pandey, 2020; Bisht, 2020).

We use labor force surveys and household-level economic surveys from prior years to measure the extent of out-migration for each of 755 districts in the three countries and to distinguish between international and domestic migration. We then use data on migration patterns to predict the incidence and number of confirmed COVID-19 cases in each district using district-day data, beginning on the day of the 100th confirmed case in each country and, subject to data availability, continuing for the following 1.5 months.[3] Out-migration in earlier periods is used as an indicator of reverse migration in February, March, and April 2020. To capture broader patterns of diffusion, we flexibly control for trends over time with day-level fixed effects.

The data establish that international migration predicts the spread of the coronavirus across and within districts in India and Pakistan. In the 45 days following the first 100th case in each country, a 1 standard deviation (SD) increase in prior international out-migration from the district-wise average (measured as the number of out-migrants per capita and averaged over the cross-section) predicts a 48% higher number of cases per capita. In Bangladesh, however, where COVID-19 testing rates were substantially lower, the correlation of international out-migration and the number of confirmed cases is imprecisely measured and flips sign from positive to negative in some specifications.

Domestic migration is a weaker and less consistent predictor of contagion. While domestic migrants who returned to their home towns and villages were a focus of local fears and national policy debate (Ray and Subramanian, 2020), the detectable effects of domestic migration are low in Bangladesh (a 3% increase in the probability of any COVID-19 cases is predicted by a 1 SD increase in domestic migration in the average district) and negative in Pakistan. In India, where fears of contagion from migrants were reported widely (Mitra et al., 2020), a 1 SD increase in domestic migration in a given district predicts a 11% increase in whether or not a district reports any cases of COVID-19 during the study period, but predictions of the number of cases are measured imprecisely.

The predictive power of international migration patterns is not driven mechanically by the COVID-19 experiences of the returning migrants themselves. Instead, the results are consistent with community spread seeded by migrants (and others) from abroad. First, Section 3.3 shows that plausible magnitudes of return-migration and infection are too small to match the scale of reported COVID-19 cases. Second, Section 2 and Appendix A describe policies that limited international air travel and dramatically slowed the influx of migrants near the start of our study windows. Third, Section 5 shows that the predictive power of the measures of inter-

national migration to explain COVID-19 cases increased steadily over time, consistent with community spread and, given the policy time-line, inconsistent with COVID-19 infections suffered by international migrants directly.

Quarantines for migrants and other travelers were imposed in the three countries, but the findings, especially for international migrants, are consistent with worries that the policies were not implemented stringently. The evidence is equivocal with respect to the role of domestic migrants, showing an advantage of household-level and individual-level survey data which provides the ability to distinguish between international and domestic migrants—and thus to provide insight into health-related cross-border externalities.

The estimates are not causal parameters. The pattern of coronavirus cases is affected by demographics, climate, the stringency of the lockdown, the nature of the initial spread, and other factors. The results necessarily reflect complicated interactions of biology, policy, and human behavior, as well as omitted variables correlated with migration patterns. Yet the results are consistent with the nature of timing of COVID-related policy decisions and cross-border mobility, and they show that data on migration patterns, drawn from existing labor and household surveys, can help pinpoint patterns of diffusion and anticipate which districts are likely to face particularly acute healthcare needs.

2. Policy responses to COVID-19 in South Asia

COVID-19 cases were first reported on January 30 in India, on February 26 in Pakistan, and on March 8 in Bangladesh. Cases then grew steadily. Our analysis continues through May 20, 2020, 45 days after the 100th case in Bangladesh. By then, there were 112,028 confirmed cases in India; 45,898 cases in Pakistan, and 26,738 cases in Bangladesh.[4]

The severity and infectiousness of the virus prompted governments to begin restricting mobility in March 2020 (International Monetary Fund, 2020). Facing job losses in host countries and movement restrictions in home countries, many migrants rushed to get home. Travel bans imposed in the three countries meant that their window to return home was short. About a week after the start of our samples, international borders had largely closed (a week before the day 1 in Bangladesh, on the day 8 in India and on the day 7 in Pakistan; see Appendix A). The growth of cases in the second part of the samples is thus almost entirely due to community spread rather than to cases brought by returning international migrants.[5]

In Bangladesh, international migrants returned while a tide of domestic migrants also returned to rural areas of the country from Dhaka, in part spurred by government promises that shelter and food would be provided in rural areas (UNB, 2020). On March 14, visa requirements were made stricter, and flights from Europe (except the United Kingdom) were halted (Daily Star, 2020). Travelers were requested to self-quarantine for 14 days.

Government telecom administration data indicate that as many as 10 million subscribers initially left Dhaka in the days following the announcement of the government Independence Day holiday on March 26, 2020, which marked the start of a ten-day lockdown (Dhaka Tribune, 2020). Some of this was followed by re-migration back to Dhaka, as the Bangladesh Garment Manufacturers and Exporters Association reopened factories on April 4, 2020.

---

[1] In 2019, India accounted for 17.5 million international migrants, Bangladesh for 7.8 million, and Pakistan for 6.3 million. (Source: United Nations DESA, 2019.) Together, the three countries include 1.8 billion people or 28% of the world population outside of China.

[2] Bangladesh imposed a nationwide lockdown on March 26, 2020, India imposed a nationwide lockdown on March 25, and most provinces in Pakistan were under lockdown as of March 23, 2020.

[3] The 100th case of COVID-19 was confirmed on April 6, 2020 in Bangladesh, on March 15 in India, and on March 16 in Pakistan. The sample windows are thus April 6-May 20 for Bangladesh; March 15-April 28 for India; and March 16-April 27 for Pakistan.

[4] Data are for confirmed COVID-19 cases on May 20, 2020 from https://www.worldometers.info/coronavirus/#countries. By September 20, 2020, India had reported 5,485,612 cumulative cases, Pakistan 305,671 cases, and Bangladesh 348,916 cases.

[5] The incubation period of COVID-19 averages 5–6 days and can take up to 14 days (World Health Organization, 2020).

Shonchoy (2020) partnered with epidemiologists from the International Centre for Diarrhoeal Disease Research, Bangladesh to show that the earliest outbreaks outside of Dhaka were predicted partly by migration patterns.

In India, similarly, reverse international migration was paired with the urban-to-rural movement of domestic migrant workers. The government first announced travel restrictions on March 11, 2020, mandating quarantines for international passengers arriving from China, South Korea, Iran, France, Germany, Italy, and Spain. By March 22, India closed its borders to all international commercial flights.[6] Domestically, a national government lockdown was announced on March 24, 2020, which restricted the movement of people throughout the entire country. Faced with loss of employment, migrant workers immediately left major urban centers for rural areas (e.g., Denis et al., 2020). While some migrants successfully reached home, others were stymied by the stoppage of transport services and journeyed hundreds of miles on foot. (In early May, special trains were arranged to take migrants home; Al Jazeera, 2020.)

In Pakistan, in addition to returning economic migrants who worked internationally and domestically in urban areas, the country received travellers who had attended large religious gatherings, including in Iran, a badly-hit neighbor (Emont and Shah, 2020). On March 21, the government suspended international flights for two weeks. The Government of Sindh announced a lockdown in the province for 14 days from March 23, 2020, ordering all public transport, markets, offices, shopping malls, restaurants, and public areas to be shut down (Arain, 2020). Punjab also was put on lockdown on March 24, 2020 (ARY Web, 2020). Pakistan announced an extension of the lockdowns at the beginning of April, with further extensions after April 14 to May 31 (SNS Web, 2020).

The analysis focuses only on confirmed cases. This has at least two implications. First, actual cases are likely much higher (The Economist, 2020). Second, the nature of testing protocols, especially in Bangladesh and India, where contact with international travelers was an early screening criterion, can partly account for positive correlations with migration and confirmed cases. By the middle of the sample period, however, testing protocols in Bangladesh and India had expanded to include all hospitalized patients showing respiratory symptoms associated with COVID-19. In Section 5, we show results for the entire sample window and for the period after testing was expanded, documenting the robustness of correlations after testing protocols had broadened. We also note that the results for Pakistan and India are broadly similar, although in Pakistan international travel history was never a screen for testing (see Appendix A for policy details).

3. Data

3.1. Tracking COVID-19 cases by district

The analysis focuses on the initial stage of the pandemic, starting with the day that the 100th case was confirmed in each country: in Bangladesh, April 6; in India, March 15; and in Pakistan, March 16. Subject to data availability, we then analyze cases over the next 1.5 months. This covers 45 days in Bangladesh, 45 days in India, and 43 days in Pakistan.[7]

Indicators for daily COVID-19 cases by district for Bangladesh were obtained from the Institute of Epidemiology, Disease Control and Research (IEDCR) in Bangladesh.[8] Data on COVID-19 cases by

district for India were taken from the CovIndia website, which provides daily updates on the number of cases by district and day.[9] Data on daily COVID-19 cases by district for Pakistan were obtained from the Government of Pakistan COVID-19 dashboard.[10]

3.2. Using national surveys to measure migration

Data on international and domestic migration by district for Bangladesh were computed from the Household Income and Expenditure Survey (HIES) 2016, a large nationally representative household survey. The indicator is calculated as the number of migrants identified in the survey divided by the number of individuals in the survey (in hundreds), adjusted using survey weights. International migration is accounted for separately from domestic (within-country) migration. When estimating, the measures are normalized as z-scores calculated using country-specific distributions.

The migrants captured in the HIES 2016 surely did not all return to Bangladesh in response to the COVID-19 crisis, some may have returned earlier, and another group of migrants (those who left post-2016) are not captured at all. Still, migration patterns tend to be relatively stable, and recent evidence validates the use of the 2016 HIES data for Bangladesh to capture reverse-migration in 2020. Ahsan et al. (2020) show a significant correlation between the HIES 2016 district-level data on migration and coronavirus-related quarantines in districts in 2020 (correlation = 0.51, *p*-value < 0.01) and between the migration data and distress calls to a government coronavirus hotline (correlation = 0.54, *p*-value < 0.01). As Ahsan et al. (2020) note, this makes widely-available surveys like the HIES particularly valuable when contemporaneous data on population mobility is unavailable. Surveys like the HIES also have the advantage of distinguishing between international and domestic migration.

Measures of international and domestic migration for India were similarly calculated from the most recent migration module in the National Sample Survey (NSS), the 2007–2008 round. Respondents were asked to report the number of migrants who left the district for another country or another district within the past five years. The indicator is calculated as the number of migrants in the survey divided by the number of individuals in the survey (in hundreds), adjusting using survey weights.

Data on international migration for Pakistan were taken from the Pakistan Social and Living Standards (PSLM) 2014–2015 survey. Assuming one international migrant per surveyed household that receives international remittances, the indicator is the number of international migrants per 100 people in the households surveyed in the district, adjusted using survey weights. Data on domestic migration for Pakistan come from the Pakistan Labour Force Survey (LFS) 2007–2008.[11] The indicator for domestic migrants is calculated as the number of internal migrants from a district per 100 people surveyed from the district.

3.3. Summary statistics and analysis of relative magnitudes

Table 1 shows summary statistics for the sample. Over the full sample, the average share of districts with any COVID-19 cases

---

[6] For more details, see: https://www.mohfw.gov.in/pdf/Traveladvisory.pdf.

[7] For Bangladesh, district-day level data were not released for two of the 45 days. We impute missing data for May 14 using data on May 13 and missing data for May 18 using data on May 17.

[8] See https://www.iedcr.gov.bd.

[9] See https://covindia.com. These data are highly correlated with data released by the Ministry of Health and Family Welfare, but the government data are not available on a daily basis.

[10] See http://covid.gov.pk/stats/pakistan. District-level data were retrieved on April 27, 2020.

[11] The 2007–2008 round was used because of the availability of information on the previous district of the migrant. In this dataset, about 10% of individuals aged 10 and above report having moved from a different district to their current district of residence. This rate (10%) is unchanged in the 2014–2015 round, which uses an identical question to identify migrants.

Table 1
Summary Statistics.

	Mean	Standard Deviation	N
Panel A: Bangladesh (April 6–May 20)			
Any Cases	0.84	0.37	2,835
Cases per Million People	19.17	42.23	2,835
Domestic Migrants per 100 People	1.18	1.03	63
International Migrants per 100 People	2.41	2.38	63
Population (millions)	2.10	1.23	63
Population Density (thousands per sq km)	1.01	0.55	63
Fraction Urban Population	0.17	0.07	63
Fraction Below Poverty Line	0.33	0.12	63
Hospital Beds (per million)	0.03	0.02	63
Number of Testing Labs	0.33	0.62	63
Panel B: India (March 15–April 28)			
Any Cases	0.35	0.48	26,145
Cases per Million People	2.56	7.72	26,145
Domestic Migrants per 100 People	4.88	3.10	581
International Migrants per 100 People	0.24	0.67	581
Population (millions)	2.17	1.65	581
Population Density (thousands per sq km)	1.32	5.70	581
Fraction Urban Population	0.24	0.18	581
Fraction Below Poverty Line	0.52	0.24	581
Primary Health Centers per Million People	28.07	22.27	581
Number of Testing Labs	0.42	1.34	581
Panel C: Pakistan (March 16–April 27)			
Any Cases	0.60	0.49	4,773
Cases per Million People	11.81	21.74	4,773
Domestic Migrants per 100 People	0.57	0.82	111
International Migrants per 100 People	1.32	1.83	111
Population (millions)	1.49	1.34	111
Population Density (thousands per sq km)	0.44	0.54	111
Fraction Urban Population	0.21	0.14	111
Fraction Below Poverty Line	0.38	0.15	111
Fraction Health Access	0.41	0.24	111
Number of Testing Labs	0.12	0.53	111

Notes: State/Provincial/Country capitals have been omitted. The sample includes all days for which we have available data since the 100th COVID-19 case was reported in each country.

was 0.84 in Bangladesh, 0.35 in India, and 0.60 in Pakistan. Cases per million people are much higher in Bangladesh and Pakistan, at 19.17 and 11.81 respectively, while the number of cases per million people is 2.56 in India. Data sources and definitions for control variables are reported in Appendix B.

Bangladesh and Pakistan also report much higher rates of international migration (2.41 and 1.32 migrants per 100 people, respectively) relative to India (0.24 migrants per 100 people). Conversely, rates of domestic migration are much higher in India data than in Bangladesh or Pakistan, with Pakistan having substantially lower rates of domestic migration than either Bangladesh or India.

Some of the estimates in Section 5 can be interpreted in terms of 1 standard deviation variations in district-wise out-migration rates. Table 1 shows that for Bangladesh, 1 SD corresponds to 2.38 migrants on a base of 2.41 migrants per 100 people (a 99% ratio). For India, 1 SD in international out-migration corresponds to a change of 0.67 migrants on a base of 0.24 migrants per 100 people (i.e., a 279% ratio). In Pakistan, 1 SD corresponds to a change of 1.83 migrants on a base of 1.32 migrants per 100 people (a 139% ratio).

The summary statistics also provide relative magnitudes of cases plausibly experienced by returning migrants versus those due to broader community spread. The magnitudes suggest the presence of substantial community spread beyond infections of migrants themselves.

For example, in India, the summary statistics show an average of 0.24 international migrants per hundred people in the survey data. The coronavirus tracker shows 2.56 cases per million people – or 0.000256 cases per hundred people – during the study window. If 3% of the migrants returned and 1% were infected with COVID-19 (a conservative assumption given that the window is early in the pandemic), then $0.24 * 3\% * 1\% = 0.000072$ infected migrants returned per hundred people.[12] Dividing 0.000072 infected migrants by 0.000256 cases yields that 28% of the reported cases could plausibly be accounted for by infected migrants.

In Pakistan, there are 1.32 international migrants per hundred people and 11.81 cases per million people – or 0.001181 cases per hundred people. If 3% of the migrants returned and 1% had COVID, then $1.32 * 0.03 * 0.01 = 0.000396$ infected migrants returned per hundred people. This would account for about 34% of cases in our study window. The calculation rests on conservative assumptions and shows that it is implausible that the correlations are solely due to infections of migrants themselves. By the final 2 week period in the sample window, the corresponding calculation suggests that international migrants could directly account for just 15% of cases in India and 18% of cases in Pakistan.

4. Empirical strategy

To frame negative externalities from global and local coronavirus containment policies, we combine district-day level data on COVID-19 cases with district-level data on international and domestic out-migration and covariates. We start by predicting variation within a given day across districts using a day fixed-effects model that starts on the day of the 100th confirmed case in each country.

For district i and day t, where t is the number of days since the 100th case was confirmed in each country, we estimate the relationship between international out-migration and COVID-19 with a linear specification for each country:

$$Y_{it} = \beta_0 + \beta_1 InternationalOutmigration_i + \beta_2 DomesticOutmigration_i + \mathbf{X_i}\beta + \alpha_t + \varepsilon_{it} \quad (1)$$

where Y_{it} is either (a) an indicator term equal to 1 if the district had any COVID-19 cases on day t and 0 otherwise, or (b) the number of COVID-19 cases per million people in the district on day t. The main coefficient of interest is β_1, the predictor of COVID-19 cases related to international out-migration, conditional on the control variables and fixed effects. $InternationalOutmigration_i$ is the number of migrants who had previously left district i for another country per 100 people as calculated from the household and labor surveys for India, Pakistan, and Bangladesh. $DomesticOutmigration_i$ is the number of migrants who left district i for another district in the same country per 100 people as calculated from the survey data for the three countries. $\mathbf{X_i}$ is a set of district-level control variables comprising population, population density, the fraction of the population residing in urban areas of the district, the fraction of the population below the poverty line, and a measure of access to health facilities.[13] Country-wide trends are captured flexibly by

Table 2
Any COVID-19 Cases in District.

	All Days				Days 29–42	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Bangladesh (April 6–May 20)						
International Migrants	0.045**	0.027	0.041**	0.012	0.005	0.002
per 100 People (z-score)	(0.018)	(0.017)	(0.018)	(0.016)	(0.004)	(0.003)
Domestic Migrants			0.017	0.022	−0.008	−0.009
per 100 People (z-score)			(0.014)	(0.014)	(0.008)	(0.008)
R^2	0.389	0.427	0.391	0.430	0.026	0.044
Dependent Variable Mean	0.84	0.84	0.84	0.84	0.99	0.99
Observations	2,835	2,835	2,835	2,835	882	882
Panel B: India (March 15–April 28)						
International Migrants	0.112***	0.087***	0.108***	0.081***	0.112***	0.084***
per 100 People (z-score)	(0.010)	(0.009)	(0.010)	(0.010)	(0.011)	(0.012)
Domestic Migrants			0.034**	0.040***	0.046**	0.053**
per 100 People (z-score)			(0.014)	(0.013)	(0.019)	(0.019)
R^2	0.217	0.328	0.222	0.334	0.069	0.209
Dependent Variable Mean	0.35	0.35	0.35	0.35	0.54	0.54
Observations	26,145	26,145	26,145	26,145	8,134	8,134
Panel C: Pakistan (March 16–April 27)						
International Migrants	0.137***	0.013	0.117***	0.005	0.133***	−0.021
per 100 People (z-score)	(0.026)	(0.033)	(0.024)	(0.033)	(0.026)	(0.028)
Domestic Migrants			0.081***	−0.058*	0.150***	−0.024
per 100 People (z-score)			(0.028)	(0.029)	(0.028)	(0.024)
R^2	0.175	0.417	0.202	0.426	0.240	0.598
Dependent Variable Mean	0.6	0.6	0.6	0.6	0.68	0.68
Observations	4,773	4,773	4,773	4,773	1,554	1,554
Controls		✔		✔		✔
Day Fixed Effects	✔	✔	✔	✔	✔	✔

Notes: State/Provincial/Country capitals are omitted. The sample for columns (1)–(4) includes all days for which data are available since the 100th COVID-19 case was reported in each country. The sample for columns (5)–(6) includes days 29–42 since the 100th COVID-19 case was reported in each country. All regressions include day fixed effects, while columns (2), (4), and (6) additionally include the following district-level controls: population, population density, the fraction of urban population, the fraction of population below the poverty line, and a measure of health access (Bangladesh: number of hospital beds per capita, India: number of primary health centers per capita, Pakistan: percentage of population which has access to a health clinic or hospital within 15 min of their dwelling). Standard errors in parentheses and double-clustered by day and district. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

day fixed-effects via the variable α_t. Standard errors are clustered two-way by district and day.

We also estimate Eq. (1) in successive 2-week windows to study the relationships between international out-migration and the spread of COVID-19 over time. These results are shown in Section 5 and in Appendix Tables 6–8. The final window provides results after international travel bans were in force and after testing protocols were broadened to include all symptomatic individuals with respiratory issues in the three countries.

The second specification flexibly accounts for changing relationships over time. We exploit the panel structure of the district-day level data, using day fixed effects to exploit variation within a given day across districts:

$$Y_{it} = \beta_0 + \sum_{t=1}^{T} \theta_t InternationalOutmigration_i * Day_t$$
$$+ \sum_{t=1}^{T} \gamma_t DomesticOutmigration_i * Day_t + \mathbf{X_i}\beta + \alpha_t + \varepsilon_{it} \quad (2)$$

where $Y_{it}, InternationalOutmigration_i, DomesticOutmigration_i, \mathbf{X_i}$, and α_t are as defined earlier. $\{Day_t\}_{t=1}^{T}$ is a set of indicator variables equal to 1 if the day is equal to t, and 0 otherwise. $\{\theta_t\}_{t=1}^{T}$ is a set of coefficients of interest capturing the relationship between international out-migration and COVID-19 cases on day t, while $\{\gamma_t\}_{t=1}^{T}$ is an analogous set of coefficients for domestic out-migration. Again, all standard errors are clustered two-way by district and day.

The pattern of the coefficients θ_t and γ_t show the shifting predictive power of migration pattern on the spread of COVID-19. If

the return of migrants to districts seeded cases of COVID-19 that led to increases in cases through community transmission over time, we would expect the coefficients θ_t and γ_t to increase during the period, even after accounting for day fixed effects.

To capture the effect of domestic migration outward from large cities, we restrict attention to districts without state, provincial, and country capitals. For consistency, we do the same when analyzing associations of international migration. We assess robustness of the results to the inclusion of these districts in Appendix Section C.2. Including the full set of districts strengthens the results on international migration in Pakistan, very slightly weakens them in India, and leaves a mixed picture in Bangladesh.

5. Results

The results connect patterns of migration by district to cases of COVID-19 in the first 1.5 months after the first 100th case in each country. First, we predict whether or not a district reports any cases. Next, we predict the number of cases per million people.

5.1. Predicting the Incidence of COVID-19

In all three countries, an increase in international out-migration predicts diffusion of COVID-19 on the extensive margin—i.e., they predict a higher probability that a district reports any cases—but the results are sensitive to the addition of controls. Table 2 presents the relationship between international out-migration and COVID-19 cases using empirical specification (1). The results for Bangladesh, India, and Pakistan are reported in panels A, B, and C, respectively.
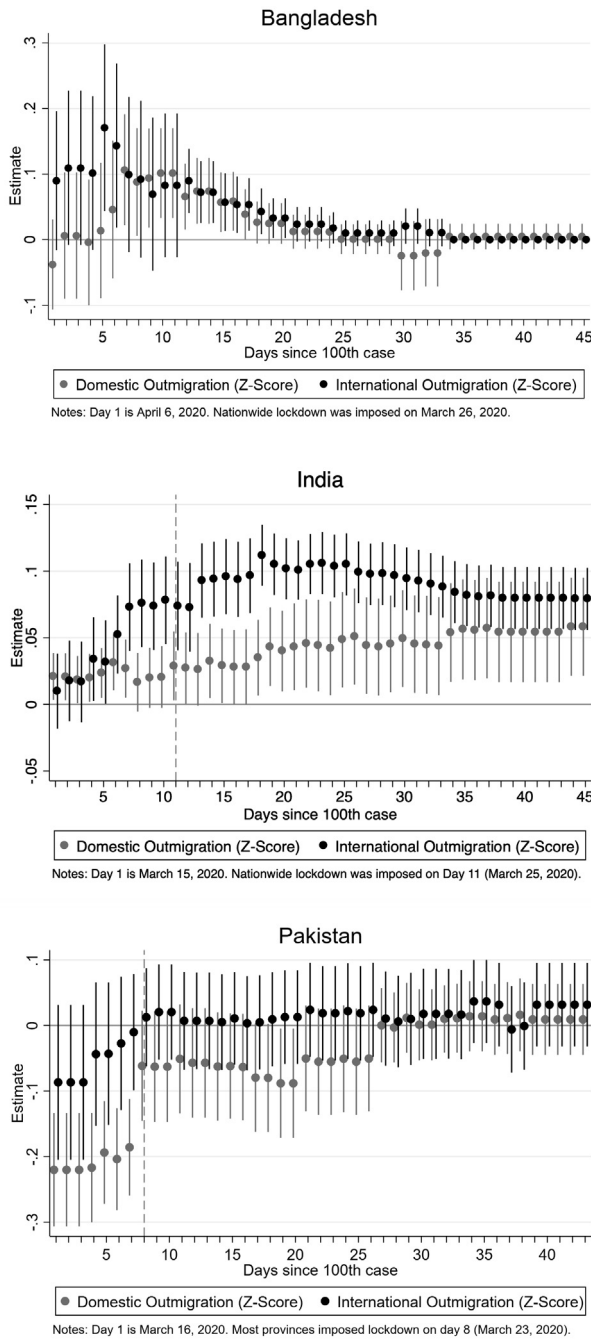
Fig. 1. Relationship between Out-migration & Probability District Reports Any Cases. Notes: Dependent variable: Indicator variable equal to 1 if the district had any COVID-19 cases on day t and 0 otherwise. Coefficients plotted in black (θ_t in Eq. (2)) illustrate the relationship between international out-migration and an indicator for any cases in the district on day t. Coefficients plotted in gray (γ_t in Eq. (2)) illustrate the relationship between domestic out-migration and an indicator for any cases in the district on day t. State/Provincial/Country capitals have been omitted. The regressions include district-level controls: population, population density, the fraction of urban population, the fraction of population below the poverty line, and a measure of health access (Bangladesh: number of hospital beds per capita, India: number of primary health centers per capita, Pakistan: percentage of population which has access to a health clinic or hospital within 15 min of their dwelling) and day fixed effects. Standard errors are double-clustered by day and district. 95% confidence intervals shown.

In Bangladesh, COVID-19 had spread widely across districts during the period, and on average during the period 84% of districts had confirmed cases. Even with the high average, column (1) shows that a 1 SD increase in international out-migration is asso-

ciated with a relatively large 5% (0.045/0.84) increase in the probability of any district-wise COVID-19 cases. In India, the association is 32%, and it is 23% in Pakistan. These results are without controls and $p < 0.05$ for each coefficient.

The prevalence of international migration is in part a proxy for other attributes of districts, however, and Column (2) shows that adding controls for district size, demography, health facilities and poverty levels diminishes net associations. In Bangladesh, the association falls to 3% (0.027/0.84). In India, the association falls to 25% (0.087/0.35) and remains statistically significant at the 1% level, and in Pakistan the association is now essentially zero (coefficient = 0.013 with a standard error of 0.033).

The pattern is robust to adding the rate of domestic migration to the specification following Eq. (1). The coefficients on the international migration variables in columns (3) and (4) are of a similar order but slightly smaller than in columns (1) and (2), and levels of statistical significance are unchanged. The results also show that international and domestic migration differ in their associations with COVID diffusion. In India, the coefficient on domestic migration is one third to half as large as the coefficient on international migration. In Pakistan, the coefficients are similarly-sized in Column (3) and flip signs in column (4).[14]

As noted in Section 2, part of the positive correlations can be explained by testing protocols and the lack of travel restrictions in the beginning of the samples. Columns (5) and (6) thus present results restricted to two-week periods at the end of the sample window, after international travel restrictions had been in force and testing protocols had broadened.[15]

In Bangladesh, 99% of districts had reported cases by this point, and there is so little variation in the dependent variable that coefficients on the migration measures are close to zero. In India, however, the coefficients increase in size, with and without controls (Panel B, columns 5 and 6). In Pakistan the coefficients also increase before controls are added (Panel C, column 5), but, as in columns (2) and (4), adding controls in column (6) eliminates the predictive power of the migration variables in explaining the extensive margin of COVID-19 diffusion. The column (5) results for India and Pakistan are consistent with community spread seeded by returning migrants.

Fig. 1 summarizes these patterns, presenting the development of the relationship between out-migration and the probability that a district reports any cases of COVID-19. Coefficients plotted in black depict the relationship between international out-migration and an indicator for any cases over time. Coefficients plotted in gray illustrate the relationship for domestic out-migration; 95% confidence intervals are shown.

The three panels are broadly consistent with the coefficients in column (4) of Table 2. For India, the domestic out-migration coefficients are positive but smaller than the international out-migration coefficients. For Bangladesh and Pakistan, the coefficients are generally statistically indistinguishable from zero, although they are positive and statistically significant between days 7 and 17 in Bangladesh, and negative and statistically significant before day 8 in Pakistan.

Taking everything together on the extensive margin: The results for India are strong and robust. But the non-results for Bangladesh, and the sensitivity to specification of the results for Pakistan, suggest that migration may more reliably predict COVID-19 on the intensive margin than the extensive margin.

---

[14] Demographers note that the mobility of migrants in Pakistan in early 2020 was considerably lower than in India (The Economist, 2020).

[15] Appendix D presents the relationships between out-migration and cases separately for each of the first, second, and third 14-day periods since the 100th case was reported in each country.

Table 3
Cases per Million People in District.

	All Days				Days 29–42	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Bangladesh (April 6–May 20)						
International Migrants	4.492	−4.122	5.621*	−4.319	9.734	−6.082
per 100 People (z-score)	(2.777)	(2.968)	(3.285)	(3.238)	(5.548)	(5.794)
Domestic Migrants			−5.276	0.629	−9.019	0.126
per 100 People (z-score)			(4.508)	(2.147)	(7.344)	(3.578)
R^2	0.133	0.569	0.148	0.569	0.060	0.719
Dependent Variable Mean	19.17	19.17	19.17	19.17	32.14	32.14
Observations	2,835	2,835	2,835	2,835	882	882
Panel B: India (March 15–April 28)						
International Migrants	1.592***	1.253***	1.604***	1.235***	2.264***	1.565**
per 100 People (z-score)	(0.452)	(0.460)	(0.450)	(0.458)	(0.627)	(0.662)
Domestic Migrants			−0.100	0.119	−0.239	0.138
per 100 People (z-score)			(0.198)	(0.212)	(0.405)	(0.426)
R^2	0.138	0.188	0.138	0.188	0.059	0.167
Dependent Variable Mean	2.56	2.56	2.56	2.56	5.11	5.11
Observations	26,145	26,145	26,145	26,145	8,134	8,134
Panel C: Pakistan (March 16–April 27)						
International Migrants	6.657***	6.023***	6.816***	5.659**	14.584***	11.678**
per 100 People (z-score)	(1.800)	(2.174)	(1.838)	(2.130)	(3.299)	(3.972)
Domestic Migrants			−0.616	−2.744**	−0.673	−4.653**
per 100 People (z-score)			(0.915)	(1.255)	(1.698)	(2.120)
R^2	0.280	0.323	0.281	0.333	0.301	0.390
Dependent Variable Mean	11.81	11.81	11.81	11.81	22.00	22.00
Observations	4,773	4,773	4,773	4,773	1,554	1,554
Controls		✔		✔		✔
Day Fixed Effects	✔	✔	✔	✔	✔	✔

Notes: State/Provincial/Country capitals are omitted. The sample for columns (1)–(4) includes all days for which data are available since the 100th COVID-19 case was reported in each country. The sample for columns (5)–(6) includes days 29–42 since the 100th COVID-19 case was reported in each country. All regressions include day fixed effects, while columns (2), (4), and (6) additionally include the following district-level controls: population, population density, the fraction of urban population, the fraction of population below the poverty line, and a measure of health access (Bangladesh: number of hospital beds per capita, India: number of primary health centers per capita, Pakistan: percentage of population which has access to a health clinic or hospital within 15 min of their dwelling). Standard errors in parentheses and double-clustered by day and district. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.2. Predicting COVID-19 cases

The results in India and Pakistan are clearer and stronger on the intensive margin. Table 3 presents predictions of COVID-19 cases per million people in each district estimated using Eq. (1).

Panel A shows that the relationship between international out-migration and cases per capita in Bangladesh is measured imprecisely with or without controls. The number of cases per capita was relatively high in the sample window (more than seven times India's rate), however, and the coefficients are large. The coefficient in column (1) of Panel A implies that a 1 SD increase in international out-migration in Bangladesh is associated with a 23% increase in the district-wise average number of cases per million. The coefficient flips from positive to negative when controls are added in column (2), although with a relatively large standard error. A similar pattern is repeated in columns (3) and (4) and in columns (5) and (6).

In India and Pakistan, in contrast, the predictions are far more precisely estimated and are robust across specifications.[16] Column (1) shows that, without controls, a 1 SD increase in international out-migration in India is associated with a 62% (1.592/2.56) increase in the district-wise average of cases per capita. In Pakistan, the association is 56% ($p < 0.01$ for each coefficient). The results fall only slightly when adding controls for other district characteristics. Column (2) shows that the association falls to 49% (1.253/2.56) in India, and it decreases to 51% (6.023/11.81) in Pakistan. The estimates remain statistically significant at the 1% level. Columns (3) and (4)

show that the results are also robust to the inclusion of measures of domestic migration.[17] The estimates in column (4) are the most conservative—with controls and the inclusion of measures of domestic migration—but the predictions remain large: 1 SD increase in international out-migration in India is associated with a 48% (1.235/2.56) increase in the district-wise average of cases per capita. In Pakistan, the corresponding figure is also 48% (5.659/11.81).

As in Section 5.1, in columns (5) and (6) we restrict attention to the final two-week sub-sample (days 29–42) after international travel restrictions were in place and testing protocols were broadened. Column (5) shows that, without controls, a 1 SD increase in international out-migration in India is associated with a 44% (2.264/5.11) increase in the district-wise average of cases per capita. With the inclusion of controls in Column (6), this association falls to 31%. In Pakistan, the associations are 66% and 53%, without and with controls, respectively.

The relatively large size of these predictions in India and Pakistan is echoed in Fig. 2. The figure presents the development of the relationship between out-migration and COVID-19 cases per million people using Eq. (2). Again, coefficients plotted in black depict the relationship between international out-migration and cases per million people over time. Similarly, coefficients plotted in gray illustrate the relationship between domestic out-migration and cases per million people over time. 95% confidence intervals are shown.

---

[16] The results are also robust to controlling for the number of testing labs in the country and the inclusion of state, provincial, and country capitals (Appendix C).

[17] While domestic migration rates are not very predictive of increases in cases per capita in India, column (4) shows that a 1 SD increase in the rate of domestic out-migration in Pakistan is associated with a 23% (-2.744/11.81) decrease in the district-wise average number of cases per capita.
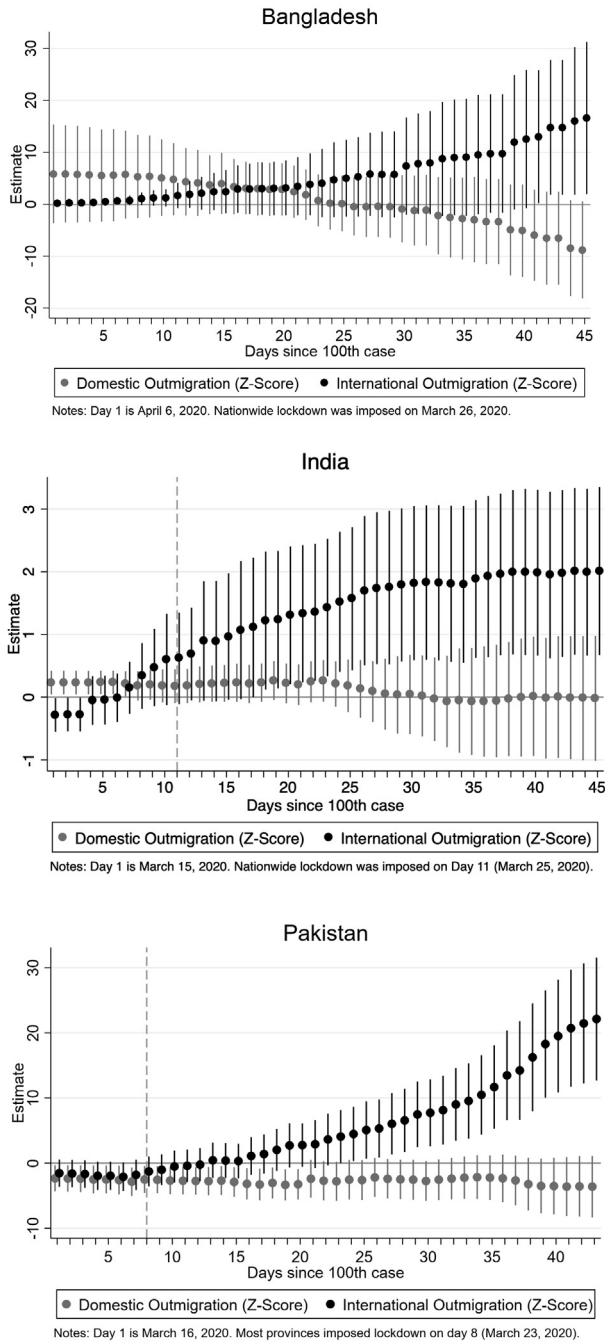
Fig. 2. Relationship between Out-migration & Cases per Million People. Notes: Dependent variable: COVID-19 cases per million people on day t. Coefficients plotted in black (θ_t in Eq. (2)) illustrate the relationship between international out-migration and cases per million people on day t. Coefficients plotted in gray (γ_t in Eq. (2)) illustrate the relationship between domestic out-migration and cases per million people on day t. State/Provincial/Country capitals have been omitted. The regressions include district-level controls: population, population density, the fraction of urban population, the fraction of population below the poverty line, and a measure of health access (Bangladesh: number of hospital beds per capita, India: number of primary health centers per capita, Pakistan: percentage of population which has access to a health clinic or hospital within 15 min of their dwelling) and day fixed effects. Standard errors are double-clustered by day and district. 95% confidence intervals shown.

The figure shows that the predictive power of domestic out-migration changes only minimally. The point estimates are small and slightly negative by the last day for which we have data for the three countries (and statistically indistinguishable from zero).

In contrast, the relationship between international out-migration and cases per capita grows over time in each country. For India, the coefficients are positive and statistically different from zero by day 17, and they continue to increase over time. For Pakistan, the coefficients are positive and statistically different from zero by day 25. For Bangladesh, the coefficients are positive but less precisely estimated, becoming statistically different from zero after day 40.

A back-of-the-envelope calculation helps to frame the magnitudes. For Pakistan, we begin with the change in the point estimates for international out-migration (23.72) relative to the change in cases per capita (31.56 cases) between day 1 and T (the last day for which we have data for each country). That difference implies that for a district that starts with the average level of out-migration, a one standard deviation increase in out-migration predicts an increase in COVID-19 cases per capita equivalent to 75% of the actual increase that was experienced on average across the sample. For India, the change in the point estimates for international out-migration (2.29) relative to the change in cases per capita (7.17) implies a result of 32% in a parallel comparison. For Bangladesh, the change in the point estimates for international out-migration (16.36) relative to the change in cases per capita (52.14) implies a figure of 31%. The estimates are large but most of the spread of COVID-19 comes from other sources.

6. Conclusion

The coronavirus outbreak started in China and soon spread, including to the Gulf, Europe and North America, countries in which migrants from South Asia were concentrated. Migrants returned home as the migrant-employing countries shut down their economies to protect their citizens. We show evidence consistent with important negative externalities for migrants' countries of origin as the COVID-19 virus spread to South Asia in early 2020.

International migration has been an important source of growth and opportunity in South Asia, and we expect that it will continue to be. But the data suggest that in early 2020 the movement of migrants back to their homes substantially increased health risks in the region. We find that 1 standard deviation increases in prior international out-migration in India and Pakistan (relative to the cross-sectional average across districts in each country) predicts increases of 48% in cases per capita.

The evidence shows that survey data on migration in previous years systematically predicts patterns of confirmed COVID-19 cases in 2020, especially in India and Pakistan. The evidence allows us to distinguish between domestic and international migration, making it complementary to real-time geo-coded data from mobile telephones and other sources that document contemporaneous population movements but lack detail on movers' identities (e.g., Milusheva, 2020).

The predictive power of the survey data should be helpful for health officials allocating resources to prepare for future global pandemics. The estimates are not causal, and could reflect other factors beyond migration rates, but they are consistent with the established epidemiological shape of infectious disease (Sattenspiel and Lloyd, 2009).

Bangladesh, India, and Pakistan share common elements of history and culture, and their governments have taken broadly similar policy decisions to contain the virus. Each imposed systematic quarantines of migrants as an important tool to slow the growth of the pandemic, but quarantines were largely voluntary. Our ability to predict COVID-19 incidence with migration data suggests that quarantines for migrants were not followed systematically, especially in the critical early months of the pandemic.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jpubeco.2020.104312.

References

Abi-Habib, M., Yasir, S., 2020. For India's Laborers, Coronavirus Lockdown Is an Order to Starve. The New York Times. https://www.nytimes.com/2020/03/30/world/asia/coronavirus-india-lockdown.html (accessed 31 March, 2020).

Ahsan, R., Iqbal, K., Khan, M., Mobarak, A.M., Shonchoy, A., 2020. Using Migration Patterns to Predict COVID-19 Risk Exposure in Developing Countries. Y-RISE Policy Brief, April 26.

Al Jazeera, 2020. Outrage after India asks migrants to pay for train journey. Al Jazeera (online).

Arain, A.L., 2020. Pakistan: Karachi's Province set for COVID-19 Lockdown. Anadolu Agency.

ARY Web, 2020. CM Punjab Usman Buzdar Announces Lockdown in Punjab. ARY News.

Bethune, Z.A., Korinek, A., 2020. Covid-19 infection externalities: trading off lives vs. livelihoods. Natl. Bureau Econ. Res. 27009.

Bisht, A., 2020. Chaos and hunger amid India coronavirus lockdown: India's strict lockdown of 1.3 billion people disrupts lives with migrant workers facing hunger and forced to walk home. Al Jazeera (online).

Dahir, A.L., 2020. Instead of Coronavirus, the Hunger Will Kill Us.' A Global Food Crisis Looms. The New York Times.

Daily Star, 2020. No flights from Europe till March 31, except UK: Foreign Minister. Daily Star.

Denis, E., Telle, O., Benkimoun, S., 2020. Mapping the lockdown effects in India: how geographers can contribute to tackle Covid-19 diffusion. The Conversation (online).

Dhaka Tribune, 2020. Covid-19: 10 million mobile subscribers left Dhaka. Dhaka Tribune.

Ellis-Petersen, H., Chaurasia, M., 2020. India racked by greatest exodus since partition due to coronavirus, Guardian (UK).

Emont, J., Shah, S., 2020. Coronavirus is spreading at religious gatherings, ricocheting across nations. Wall Street J.

International Monetary Fund, 2020. Policy Responses to COVID-19: Policy Tracker. https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#I.

Malik, K., Meki, M., Morduch, J., Ogden, T., Quinn, S., Said, F., 2020. COVID-19 and the Future of Microfinance: Evidence and Insights from Pakistan. Oxford Rev. Econ. Policy.

Milusheva, S., 2020. Using Mobile Phone Data to Reduce Spread of Disease. World Bank Policy Research Working Paper.

Mitra, R., Rawat, C., Varshney, G., 2020. Return migration and COVID-19: Data suggests Kerala, TN, Punjab, UP, Bihar may be future red zones for contagion risk. FirstPost.

Pandey, G., 2020. Coronavirus in India: Migrants running away from quarantine. BBC News (Delhi).

Ray, D., Subramanian, S., 2020. India's lockdown: an interim report. Natl. Bureau Econ. Res. 27282.

Sattenspiel, L., Lloyd, A., 2009. The Geographic Spread of Infectious Diseases: Models and Applications, vol. 5. Princeton University Press.

Sen, A., Rajan, R., Banerjee, A., 2020. Huge numbers may be pushed into dire poverty or starvation...we need to secure them. The Indian Express.

Shonchoy, A., 2020. Data-Driven Identification of Migration Corridors and Potential Outbreak Hotspots for Covid-19 in Bangladesh. Florida International University, Mimeo.

SNS Web, 2020. Coronavirus pandemic: Pakistan to extend lockdown for 2 more weeks as death toll reaches 31. The Statesman.

The Economist, 2020. Is Pakistan really handling the pandemic better than its rival? The Economist.

Thomsen, P., 2020. Europe's COVID-19 Crisis and the Fund's Response. International Monetary Fund, blog post. https://blogs.imf.org/2020/03/30/europes-covid-19-crisis-and-the-funds-response/.

UNB, 2020. Face corona 'war' with compassion. The Independent, Dhaka.

United Nations DESA, 2019. International migrants numbered 272 million in 2019, continuing an upward trend in all major world regions. Population Facts, 2019/4.

Viscusi, W.K., 2020. What Is a Life Worth? COVID-19 and the Economic Value of Protecting Health. Foreign Affairs.

World Bank, 2020. COVID-19 Crisis Through a Migration Lens. Migration and Development Brief, no. 32. https://openknowledge.worldbank.org/handle/10986/33634.

World Health Organization, 2020. Coronavirus disease 2019 (COVID-19) Situation Report – 73. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_6.