

FEATURED ARTICLE



Using international migration links for early detection of COVID-19 risk exposure in low- and middle-income countries

Mahreen Khan¹ | Reshad Ahsan² | Kazi Iqbal³ |
Ahmed Mushfiq Mobarak⁴ | Abu S. Shonchoy⁵

¹Oxford Martin School, University of Oxford, Oxford, UK

²Department of Economics, University of Melbourne, Melbourne, Australia

³Industry and Physical Infrastructure Division, Bangladesh Institute of Development Studies, Dhaka, Bangladesh

⁴Department of Economics, Florida International University, Miami, Florida, USA

⁵Department of Economics, Yale University, New Haven, Connecticut, USA

Correspondence

Mahreen Khan, University of Oxford, Oxford, UK.

Email: mahreen.khan@oxfordmartin.ox.ac.uk

Ahmed Mushfiq Mobarak, Yale University, School of Management, New Haven, Connecticut, USA.

Email: ahmed.mobarak@yale.edu

Editor in charge: Alex Winter-Nelson

Abstract

Reliable testing data for new infectious diseases like COVID-19 is scarce in developing countries making it difficult to rapidly diagnose spatial disease transmission and identify at-risk areas. We propose a method that uses readily available data on bi-lateral migration channels combined with COVID-19 cases at respective migrant destinations to construct a spatially oriented risk index. We find significant and consistent association between our measure and various types of outcomes including actual COVID-19 cases and deaths, indices of government policy responses, and community mobility patterns. Results suggest that future pandemic models should incorporate migration-linkages to predict regional socio-economic and health risk exposure.

KEYWORDS

coronavirus, COVID-19, international migration, labor migration, labor mobility, migrant workers

JEL CLASSIFICATION

F22, J11, I19, O10

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Applied Economic Perspectives and Policy* published by Wiley Periodicals LLC on behalf of Agricultural & Applied Economics Association.

The rapid proliferation of infectious diseases such as COVID-19 can pose substantial threats to the economies of low- and middle-income countries (LMICs) both from public health and socio-economic perspectives. The crisis can be compounded by limitations in health system capacity and social safety nets (Ivers & Walton, 2020, p. 1149; Walker et al. 2020, pp. 413–414; Tondl, 2021). For example, an early evaluation of global health systems revealed that there were fewer than 2000 working ventilators to serve the hundreds of millions of people in Africa (MacLean & Marks, 2020). Similarly, COVID-19 was also associated with widespread income drop combined with large increases in food insecurity in many LMICs in Africa, Asia, and Latin America (Egger et al., 2021; Ahmed et al., 2021).

To mitigate the cost of a pandemic's shock to health and economic systems, policy makers in LMICs need to employ various tools to swiftly identify the spatial distribution of vulnerable regions at some reasonable level of granularity such that they can target economic and health resources efficiently. Similarly, at the global level, a mapping of at-risk countries can be helpful in directing the flow of development aid. Deficiencies in health-care infrastructure and resources in the LMIC context can prevent rapid deployment of testing services for early disease identification, thus creating a multi-pronged attack on the lives of the economically vulnerable.

Considering this imminent need amongst regional and global policy makers to swiftly identify vulnerable countries and regions at the onset of virally transmitted pandemics such as COVID-19, we propose a novel approach to address this problem by exploiting migration-based linkages between countries to make predictions about the spatial risk distribution stemming from the contagion. This tool can serve as a proof-of-concept to infer the future exposure to socio-economic and health related vulnerabilities for LMICs where large-scale migration is used as an active labor market policy.

Our COVID-19 related risk measurement approach relies on the human-to-human transmission of viral pandemics that lead to disease spread in new locations. Past pandemics and epidemics of infectious diseases like HIV, SARS and MERS have been slinked to human mobility and migration patterns (Greenaway & Gushulak, 2017, p. 316). Correspondingly, in the context of COVID-19's exceptionally infectious nature (Guan et al., 2020, p. 428), pre-existing bilateral migration links with COVID-19 affected areas can be informative about disease risk in new locations, particularly in the migrants' countries of origin. This is partly because travelers returning to countries of origin, many of whom were migrants working under various types of labor contracts, were an inadvertent vector for early COVID-19. Lee et al. (2021) established the link between migration and the early spread of COVID-19 across national borders and across regions for three countries, namely, Bangladesh, India, and Pakistan (Lee et al., 2021).

After constructing our migration-linked disease exposure index, we validate the reliability of the measure by comparing our risk predictions to the number of confirmed COVID-19 cases in the first five weeks of the COVID-19 pandemic. We find a strong positive correlation between our index and confirmed COVID-19 cases—a 1% increase in our COVID-19 risk exposure measures is predicted to significantly increase confirmed COVID-19 cases by about 0.2% with a 2 week-lag from exposure at the migrants' destination countries. We also carry out robustness checks using various indicators of the severity of the disease, namely, a wide range of governments' response to the spread of infection (Hale et al., 2020); restrictions to citizen mobility (Google, 2020); and the number of COVID-19 deaths. While each of these indicators come with their respective set of limitations, and the size of the impact shows notable heterogeneity, we find consistently significant correlation with each of these separate indicators. This provides confidence in the reliability of our index as a risk-measurement tool. The strong predictive power of our index is retained even after controlling for a large set of country and week fixed effects. Furthermore, we also conduct some sensitivity analysis

using receiver operating characteristics (ROC) graphs for different cut-off values of our index and its ability to predict substantial and high-transmission regions using the United States Center for Disease Control (CDC) Indicator of Community Transmission (CDC, 2022) and find reliable results (see Figure A1).

We provide a simple application of our method for the sub-regional COVID-19 risk analysis for Bangladesh, a country characterized both by large annual labor out-migration along with severely under-resourced health systems (UNDP, 2020). We use survey and administrative data to create district and sub-district (*upazila*) indices. These provide important insights for local policy makers in the early stages of the pandemic. Our hypothesis linking COVID-19 and human mobility complements a related study where authors find that respondents in Bangladeshi communities, where at least one migrant returned in the 2 weeks prior to the survey, were significantly more likely to report one or more symptoms associated with COVID-19 (Lopez-Pena et al., 2020).

Many countries in the Global South experienced high rates of migration in the past decade (McAuliffe & Khadria, 2020) and have large emigrant populations residing in the high-income countries. Some notable high-frequency migrant destinations like Italy and the United States were affected by COVID-19 in the early stages of the pandemic (Kuchler et al., 2022). Our approach suggests that migrant-sending countries and regions having stronger links to these destination countries via their migrant populations are at higher risk of being exposed to the health and socio-economic consequences of the pandemic compared to countries with weaker links. Consequently, these vulnerable migrant-sending countries may need to pay strong attention to mitigate these effects. Returning migrant movements during pandemics could be triggered by several different factors such as the financial crises initiated by country-wide lockdowns, economic closures, and other uncertainties while also creating unintentional spread of the disease in their countries of origin (Guadagno, 2020). Since India had over 138,000 migrants in Italy in 2017 compared to Tanzania, which had only around 1600 migrants (WHO, 2022), our hypotheses would suggest that India, in particular the regions with higher migrant stocks to Italy, would be most at risk of exposure, relative to Tanzania after scaling for the appropriate population sizes in the countries of origin following migratory movements during a pandemic such as COVID-19.

Our paper contributes to several strands of research. Firstly, we contribute to a growing literature linking the impact of COVID-19 and degree of inter-connectedness between countries due to social connections and migratory movements (Chan et al., 2020; Kuchler et al., 2022; Lee et al., 2021; Milani, 2021, pp. 225–226). Secondly, our paper focuses on understanding pandemic risks for especially vulnerable populations: international migrants, their households, and their communities. We thus contribute to development and migration research by studying the socio-economic risks posed to migrant communities (Guadagno, 2020). Finally, by linking human mobility and risk of exposure to COVID-19, we contribute to the broader public health literature linking human mobility and population health (Castelli & Sulis, 2017, 284; Gushulak et al., 2009; Hirsch, 2014, pp. 42–43). In contrast to epidemiological studies that predict the evolution of the number of infected individuals in a population, we focus on using bi-lateral migration channels, driven by historic economic relationships between countries, to predict disease risk exposure and the subsequent socio-economic vulnerabilities of populations in developing countries, thus highlighting the importance of incorporating spatial elements in modeling future pandemics (Gatto et al., 2020, p. 10484).

In the remaining paper, we proceed as follows. First, we provide a theoretical framing that links migration with COVID-19 risk exposure for LMICs. Next, we provide a detailed description of the data and sample period for validating the reliability of our index. This is followed by

a summary of the results and discussion of the limitations. We then provide a sub-national application of our risk exposure to the case of Bangladesh to illustrate the applicability of our approach. Finally, we conclude with next steps and policy applications.

THEORETICAL FRAMEWORK LINKING MIGRATION AND COVID-19 RISK EXPOSURE

We theorize that migration-based links can provide important insights on exposure to viral pandemic risk exposure. We build on the migration and COVID-19 link established in Lee et al. (2021) to construct a disease risk index that predicts exposure for every LMIC country because of their migration-based links. As a starting point, we use their results that the number of COVID-19 cases in a migrant-sending region i at time t , COV_{it} , is proportional the total out-migration rate. This is illustrated by the following equation:

$$COV_{it} \propto \frac{Mig_{it}}{Pop_{it}}. \quad (1)$$

Building on these results, we construct an index of COVID-19 risk exposure by combining destination specific out-migration with COVID-19 case exposure. This index exploits the insight that past bi-lateral migration channels are linked to inter-country mobility patterns in the early days of COVID-19 pandemic and have predictive value for early detection of pandemic risk exposure.

Conceptually, we theorize that an LMIC's exposure to an infectious disease such as COVID-19, via migration channels, can depend on the number of return migrants from each destination country d , A_{id} , and the probability that each returning migrant from d is infected with COVID-19, φ_d . While we do not observe A_{id} directly for the countries in our sample, we can proxy for this using the total stock of migrants, M_{id} , from origin country i residing in destination d , prior to the pandemic. The key assumption is that the number of returning migrants from d to i in 2020 is proportional to the stock of pre-COVID migrants from i that reside in d , that is:

$$A_{id} \propto M_{id}. \quad (2)$$

In order to proxy φ_d , we assume that the infection probability of a returning migrant from d is an increasing function of the COVID-19 infection rate in d . That is, all else equal, a returning migrant from a country with a higher infection rate is more likely to be infected themselves. Thus, we use the number of COVID-19 infections per capita in a destination to proxy φ_d .

$$\varphi_d = \frac{COV_{dt}}{POP_d}. \quad (3)$$

With these proxies in hand, we define our LMIC's migration-based exposure to COVID-19 as:

$$EXP_{it} = \sum_{d=1}^D M_{id} \left(\frac{COV_{dt}}{POP_d} \right). \quad (4)$$

where i indexes migrant-origin LMICs and $d = 1, 2, \dots, D$ indexes migrant-receiving destinations. M_{id} is the total stock of migrants from country i residing in destination d (United Nations, 2017).¹

For each destination, we multiply country i 's stock of out-migrants, M_{id} , with the number of COVID-19 cases per capita in destination, d . COVID-19 infections, COV_{dt} , is the number of confirmed cases reported by the European Centre for Disease Prevention and Control (ECDC, 2022) in destination d on week t of 2020. We divide this by the total population in d in mid-2017, POP_d , (United Nations, 2017) to get a per capita rate.

Our COVID-19 risk index, EXP_{it} , is an increasing measure of COVID-19 exposure: a higher value is associated with greater COVID-19 infections in a destination country with which country i has strong migration links. Our index varies by origin country due to differences in pre-COVID-19 migration patterns. It also varies by time due to the evolution of COVID-19 cases in destination countries. The index value is thus determined not only by overall emigration rates (Lee et al., 2021), but also each country's migration links to specific destinations weighted by the COVID-19 cases, respectively. This allows us to create a nuanced assessment of the regional pandemic risk distribution for LMICs with non-trivial exposure to out-migration.

Our COVID-19 risk index can be used in conjunction with other economic data to inform policy responses at sub- and cross-national levels. In Figure 1a,b, we illustrate the cross-country variation in our index in South and Southeast Asia, and Africa, respectively. For example, in the case of Asia, amongst countries where COVID-19 was not widespread early in the pandemic (defined as having fewer than 2000 cases), our index predicted that India, the Philippines, and Vietnam are relatively more exposed as shown by the darker shades.

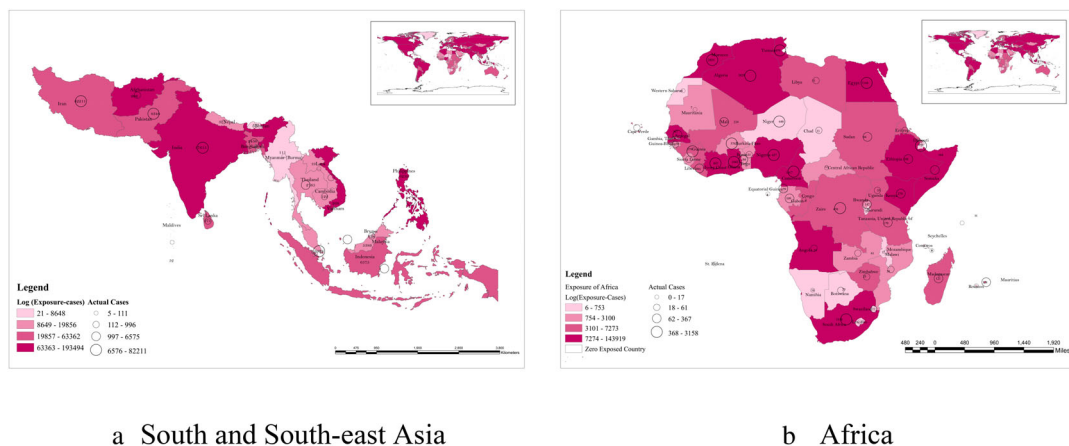


FIGURE 1 Region-wise COVID-19 risk exposure and case distribution using COVID-19 Index. (a) South and South-east Asia (b) Africa. Figures 1(a) and (b) show the heat maps based on the COVID-19 risk index for South and South-east Asia and Africa, respectively. The darkest shade indicates countries in the top-most quartile of the risk categorization according to the index. We also superimpose the cumulative number of COVID-19 cases reported in each country (as of March 24, 2020) with larger circles illustrating higher caseloads. Case data is taken from European Centre for Disease Prevention and Control (ECDC, 2022).

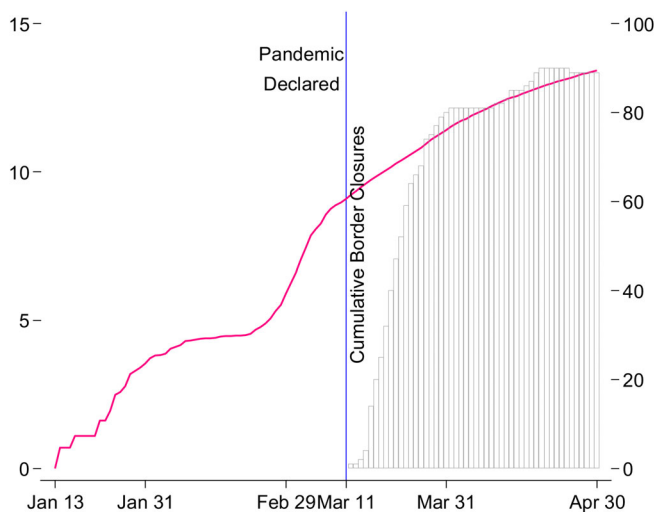


FIGURE 2 Cumulative global COVID-19 cases and Government Response Index across countries. The pink line denotes cumulative COVID-19 cases using data from European Centre for Disease Prevention and Control (ECDC, 2022) where the y-axis on the left reflects number of COVID-19 cases per 10,000. The bars denote the Government Response Index (GRI) as proxied by Hale et al. (2020) where the y-axis on the right reflects the degree border closures measured by the cumulative GRI measured as a percentage. The blue line marks the announcement of COVID-19 as a global pandemic by the World Health Organization (WHO). Data is included only for LMICs countries as per the study sample.

VALIDATION OF THE COVID-19 RISK INDEX

In the following section, we provide an overview of the data and methods to validate the reliability of our index in assessing actual and perceived COVID-19 risk intensity across countries.

Data

Data on COVID-19 cases and deaths for each country are collected from official government sources and reported by the European Centre for Disease Prevention and Control (ECDC, 2022). The pink line in Figure 2 illustrates the cumulative cases by region during the first four months of 2020. The cumulative numbers suppress the large heterogeneity across regions such as 125,000 cumulative cases in Latin America, South Asia, the Middle East, and North Africa compared to only 25,000 or fewer in South-East Asia, Sub-Saharan Africa, and the Pacific by April 30, 2020.

We obtain migration data (United Nations, 2017),² which reports the total stock of migrants between origin and destination countries using population censuses, population registers, and nationally representative surveys at the destination countries. We use 2017 data to calculate the stock of migrants from each origin-country to all possible destination countries. By using pre-COVID-19 migration data, we ensure that our exposure measure is not contaminated by endogenous changes that may impact the extent of return migration.

We restrict the countries of migrant origin in our sample to include only LMICs (World Bank, 2020). LMICs accounted for 79% of all out-migrants in 2017. In 2019, more than 40% of

the out-migrants originated only from Asia with India being the largest sender of migrants, followed by China, Bangladesh, and Pakistan. Meanwhile about two-thirds of all international migrants resided in high-income countries in 2019 (McAuliffe & Khadria, 2020). Therefore, international migrants who travelled from COVID-19 affected high-income countries back to their home countries at the early stages of the pandemic, had a high likelihood of originating from LIMCs. As a result, we restrict our sample of migrants to this group and subsequently, our migration-linked COVID-19 risk measure is only applicable for countries in the LMIC category (World Bank, 2020).

Our focus is to estimate whether return migration was a vector of transmission for socio-economically vulnerable countries, therefore, to construct our final sample, we restrict the set of migrant-origin countries to the LMICs. Consequently, our index measure is applicable only to these countries. We do not restrict destination countries, so that our out-migration data capture all possible destinations. We also use population data for each country from United Nations (2017).

In addition to data on COVID-19 cases and deaths, we use two other sources of data which we subsequently use as proxies to measure the intensity of the COVID-19. The first set is the Oxford COVID-19 Government Response Tracker (Hale et al., 2020), which capture government policies related to closure and containment, health, and economic policy for more than 180 countries, plus several countries' subnational jurisdictions. Policy responses are recorded on ordinal or continuous scales for 19 policy areas, capturing variation in degree of response to create four indices that group different families of policy indicators in the following areas: (i) Stringency index ("containment and closure" or "lockdown" policies); (ii) Government Response Index (GRI) for all categories; (iii) Containment, closure, and Health Index (CHI); and (iv) Economic Support Index (ESI). These indices have been used widely to measure government responses to real or perceived risk of COVID-19 intensity in a country. For example, cross country studies show that cancellation of public events, restriction on private gathering, and closing of schools and workplaces had significant impact on reducing COVID-19 infections (Askitas et al., 2021, p. 1).

The second set of measures on proxies of COVID-19 intensity is the change in mobility-based data from Google Maps (Google, 2020). The Google Community Mobility Reports reveal what changed in response to policies aimed at combating COVID-19 and measure movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential areas.

In selecting our sample period, we take into consideration both the incubation period of COVID-19, as well as information on international travel bans, since our index is calculated based on the assumption that international borders had to be open to enable migrants to travel back to their home countries (Equation 2). We use the insight that many international migrants were faced by uncertainties at the onset of the pandemic and decided to return to their countries of origin, which also led them to inadvertently become carriers of COVID-19 in some cases. Furthermore, we use the knowledge that 99% of the infected population showed symptoms within 14 days of exposure for the early strains of COVID-19 (Lauer et al., 2020, p. 580). Since the World Health Organization (WHO) declared COVID-19 a pandemic on March 11, 2020, we add the 14-day incubation period leading to March 24, 2020, to mark a reasonable end date to our sample. To illustrate the timelines, we graph the measure of international border closure proxied by the Government Response Index (Hale et al., 2020) alongside cumulative COVID-19 cases in Figure 2. We deduce the final baseline sample period to be of 10-week duration, starting from January 15, 2020, when ECDC started recording the COVID-19 cases, and ending on March 24, 2020.

Econometric specification

Our benchmark econometric specification takes the following form:

$$\ln(COV_{it}) = \alpha + \beta EXP_{it} + \theta_i + \theta_t + \epsilon_{it}, \quad (5)$$

where COV_{it} is the number of confirmed COVID cases per million people in country i during week t .³ EXP_{it} is country i 's exposure to COVID-19 via return migration and is defined in Equation 4. The coefficient of interest is β , which we expect to be positive if return migration was a vector of transmission to LMICs during the initial phase of the pandemic. ϵ_{it} is the error term. All standard errors are clustered at the country level.

One possible concern with this empirical approach is that the number of COVID-19 cases in i is likely to be driven by country-level characteristics such as its health infrastructure, demographic profile, population density, and rate of urbanization. If these characteristics are also correlated with the stock of out-migrants from i , then β will not be identified. We address this concern in two ways. First, we include origin-country fixed effects, θ_i , in all regressions, which will absorb the time invariant confounding effects. We also include θ_t , which are week fixed effects that absorbs the week specific shocks. All observations are at the country-week level.

A second possible concern is the use of COVID-19 cases for validation of the index. Since our index is constructed using COVID-19 case intensities at the respective migrant-destinations of the LMICs, there could be a potential problem if contemporaneous COVID-19 cases at the migrant origin, COV_{it} , is impacted by any reverse immigrant stock from their destination countries, d . In order to mitigate this concern, we measure the degree of correlation between the destination countries of LMIC migrants, and the destination countries of the emigrants from these respective LMIC destination countries. We find this correlation to be negative and significant indicating that there is a negative likelihood that immigrants into LMICs originate from the same destination countries to which the LMIC population out-migrates (see Figure A2).

A third concern arises from measurement error in the COVID-19 case load in the LMICs. To mitigate this, we carry out robustness checks with several alternate dependent variables to capture various aspects of COVID-19 intensity. These measures include policy restrictions to measure the government's actual or perceived responses to COVID-19 risk (Hale et al., 2020); changes in mobility during COVID-19 (Google, 2020); and COVID-19 deaths per capita.

Finally, we test the assumption that existing stocks of migrants is associated with incoming returnees in the early stages of COVID-19 (see Equation 2). We use a case study from Bangladesh, where we obtained data on airport arrivals in Bangladesh between December 2019 and March 2020 from the Civil Aviation Authority of Bangladesh (CAAB). We estimate the correlation between the CAAB data with data on migrant stock figures from surveys and administrative data in Bangladesh. These are described in further detail in a subsequent case study on Bangladesh.

RESULTS AND DISCUSSION

We present descriptive statistics for the main variables in Table 1. Panel A provides the summary statistics for the logarithm of our main independent variable, which is the logarithm of COVID-19 exposure index as well as COVID-19 cases and deaths from ECDC data.

TABLE 1 Summary statistics

Vars	(1) Obs	(2) Mean	(3) SD	(4) Min	(5) Max
Panel A					
Lg (exp)	1230	0.93	1.53	0.00	7.44
Lg (cases pc)	1230	0.00	0.01	0.00	0.10
Lg (deaths pc)	1230	0.00	0.00	0.00	0.01
Panel B					
Stringency Index	1230	15.83	23.18	0.00	100.00
GRI	1230	12.86	17.77	0.00	84.52
CHI	1230	14.75	20.18	0.00	88.89
ESI	1230	1.57	8.05	0.00	75.00
Panel C					
Retail/ recreation	1230	-0.25	14.52	-80.00	28.00
Grocery	1230	5.18	9.73	-60.00	34.00
Parks	1230	3.81	16.29	-73.00	86.00
Transit	1230	-0.77	14.86	-84.00	27.00
Workplace	1230	4.81	12.05	-59.00	26.00
Panel D					
Share health exp/GDP (%)	123	5.95	2.51	2.27	16.62
Share working age pop (%)	123	63.97	17.62	31.88	111.34
Share pop>65 (%)	123	5.87	3.84	1.92	20.76
Pop density (per/sqkm)	123	134.37	215.21	2.00	1654.67
Share urban pop (%)	123	51.40	20.45	12.71	91.75

Note: Panel A provides the summary statistics for the main variables, including our measure of the COVID-19 exposure index, Covid-19 cases, and deaths. Observations are at the week-country level. Panel B provides the summary statistics for the four main indices: Stringency Index, Government Response Index (GRI), Containment, closure and Health Index (CHI), and Economic Support Index (ESI) (Hale et al., 2020) measured at the country-week level. Panel C provides the summary statistics for the Google Community Mobility variables (Google, 2020), expressed as percentage changes in mobility compared to a baseline day, which represents a normal value for that day of the week and is the median value from the 5-week period January 3–February 6, 2020. Panel D provides the country specific characteristics used as country-level controls.

Given the lower reported numbers in COVID-19 deaths in the sample period, there is very limited variation in the data. Observations are at the week-country level. Panel B provides the summary statistics for the four main stringency indices from Hale et al. (2020) measured at the country-week level. Panel C provides the summary statistics for the Google Community variables. These variables are expressed as percentage changes in mobility compared to a baseline day, which represents a normal value for that day of the week and is the median value from the 5-week period January 3, 2020, to February 6, 2020. All variables in Panels A, B, and C are measured at the country-week level. Panel D provides the country specific characteristics used as country-level controls (World Bank, 2020).

TABLE 2 Contemporaneous and lag effects of COVID-19 exposure index on COVID-19 cases.

Variables	(1) Contemporaneous	(2) 1-wk lag	(3) 2-wk lag	(4) 3-wk lag	(7) All lags
L0. Exp	0.001** (0.000)				0.001 (0.001)
L1. Exp		0.001*** (0.001)			0.000 (0.001)
L2. Exp			0.002*** (0.001)		0.002** (0.001)
L3. Exp				0.002* (0.001)	-0.001 (0.001)
Cons	0.000 -0.001	0.000 -0.001	0.000 0.000	0.000 0.000	0.000 0.000
Obs	1230	1107	984	861	861
R ²	0.16	0.16	0.16	0.15	0.17
# countries	123	123	123	123	123

Note: Dependent variable is logarithm of confirmed COVID-19 cases/person. Exposure is defined in Equation 4 (in logarithm). We control for country and week fixed effects in all specifications. Sample period is 10 weeks from January 15, 2020 to March 24, 2020. Reported standard errors are clustered at the country. Sample includes LMICs as classified by the World Bank (2020). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The main results (based on Equation 5), are reported in Table 2. Results show that our index is significantly and positively correlated with the subsequent COVID-19 cases in the sample of LMICs. We regress confirmed cases on the exposure measures lag on contemporaneous, one, two and three-week lags of the cases (columns 1 to 4, respectively) to observe the effectiveness of the index over different time periods. Results show that the effect on the coefficient is positive and significant at the 5% level for all three lagged cases although the level of significance decreases with the third lag. Only the second lag remains positive and significant at 5% level when all three are included (column 5), which corresponds to the 2-week virus incubation period. A 10% increase in exposure measure results in about 0.02 percentage increase in contemporaneous COVID-19 cases and significant at the 5% level.

While size of the effect is small, the coefficient is positive and significant across all specifications showing a positive relationship between our index and the degree of observed COVID-19 intensity. Thus, Table 2 shows that firstly, our index, which used the stock of international migrants, is a good predictor of the spread of COVID-19. Secondly, the index is most effective at predicting spread in the first 2 weeks of the migrants' return to the country of origin, confirming the epidemiological evolution of the virus.

Robustness checks

While our results in Table 2 show strong positive association between our COVID-19 index measure and the prevalence of COVID-19 in LMICs using reported COVID-19 case data, we are aware of the limitations and heterogeneity in the COVID-19 case data for many LMICs. The

effect size is also quite small and therefore required further robustness checks. Given the results of column 5 of Table 2, we use second lag as the main specification in the robustness analysis.

One key issue is measurement error on the dependent variable, which does not impact the consistency in estimating our coefficient as long as the errors are uncorrelated with our COVID-19 index. Since our index is calculated using out-migration rates and COVID-19 caseloads in migrant-destinations, we expect this to be a valid assumption and expect this risk to be mitigated.

However, if measurement errors are correlated with our COVID-19 index, there might be a positive bias on our estimate. To mitigate these concerns, we study the association between our index and several additional outcomes to test if the sign and direction of our findings remain consistent.

We validate our measure against two sets of variables that measure real and perceived responses to COVID-19 intensity at the country level. The first set of indices measures variation in policy measure undertaken by different governments to contain the spread of disease (Hale et al., 2020). These results are presented in columns 1–4 of Table 3. All four measures show a positive and significant relationship with our COVID-19 exposure index. This indicates that as the risk predicted by our COVID-19 index for an LMIC increases for a country, we also see more restricted government measures in those areas due to real or perceived threat of disease, thus corroborating the validity of an index for pandemic risk assessment.

In columns 5–9 of Table 3, we use five variables that capture that the extent of community mobility of the citizens from Google's Community Mobility Report (Google, 2020). The Google mobility data from Google measure visitor numbers to specific categories of location (e.g., grocery stores; parks; train stations) every day and compares this change relative to baseline days before the pandemic outbreak. Baseline days represent a normal value for that day of the week and are calculated as the median value over the five-week period from January 3, 2020, to February 6, 2020. Measuring it relative to a normal value for that day of the week is helpful because people often have different routines on weekends versus weekdays. We find a significant, negative correlation between these mobility measures and our index indicating that a decline in social and economic activities is associated with an increase in COVID-19 risk. This is consistent with the prediction from our risk index since we would expect that as perceived and actual risk of COVID-19 increases, the observed mobility in a country will decline.

We also test our index against confirmed COVID-19 deaths and present these results in column 10 of Table 3. We find a very small and positive but significant relationship between the log of death per capita with our exposure measure. Unfortunately, given the very small variation in the COVID-19 cases, the size of the coefficient produced very small effects and therefore not large enough to produce sufficient evidence on this measure.

Variation in country level policies over time to combat COVID-19 can impact confirmed cases in the migrant-origin countries. The most rigorous control for this would require week interacted with country fixed effects. However, including the country-week fixed effects would eliminate all variation in our analysis including the one captured by our index, which we ultimately want to measure. We take a less conservative approach and control for country specific confounders by including the interaction terms between month fixed effects and the following in some additional robustness checks: (i) percentage of health expenditure in GDP; (ii) share of working age population; (iii) share of population over 65; (iv) population density; and, (v) share of urban population (columns 1–5, Table A1). This is the closest to mitigating variation between countries over time and we find that all our results hold for the inclusion of all five interactions.

TABLE 3 Effect of exposure measure on alternative outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Stringency	GRI	CHI	ESI	Retail/ Rec	Grocery	Parks	Transit	Workplace	Deaths/ cap
L2.Exp	8.937*** (1.687)	6.344*** (1.237)	6.762*** (1.353)	3.832*** (1.435)	-7.099*** (1.988)	-4.385*** (1.283)	-6.839*** (2.163)	-6.701*** (1.744)	-6.197*** (1.464)	0.000*
Constant	5.162*** (0.611)	4.320*** (0.457)	5.041*** (0.528)	-0.008 (0.393)	4.807*** (0.729)	3.910*** (0.494)	6.701*** (1.013)	3.872*** (0.785)	7.542*** (0.517)	0.000 (0.000)
Obs	888	888	888	888	468	468	467	462	468	984
R ²	0.79	0.80	0.80	0.22	0.61	0.26	0.39	0.59	0.59	0.03
Countries	111	111	111	111	78	78	78	77	78	123

Note: Columns indicate dependent variables for each regression. Columns 1 to 4 represent indices of governments' response to the COVID-19 pandemic (Hale et al., 2020). Columns 5 to 9 report five measures of community mobility (Google, 2020). The final column uses confirmed death/capita. Default lag period is two. Controls include country and week fixed effects in all specifications. Sample period is 10 weeks (January15, 2020 to March 24, 2020). Reported standard errors are clustered at the country level. Sample includes LMICs (World Bank, 2020). ****d* < 0.01, ***d* < 0.05, **d* < 0.1.

We also check the robustness of our baseline results using three different sample periods: (i) 12 weeks from 15 January 2020 to 7 April 2020; (ii) 14 weeks from 15 January 2020 to 21 April 2020; and (iii) 15 weeks from 15 January 2020 to 28 April 2020 (columns 6–8, Table A1). Results remain positive and significant to these specifications. Given the large number of zeros in COVID-19 cases in our data, we express our variables in inverse hyperbolic sine function and find consistent estimates (column 9, Table A1). Finally, we sense-check our hypothesis by restricting the sample to only high-income countries (HIC) (column 10, Table A1). The results are not significant showing no notable association between our COVID-19 risk measure and COVID-19 cases. This supports our initial hypotheses that this approach to understanding COVID-19 risk exposure using international migration links is applicable for LMICs that are prone to be net migrant-sending countries.

Limitations

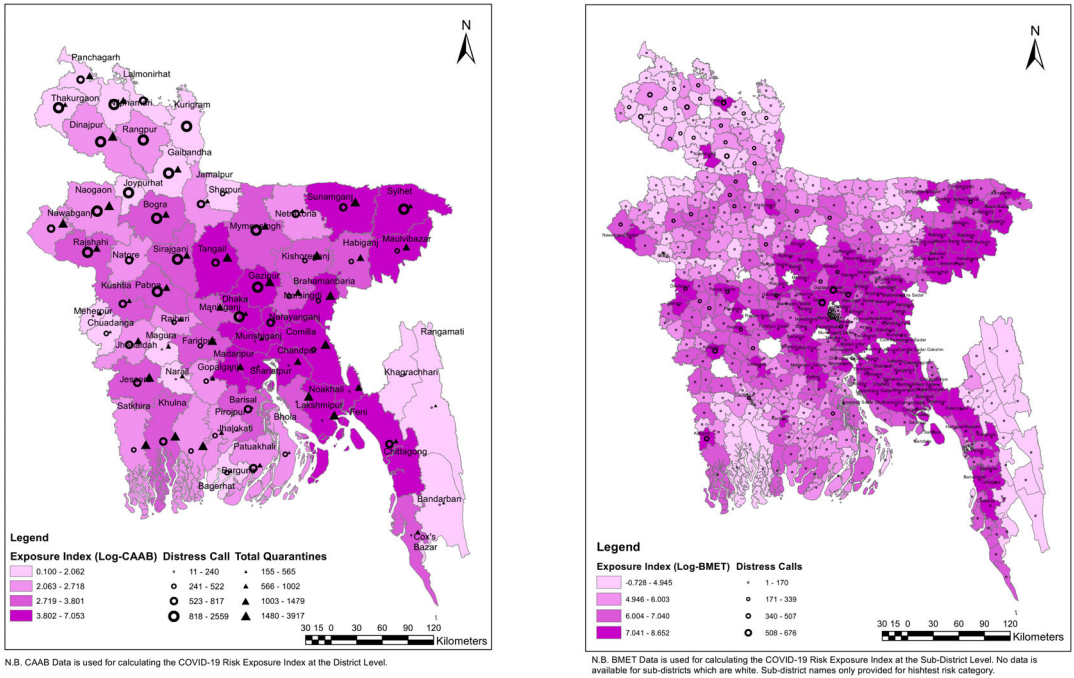
One of the main limitations in this analysis is the errors in calculating COVID-19 case and deaths data as well as the policy restrictions undertaken by governments, which can vary by country over time. We include as many alternate measures of COVID-19 intensity as possible including interactions between month and country characteristics to check if our index consistently reveals the same pattern of association. We find consistent support of our hypothesis using alternate measures of perceived and actual COVID-19 intensity, thus concluding that our index appropriately reflects the COVID-19 risk for LMICs.

In addition to limitations discussed above, we note that COVID-19 spread in the early pandemic days was associated with returning travelers who were also tourists and non-migrants, which is not captured by our index. Additionally, the stock of migrants' data (United Nations, 2017) are heterogeneous and include a combination of permanent, semi-permanent, and temporary migrants, which we do not distinguish in our analysis. The probability of return for these different classes of migrants might be different and future work can extend to incorporate this heterogeneity.

APPLICATION OF PANDEMIC RISK EXPOSURE: CASE OF BANGLADESH

We apply our methodology to estimate pandemic risk via exposure to international migration by constructing corresponding sub-national indices for Bangladesh, which has a significant exposure to international migration. We use this replication to illustrate that administrative and survey data on migrant stock can predict incoming traveler data. We also show the application of our risk exposure to spatially identify socio-economically vulnerable regions. At the regional level, our index is positively and significantly correlated with regions having higher COVID-19 cases, quarantines, and COVID-19 related distress calls to a public hotline (A2I, 2020).

To illustrate the relationship in Equation 2, we first calculate the stock of migrants that traveled to each destination at the district and sub-district level prior to COVID-19 (2018 and 2019) using the database from the Bureau of Manpower, Employment and Training (BMET), under the Bangladesh Government's Ministry of Expatriates' Welfare and Overseas Employment. This database contains information on every migrant registered to go abroad for employment purposes including location at the origin (sub-district level), destination country, and expected departure date. We then estimate the number and destination of migrants at the household level and aggregate to the district level using survey weights from the Household Income and



a District level exposure

b Sub-district (*upazila*) level exposure

FIGURE 3 COVID-19 risk exposure in Bangladesh. (a) District level exposure. (b) Sub-district (*upazila*) level exposure. In panels (a) and (b) we show the regional variation in Bangladesh of the COVID-19 risk exposure index at the district and sub-district levels, respectively. The darkest shade indicates countries in the top-most quartile of the risk categorization according to the index. We also superimpose the cumulative number of COVID-19 related distress calls to a government hotline (as of March 24, 2020) with larger circles illustrating higher calls. For the districts, we also superimpose the cumulative number of COVID-19 related calls (as of March 24, 2020) with larger circles illustrating higher calls. Case data is taken from European Centre for Disease Prevention and Control (ECDC) and migrant stock data from the Bangladesh Bureau of Manpower, Employment and Training (BMET). Distress call and quarantine data was obtained directly from the Government of Bangladesh (A2I, 2020).

Expenditure Survey (HIES, 2016), a comprehensive nationally representative. Finally, using records incoming travelers to Bangladesh between 17 December 2019⁴ and 18 March 2020 from the CAAB, we estimate the number of returnees from each destination at the district level. Given that CAAB data tracks actual returns, it provides the most reliable signal of virus transmission amongst all the sources of migration data that are available to us.

Using the above data, we analyze the association between migrant stock data from administrative and survey data with the CAAB data. We find that airport returnees from CAAB data are positively and significantly correlated with the number of migration permits issued in that district by BMET in the previous 5 years and the number of migrants calculated from the HIES data (Table A2). Our results show that the approach presented in Equation 4, can be conducted credibly using administrative data on migration permits and national surveys thus broadening the scope and applicability of our analytical approach.

We then apply our COVID-19 risk index at the sub-national level to illustrate the spatial risk distribution based on migration-based links (Figures 3a,b). For example, a higher index value

implies strong migration links to destinations such as Italy, Singapore, or the United States, where the disease was already more prevalent early in the pandemic. We find a positive association between our exposure index and the actual number of COVID-19 cases at the regional level. Results at the sub-district level are significant at the 5% level (Table A3). We also found positive correlations between number of people quarantined at the district level and the number of COVID-19 related distress calls placed to a national hotline obtained from the Government of Bangladesh (A2I, 2020) at the sub-district level.⁵

Implementation of localized lockdowns or other targeted policies using district level exposure variation is constrained by the size of the area and the population exposed at this level. The top 10 high risk Bangladeshi districts identified by our district-level index have an average population of 4.6 million while the average population of the top 10 most risky sub-districts is 542,000 (BBS, 2011). Thus, the latter is more feasible for relief or public health targeting and our measures of risk exposure can help identify their vulnerable regions.

Heat maps based on our COVID-19 risk exposure indicate places where development indicators might be negatively impacted by exposure to COVID-19 through international migration channels. For example, Bangladeshi districts that sent many migrants to Italy could be expected to experience larger adverse shocks to remittance income and need greater social safety nets early in the pandemic. Overall, remittances into Bangladesh fell by over 30% (by US\$ 500 million) year-on-year in April 2020 (TBS Report, 2020). Thus, even though migrants may become less predictive of COVID-19 occurrence over time, our sub-national heat maps are still informative about the nature of economic stressors over time. Furthermore, consistent with the logic of our risk exposure index, human mobility was the strongest predictor of at least one symptom associated with COVID-19 amongst a sample of representative households from a survey conducted in Bangladesh's Cox's Bazar district (Lopez-Pena et al., 2020).

CONCLUSION

In our paper, we provide a novel approach in estimating the spatial variation in infectious disease risk for low- and middle-income countries based on their exposure to international migration with a specific application to the COVID-19 pandemic. Our methodology uses readily available datasets on origin–destination linked migration stocks to identify the variation in exposure to COVID-19 across countries and sub-national regions. We validate our index measure using subsequent cases of COVID-19 to demonstrate its reliability. The approach proposed in this study can add to the policy makers toolkit in LMIC contexts for identifying at-risk regions in early stages of viral pandemics, especially when testing data can be inadequate.

LMICs need geographically disaggregated information to determine how to spatially target resources within each country. Widespread, nation-wide lockdowns are either too costly or infeasible in poorer countries (Barnett-Howell & Mobarak, 2020). Thus, alternative means of swift identification of vulnerable regions during pandemics is crucial for policy makers. Data deficiencies can hamper resource allocation both at the sub-national and global levels. International bodies such as WHO need analogous comparative information across countries to spatially target resources and provide support to LMICs at greater risk. The lack of uniformity in testing frequency and different protocols across countries⁶ can make it difficult to identify relative disease risk and target economic support.

Consequently, whether it is targeting public health measures, lockdowns, and quarantines, or providing financial support, international policy makers need to identify at-risk countries quickly. Meanwhile, national- and regional-level decision makers need to prioritize regions that

require rapid response in terms of enhancing hospital and screening capacity, flow of medical resources, or imposing more stringent social distancing and lockdown measures. Vulnerable areas may also need immediate social protection support and targeted relief for those at greatest risk of food insecurity. Due to limitations in the health sector and public resources in LMICs, timely detection of cases and accurate data on the spread of highly infectious diseases such as COVID-19 can be challenging. Despite various data limitations, we found consistent support that our method provides a practical and credible approach for decision makers operating in resource-constrained environments to identify potentially vulnerable regions exposed to pandemic risk as a consequence of international migration-based linkages.

The methods we developed can be applied to create heat maps in other developing countries and future pandemics where decision makers are constrained by inadequate testing capacity. This migration-based exposure index adds to the policy makers toolkit in early stages of a pandemic and can be combined with epidemiological modeling and other data sources such as night-time lights, mobile phone communications, and transport flows to improve predictions on the specific spatial patterns of disease spread within countries. The same exposure concept underlying our index can also be applied to data on internal-migration links to model the community spread of disease over time. Other research papers have also documented how various forms of social and economic connectedness is predictive of the spread of COVID-19 (Chan et al., 2020; Kuchler et al., 2022; Lee et al., 2021).

In a sub-national application of our approach, we worked with multiple sources of data in Bangladesh because our goal is to establish a “proof of concept” that can be applied to other LMICs to make sub-national predictions. The comparison of different measures provides insight on the relative advantages of different data sources that range from administrative record-keeping to national surveys. Many LMIC governments have been collaborating with mobile service providers to collect information on (distress) call patterns made to helplines to implement contact tracing⁷ and these can further build on the approach provided in this paper.

Furthermore, our paper contributes to a growing literature studying the links between infectious diseases like COVID-19 and socio-economic outcomes by focusing on international migration links. With increasing risks of future viral outbreaks and the prominence of international migrants globally, this paper makes an important contribution by quantifying the risk of a country or region to disease outbreak based on its stock of international migrants.

ENDNOTES

- ¹ Note that M_{id} is not normalized by the origin country's population and should not be interpreted as a weight.
- ² Migrants are defined as foreign-born residents or foreign citizens. In developing countries where refugees were not included in population censuses. Data on refugees from the Office of the United Nations High Commissioner for Refugees and the United Nations Relief and Works Agency for Palestine Refugees in the Near East were added to construct the total stock of migrants.
- ³ To account for the large number of zeroes in the COVID infection data, we add one to COV_{iw} prior to taking logs. We also show that our key result is robust to using an inverse hyperbolic sine (IHS) transformation.
- ⁴ Chinese Health officials informed WHO about a cluster of 41 patients with pneumonia in Wuhan, Hubei Province on 31 December 2019. <https://www.who.int/news-room/detail/08-04-2020-who-timeline—covid-19>
- ⁵ Quarantine data come from the Government of Bangladesh data published on the following site: <https://corona.gov.bd/>, accessed on April 16, 2020. The distress call data tracked the location of the origins for the calls placed on the hotline between March 22 and April 12, 2020. Correlations between COVID-19 cases and quarantines and distress calls remain significant after we control for district level measures of medical facilities, preparedness for COVID-19, medical staff availability, and other logistical preparation.

- ⁶ Testing per capita was three times as high in Pakistan compared to Bangladesh, four times as high in Romania compared to Ukraine, seven times as high in El Salvador compared to Guatemala, and 10 times as high in Uruguay compared to Bolivia, [Worldometers.info](https://www.worldometers.info), accessed on April 1, 2020.
- ⁷ China has been using contact tracing applications since February while India launched the Aarogya Setu on 2 April 2020. Meanwhile, Ghana has also developed a COVID-19 tracker app to help trace people infected with the virus amongst other LMICs. A full list of countries using different private and public sector launched apps including their coverage can be found here: <https://www.top10vpn.com/news/surveillance/covid-19-digital-rights-tracker/>

REFERENCES

- A2I. 2020. Government of Bangladesh: Innovate for All (A2I) <https://a2i.gov.bd/>
- Ahmed, Akhter, M. Mehrab Bakhtiar, Daniel Gilligan, John F. Hoddinott, Shalini Roy, Sadat Anowar, Julie Ghostlaw, and Giang Thai. 2021. *Signs of Recovery: Patterns of Livelihoods and Food Security before and during COVID-19 in Rural Bangladesh*. Washington DC: International Food Policy Research Institute (IFPRI) No. April 2021.
- Askitas, Nikolaos, Konstantinos Tatsiramos, and Bertrand Verheyden. 2021. "Estimating Worldwide Effects of Non-pharmaceutical Interventions on COVID-19 Incidence and Population Mobility Patterns Using a Multiple-Event Study." *Scientific Reports* 11(1): 1–13.
- Bangladesh Bureau of Statistics (BBS). 2011. Population and Housing Census 2011 <http://data.bbs.gov.bd/index.php/catalog/66>
- Barnett-Howell, Z., and A. M. Mobarak. 2020. Should Low-Income Countries Impose the Same Guidelines as Europe and North America to Halt the Spread of COVID-19? Yale University. <https://som.yale.edu/should-low-income-countries-impose-the-same-social-distancing-guidelines-as-europe-and-north-america-to-halt-the-spread-of-covid-19>
- Castelli, Francesco, and Giorgia Sulis. 2017. "Migration and Infectious Diseases." *Clinical Microbiology and Infection* 23(5): 283–9.
- Centre for Disease Control (CDC). 2022. *Indicators for Monitoring COVID-19 Community Levels and COVID-19 and Implementing COVID-19 Prevention Strategies* <https://www.cdc.gov/coronavirus/2019-ncov/science/community-levels.html>
- Chan, Ho Fai, Ahmed Skali, and Benno Torgler. 2020. "A Global Dataset of Human Mobility." In *CREMA Working Paper No. 2020–04*. Zürich: Center for Research in Economics, Management and the Arts (CREMA).
- European Centre for Disease Prevention and Control (ECDC). 2022. Latest COVID-19 Data Accessed July 31, 2022. <https://www.ecdc.europa.eu/en/covid-19/data>
- Egger, Dennis, Edward Miguel, Shana S. Warren, Ashish Shenoy, Elliott Collins, Dean Karlan, Doug Parkerson, et al. 2021. "Falling Living Standards during the COVID-19 Crisis: Quantitative Evidence from Nine Developing Countries." *Science advances* 7(6): eabe0997.
- Gatto, Marino, Enrico Bertuzzo, Lorenzo Mari, Stefano Miccoli, Luca Carraro, Renato Casagrandi, and Andrea Rinaldo. 2020. "Spread and Dynamics of the COVID-19 Epidemic in Italy: Effects of Emergency Containment Measures." *Proceedings of the National Academy of Sciences* 117(19): 10484–10,491.
- Google. 2020. Covid-19 Community Mobility Reports April 20, 2021. https://www.google.com/covid19/mobility/data_documentation.html?hl=en.
- Greenaway, Christina, and Brian D. Gushulak. 2017. "Pandemics, Migration and Global Health Security." In *Handbook on Migration and Security* 316–36. Cheltenham: Edward Elgar Publishing.
- Guadagno, Lorenzo. 2020. "Migrants and the COVID-19 Pandemic: An Initial Analysis." *Migration Research Series* 60: 4–12 <https://publications.iom.int/books/mrs-no-60-migrants-and-covid-19-pandemic-initial-analysis>.
- Guan, Jinxing, Yongyue Wei, Yang Zhao, and Feng Chen. 2020. "Modeling the Transmission Dynamics of COVID-19 Epidemic: A Systematic Review." *Journal of Biomedical Research* 34(6): 422–30.
- Gushulak, Brian D., Jason Weekers, and Douglas W. MacPherson. 2009. "Migrants and Emerging Public Health Issues in a Globalized World: Threats, Risks and Challenges, an Evidence-Based Framework." *Emerging health threats journal* 2(1): 7091.
- Hale, Thomas, Samuel Webster, Anna Petherick, Toby Phillips, and B. Kira. 2020. Oxford COVID-19 Government Response Tracker (OxCGRT). Last updated 8: 30
- Hirsch, Jennifer S. 2014. "Labor Migration, Externalities and Ethics: Theorizing the Meso-Level Determinants of HIV Vulnerability." *Social Science & Medicine* 100: 38–45.

- HIES. 2016. Household Income and Expenditure Survey 2016–2017 <https://catalog.ihnsn.org/catalog/7399/study-description>
- Ivers, Louise C., and David A. Walton. 2020. “COVID-19: Global Health Equity in Pandemic Response.” *The American journal of tropical medicine and hygiene* 102(6): 1149–50.
- Kuchler, Theresa, Dominic Russel, and Johannes Stroebel. 2022. “JUE Insight: The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook.” *Journal of Urban Economics* 127: 103314.
- Lauer, Stephen A., Kyra H. Grantz, Qifang Bi, Forrest K. Jones, Qulu Zheng, Hannah R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. 2020. “The Incubation Period of Coronavirus Disease 2019 (COVID-19) from Publicly Reported Confirmed Cases: Estimation and Application.” *Annals of internal medicine* 172(9): 577–82.
- Lee, Jean N., Mahreen Mahmud, Jonathan Morduch, Saravana Ravindran, and Abu S. Shonchoy. 2021. “Migration, Externalities, and the Diffusion of COVID-19 in South Asia.” *Journal of Public Economics* 193: 104312.
- Lopez-Pena, P., Davis, C. A., Mobarak, A. M., & Raihan, S. 2020. *Prevalence of COVID-19 symptoms, risk factors, and health behaviors in host and refugee communities in Cox's Bazar: A representative panel study*. Yale University. <http://yriase.yale.edu/wp-content/uploads/2020/05/Prevalence-of-COVID-19-symptoms-risk-factors-and-health-behaviors-in-host-and-refugee-communities-in-Cox's-Bazar.pdf>
- MacLean, Ruth, and Simon Marks. 2020. “10 African Countries Have No Ventilators. That's Only Part of the Problem.” In *New York Times*. May 17, 2020. Dakar: New York Times <https://www.nytimes.com/2020/04/18/world/africa/africa-coronavirus-ventilators.html>.
- McAuliffe, Marie, and Binod Khadria. 2020. “Report Overview: Providing Perspective on Migration and Mobility in Increasingly Uncertain Times.” *World migration report* 2020(1): e00011.
- Milani, Fabio. 2021. “COVID-19 Outbreak, Social Response, and Early Economic Effects: A Global VAR Analysis of Cross-Country Interdependencies.” *Journal of population economics* 34(1): 223–52.
- TBS Report. 2020. “Remittance Drops 34% in April amid Pandemic.” *The Business Standard*, April 30: 2020 July 31, 2021. <https://tbsnews.net/economy/remittance-drops-34-april-amid-pandemic-75796>.
- Tondl, Gabriele. 2021. *Development in the Global South at Risk: Economic and Social Effects of the COVID-19 Pandemic in Developing Countries*, Vol 65. Vienna: ÖFSE Working Paper.
- United Nation Development Programme (UNDP). 2020. *Dashboard 1: Quality of Human Development* July 31, 2021. https://hdr.undp.org/sites/default/files/2021-22_HDR/HDR2021-22_Dashboard_1.pdf
- United Nations. 2017. United Nations Global Migration Database. United Nations Department of Economic and Social Affairs July 31, 2021. https://population.un.org/unmigration/index_sql.aspx
- Walker, Patrick G. T., Charles Whittaker, Oliver J. Watson, Marc Baguelin, Peter Winskill, Arran Hamlet, Bimandra A. Djafaara, et al. 2020. “The Impact of COVID-19 and Strategies for Mitigation and Suppression in Low-and Middle-Income Countries.” *Science* 369(6502): 413–22.
- World Health Organization (WHO). 2022. WHO Coronavirus Dashboard July 31, 2021. https://covid19.who.int/?adgroupsurvey={adgroupsurvey}&gclid=CjwKCAiA866PBhAYEiwANkIneHNNH16r3fMz5il4A0DCojPOcTIVurn0Hm5H8GVO230GVj5haLyEovRoCjIEQAvD_BwE
- World Bank. 2020. Countries and Economies <https://data.worldbank.org/country>

How to cite this article: Khan, Mahreen, Reshad Ahsan, Kazi Iqbal, Ahmed Mushfiq Mobarak, and Abu S. Shonchoy. 2023. “Using International Migration Links for Early Detection of COVID-19 Risk Exposure in low- and Middle-income Countries.” *Applied Economic Perspectives and Policy* 45(4): 1780–1800. <https://doi.org/10.1002/aep.13387>

APPENDIX A

TABLE A1 Robustness checks for impact of exposure measure.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Control 1	Control 2	Control 3	Control 4	Control 5	Period 1	Period 2	Period 3	IHS	HIC
L2.Exp	0.003*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.004*** (0.002)	0.006** (0.002)	0.006** (0.003)	0.002*** (0.001)	-0.001 (0.018)
Constant	-0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.002)	0.000 (0.000)	0.000 (0.005)
Obs	936	960	960	984	976	1476	1722	1845	1230	760
R ²	0.17	0.17	0.17	0.15	0.17	0.23	0.19	0.18	0.15	0.26
# countries	117	120	120	123	122	123	123	123	123	76
MonthFE x Interactions*	X	X	X	X	X					

Note: Dependent variable is in logarithm (confirmed COVID-19 cases/person). Exposure is calculated as per Equation 4 and expressed also in logarithm. In column 1, we control for interactions between month fixed effects (FE) and share of health exp. in GDP; in column 2, we control month FE interacted with the share of working age population; in column 3, we control month FE interacted with the share of population aged over 65; in column 4, we control month FE interacted with population density; and in column 5, we control interactions between the month FE and share of urban population. In column 6, we extend the sample period by 2 weeks. In this case the sample period is 15 January 2020–7 April 2020. In column 7, we extend the sample period further by 2 weeks. In this case, the sample period is 15 January 2020–21 April 2020. In column 8, we again extend by 1 week to 28 April 2020. The reason for extending sample period by 1 week is that this period includes the date (22 April 2020) when the highest number of countries implemented international travel ban. In column 9, both dependent and independent variables are expressed in inverse hyperbolic sine. In column 10, we use restrict our sample to only include migrants of high-income countries. In columns (1)–(9), the sample includes migrants from LMICs as classified by the World Bank in 2019. In columns 1–5, 9 and 10, the sample period is 10 weeks from January 15, 2020 to March 24, 2020. We control country and month FE in all specifications. Robust standard errors are clustered at the country level and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A2 Bangladesh case study: Regression of migrant stock data from administrative and survey data on airport arrival data.

	(log) Number of arrivals from CAAB data		
	District level		
	(1)	(2)	(3)
(Log) Migrants (HIES, 2016)	0.464*** -6.54		-0.0113 (-0.09)
(Log) Migrants (BMET Avg 2015–2019)		0.731*** -8.81	0.744*** -4.51
Constant	4.698*** -6.39	3.030*** -4.13	3.031*** -4.09
R ²	0.408	0.556	0.556
<i>Number of districts (obs)</i>	64	64	64

Note: The dependent variable is the logarithm number of arrivals using airport arrival data from the Civil Aviation Authority of Bangladesh (CAAB). In column 1, the independent variable is the logarithm of the number of migrants calculated from the nationally representative Household Income and Expenditure Survey (HIES, 2016). In column 2, the independent variable is the logarithm of the number of average annual out-migrants calculated from data from the Bangladesh Bureau of Manpower Employment and Training (BMET) for the years 2015 to 2019. All regressions are at the district level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE A3 Bangladesh case study: Correlation between COVID-19 risk exposure index and COVID-19 cases at the district and sub-district level.

Variables	(1) District	(2) Sub-district
L2. Exp	0.06285 (0.054)	0.137*** (0.031)
Cons	4.396*** -0.081	1.606*** -0.038
Obs	576	4887
R-sq	0.90	0.85062
# regions	64	543

Note: The dependent variable is the log of confirmed COVID-19 cases per person at the district (column 1) and sub-district (column 2) levels. Exposure is the exposure measure as defined in Equation 4 (in logarithm). We control for respective regional and week fixed effects for all specifications. The sample period is 10 weeks from January 15, 2020 to March 24, 2020. Robust standard errors clustered at the regional level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

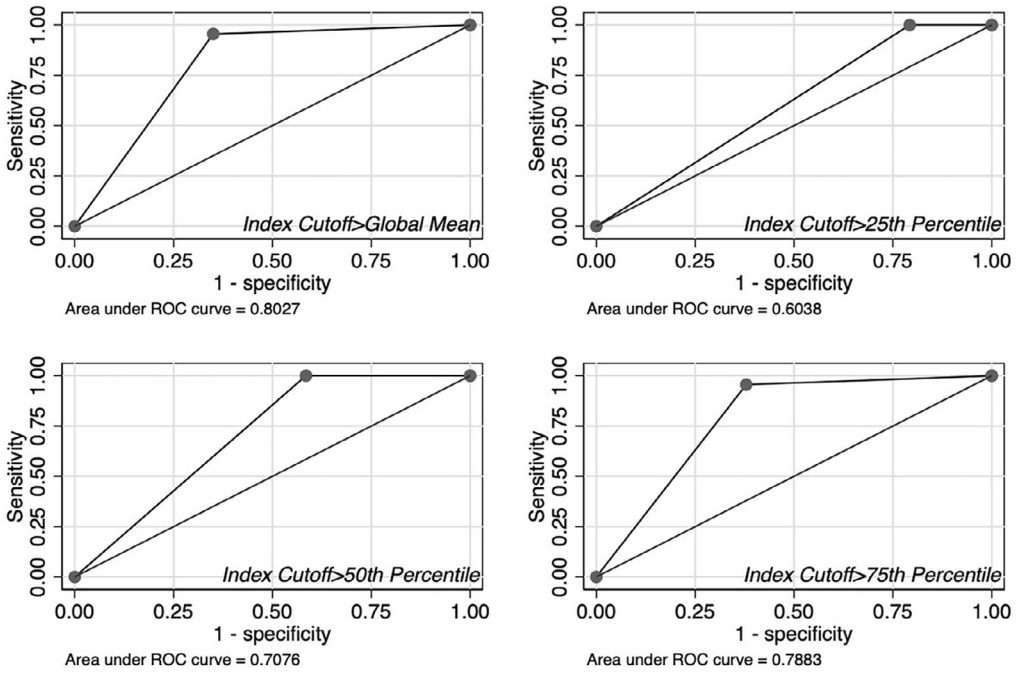


FIGURE A1 Sensitivity analysis using receiver operating characteristics (ROC). We show sensitivity analysis using receiver operating characteristics (ROC) and area under curve (AUC) for four different cut-offs of our COVID-19 index in predicting high risk substantial and high-transmission regions using the United States Center for Disease Control (CDC) Indicator of Community Transmission (CDC, 2022). For index cut-offs that are at the mean, median and 75th percentile, our analysis show that the AUC is above 0.7, indicating strong predictive power of our COVID-19 index for high transmission regions.

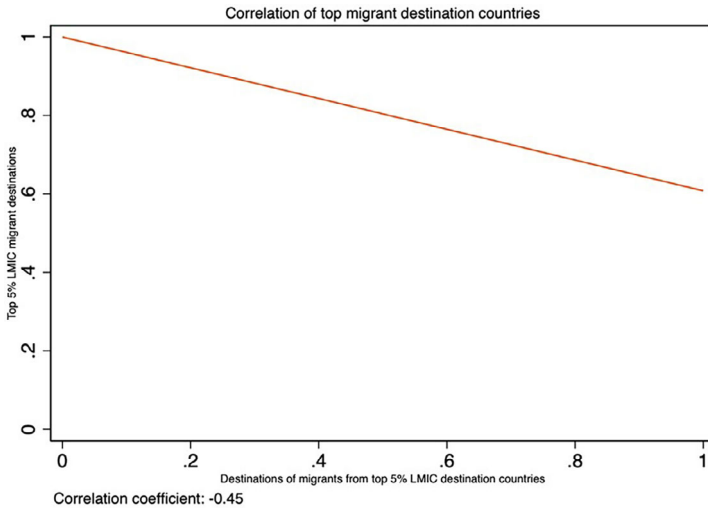


FIGURE A2 Correlation between destination of LMIC out-migrants and immigrants. Graph shows the degree of correlation between the destination countries of LMIC migrants (x-axis), and the destination countries of the emigrants from these respective LMIC destination countries (y-axis). Correlation coefficient of -0.45 is significant at the 5% level.