Photo credit: Mehedi Hasan/ Dhaka Tribune

Photo credit: Mehedi Hasan/Dhaka Tribune

— APRIL 26, 2020 —

Using Migration Patterns to Predict COVID-19 Risk Exposure in Developing Countries

Reshad Ahsan (Univ. of Melbourne), Kazi Iqbal (BIDS), Mahreen Khan (MIT Sloan), Ahmed Mushfiq Mobarak (Yale Univ)[1], and Abu Shonchoy (FIU)

EXECUTIVE SUMMARY

Inadequate COVID-19 testing capabilities is producing testing data that cannot be reliably compared across countries or across jurisdictions within low- and middle-income countries (LMICs). This is hampering the ability of LMICs to devise timely and effective policy responses, such as identifying hotspots and spatially targeting public health responses or economic relief. Data deficiencies also hamper global resource allocation. International bodies such as the World Health Organization need comparative information on disease risk across countries, to be able to direct support to regions at greater risk.

We develop a methodology to make indirect inferences about the spatial distribution of COVID-19 risk using the insight that migration is a primary driver of disease spread across jurisdictional borders. We use bilateral migration flows and prevalence of COVID-19 cases in all migration destinations (such as Italy, Spain, United States, Singapore, and South Korea) to construct a country-level index of COVID-19 risk exposure. We then use sub-district and municipality level data on the origins of migrants, and airport disembarkation data on recent returnees, to create sub-national heat maps of COVID-19 risk within Bangladesh and the Philippines. We validate our indices by comparing them to data on confirmed cases, COVID- 19 deaths, location of quarantines and distress calls to a government hotline. We use multiple data sources in Bangladesh to evaluate the broad applicability of our method to other LMICs where only a subset of such data may be publicly accessible. Our analysis proceeds in the following steps:

1. We combine the United Nations (2017) database of country-pair migration links with Johns Hopkins CSSE data on COVID-19 outbreak intensity at each migration destination. We construct an index of COVID-19 risk exposure for every country, using the number of emigrants from that country to COVID-affected destinations to infer the likelihood that return migrants are now bringing back the disease to each 'home' country.

2. We validate this COVID-19 exposure index by comparing it to the number of confirmed COVID-19 cases through testing, as well as to the number of COVID-deaths (given the aforementioned limitation of testing data). There are strong positive correlations between our index and both confirmed cases and deaths, in the order of +0.66 to +0.72. The strong predictive power of our index survives even after controlling for a broad set of country-characteristics that can proxy for within-country transmission of disease after COVID-19 importation via migrants.

Next, we apply the same insight to create sub-national COVID-risk-exposure indices for Bangladesh utilizing data on district and sub-district origins of emigrants to COVID- affected destinations.

a. The first index utilizes airport disembarkation card data collected by the Civil Aviation Authority of Bangladesh (CAAB) from arriving passengers during December 2019 - March 2020. Returnees' districts of origin strongly predict subsequent quarantines (correlation +0.52) and distress calls to a government hotline from those districts (correlation +0.77).

b. We show that district origins of airport returnees are strongly correlated with the migration permits handed out to people from those

1. Contact: ahmed.mobarak@yale.edu. Dated: May 3, 2020. Updates, including the full paper on which this policy brief is based, can be found at http://yrise.yale.edu/covid-19. The sub-national COVID-risk indices we describe for Bangladesh and Philippines can be constructed for other migrant-sending countries that have data on the district or municipality origins of migrants. We stand ready to help if required.

districts by the government in the previous five years (correlation +0.73). The permit data has fine-grained addresses, which allow us to create a sub-district-level risk exposure index. This is more useful for precise targeting of policy. That index correlates well (+0.47)

with distress calls to government hotlines at the sub-district level.

c. Moreover, district-level returnee numbers are also predicted by migrants from those districts identified in the nationally representative Household Income and Expenditure Survey (HIES). A risk-exposure index constructed based on HIES is also positively correlated with subsequent COVID-19 quarantines (+0.51) reported by the Bangladesh government and distress calls (+0.54).".

4. We apply the same method to the Philippines and construct province and municipality level COVID-19 risk exposure indices. Indices predict the spatial variation in COVID-19 cases confirmed by the Filipino government (correlation +0.71 and +0.64 respectively).

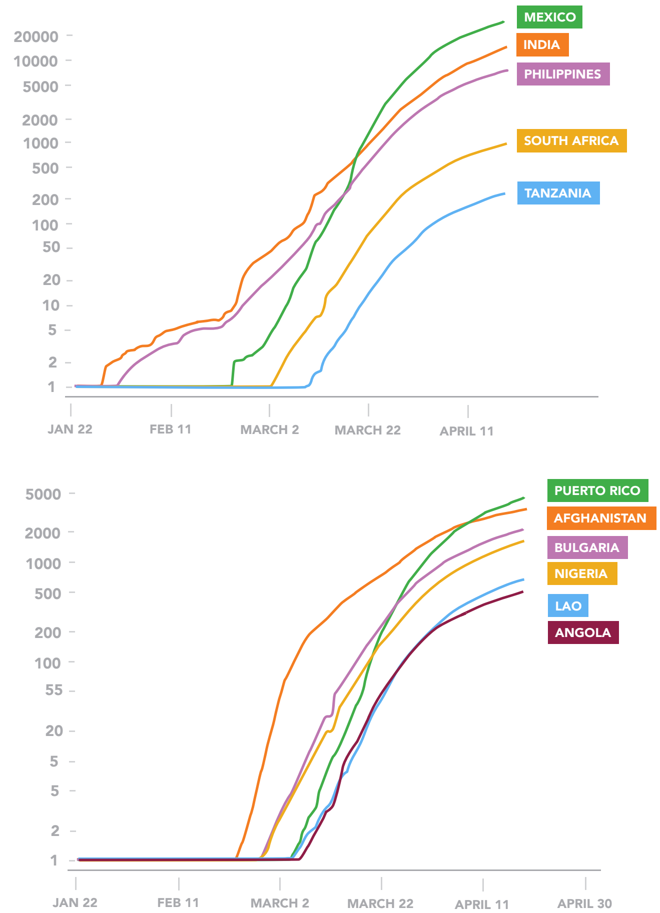
5. To ground-truth the migration-disease link underlying our method, we conduct a phone- survey of 909 households across one Bangladeshi district, and find that respondents in communities where a migrant returned in the 2 weeks prior are 242% more likely to report WHO/CDC COVID-19 symptoms. Returnee presence is *the single largest risk factor* in multivariate analysis. Traveling away from the community is also predictive of symptom prevalence, which suggests that mobility is a key transmission vector.

1 INTERNATIONAL COVID-19 EXPOSURE INDEX

We construct an index of COVID-19 risk exposure for every country using data on country- pair migration links and the prevalence of COVID cases at the migration destinations. The index (formula provided in Section 5) can be interpreted as the expected number of returning migrants to each country who have been infected with COVID-19, given the spread of the virus in the countries they return from.

Figure 1 illustrates the growth in migration-related COVID-19 exposure for a select set of migrant-sending countries. India and the Philippines were exposed early due to their emigrant exposure to the UAE and China, while Mexico was exposed much later given its close migration links to the U.S. The later and less intense exposure of South Africa, Tanzania, Bulgaria, Nigeria, Lao and others are due to their relatively low level of migration dependence, or their exposure to destinations where the outbreak occurred later.

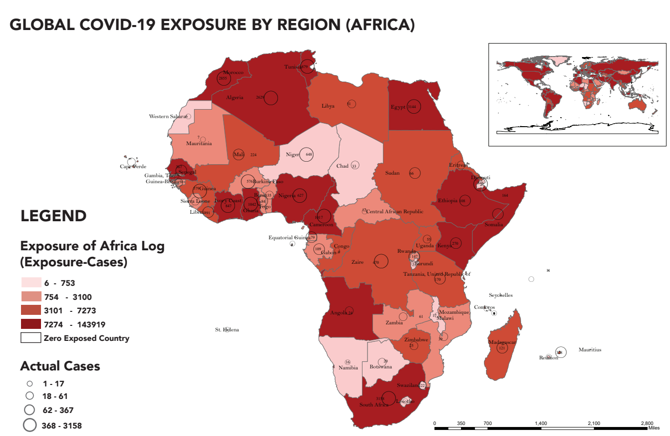
FIGURE 1: Cumulative COVID-19 exposure over time.



NOTES: the vertical axis is in log scale.

In Figure 2 below, we illustrate the cross-country variation in our index in Africa and South and Southeast Asia. Among African countries where COVID-19 is not widespread yet (defined as having fewer than 2,000 cases), our index suggests that Angola, DR Congo, Ethiopia, Kenya, Ghana, Nigeria, and Zimbabwe are relatively more exposed. Table 1 provides a list of highly exposed countries by region, given their migration links to destinations where COVID-19 is present.

FIGURE 2: COVID-19 Exposure by Region



GLOBAL COVID-19 EXPOSURE BY REGION (SOUTH AND SOUTHEAST ASIA)

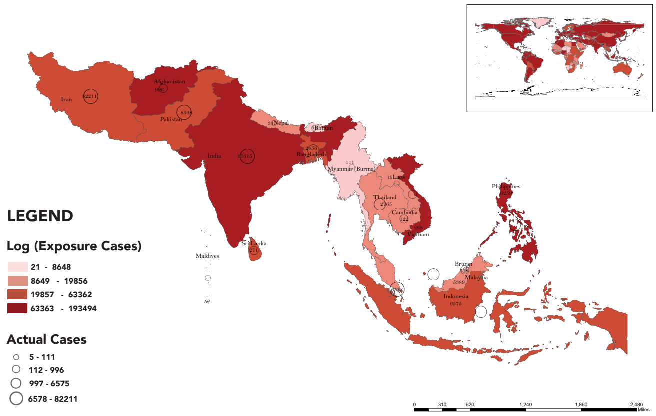


TABLE 1: Highly Exposed Countries by Region

REGION	HIGHLY EXPOSED COUNTRIES
AMERICAS	Bolivia, El Salvador, Guatemala, Guyana, Honduras, Nicaragua, Suriname, Uruguay, Venezuela
CARIBBEAN	Cuba, Haiti, Jamaica, Puerto Rico, Trinidad and Tobago
EUROPE	Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Hungary, Latvia, Lithuania, North Macedonia, Moldova, Slovakia
MIDDLE EAST AND NORTH AFRICA	Armenia, Azerbaijan, Georgia, Iraq, Jordan, Lebanon, Sudan, Syria, Tunisia, Yemen
SOUTH AND CENTRAL ASIA	Afghanistan, Kazakhstan, Nepal, Sri Lanka, Uzbekistan
SOUTH EAST ASIA	Cambodia, Hong Kong, Lao, Myanmar, Viet Nam
SUB-SAHARAN AFRICA	Angola, Cameroon, DR Congo, Ethiopia, Kenya, Somalia, Cote D'Ivoire, Ghana, Nigeria, Senegal, Zimbabwe

NOTES: only countries where COVID is not widespread yet, defined as countries having fewer than 2,000 cases on April 19, 2020 are included in this analysis. Among these countries, those that have exposure values at or above the 67th percentile are classified as being high risk.

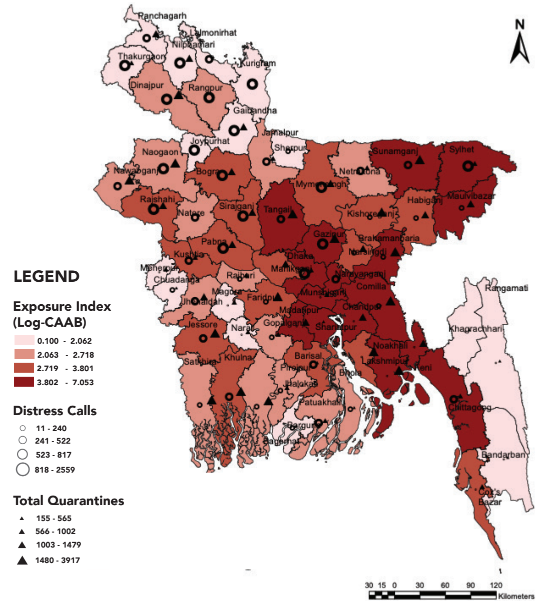
2 SUB-NATIONAL COVID-19 EXPOSURE INDEX

We apply the same method to construct sub-national COVID-19 risk exposure indices for Bangladeshi districts and sub-districts as well as Filipino provinces and municipalities. Figure 3a shows the district-level heat map of COVID-19 risk for Bangladesh, which is created using the local addresses of everyone who arrived in the country between December 17 2019 and March 18 2020. Addresses were extracted from disembarkation cards collected by the Civil Aviation Authority of Bangladesh at airports. Figure 3b displays a *upazila* (sub-district) level index, constructed using addresses extracted from the database of migration permits allocated by the Bangladesh Bureau of Manpower, Employment and Training (BMET). These are measures of each sub-district's exposure to global COVID-19 risk, in that the index value rises if the locality has strong migration links to destinations such as Italy, Singapore or the United States, where the disease was already more prevalent.

FIGURE 3: Heat Maps for COVID-19 Risk Exposure in Bangladesh

(a) DISTRICT LEVEL

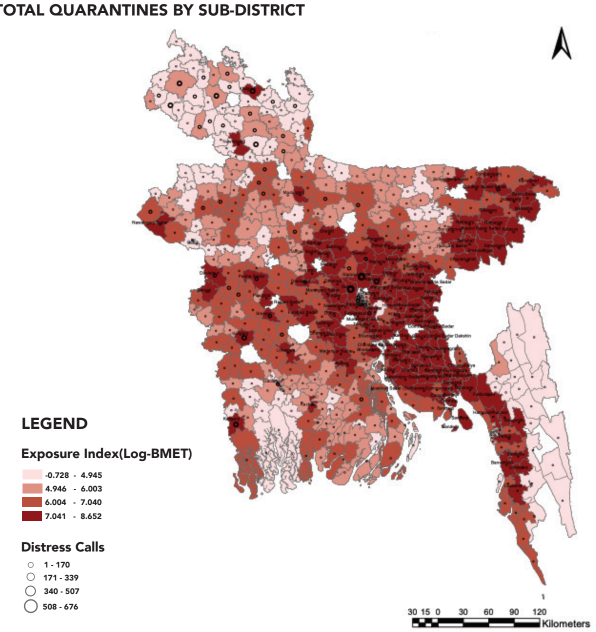
BANGLADESH COVID-19 RISK EXPOSURE (BY CASES), DISTRESS CALLS, AND TOTAL QUARANTINES BY DISTRICT



N.B. CAAB Data is used for calculating the COVID-19 Risk Exposure Index at the District Level.

(b) SUB-DISTRICT LEVEL

BANGLADESH COVID-19 RISK EXPOSURE (BY CASES), DISTRESS CALLS, AND TOTAL QUARANTINES BY SUB-DISTRICT



N.B. Administrative Data is used for calculating the COVID-19 Risk Exposure Index at the Sub-District Level. No data is available for sub-districts which are white. Sub-district names only provided for hishtest risk category.

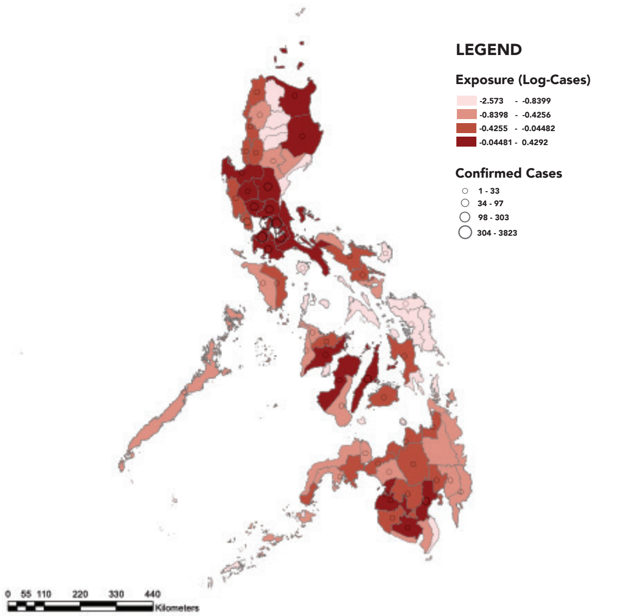
In the full paper, we also construct a risk exposure index for Bangladeshi districts using survey data on migrants reported in the Household Income and Expenditure Survey (HIES 2016). This survey is conducted by the Bangladesh Bureau of Statistics and is similar to a World Bank Living Standards Measurement Study (LSMS) survey. We show that the destinations of migrants identified in the HIES survey and the BMET migration permits predict the locations that travelers return from, as recorded by

CAAB at airports during December 2019 - March 2020. This makes our sub-national COVID-19 risk exposure index broadly applicable: this method can be applied to any country that records migration permits (e.g. Indonesia), or has an LSMS survey with a migration module (e.g. Cambodia, Ethiopia, Nigeria, Nepal, Uganda).

Next, we apply our method to the Philippines and construct province- and municipality- level risk exposure indices (Figure 4). These are constructed using administrative data on international migrants from the Overseas Worker Welfare Administration (OWWA).

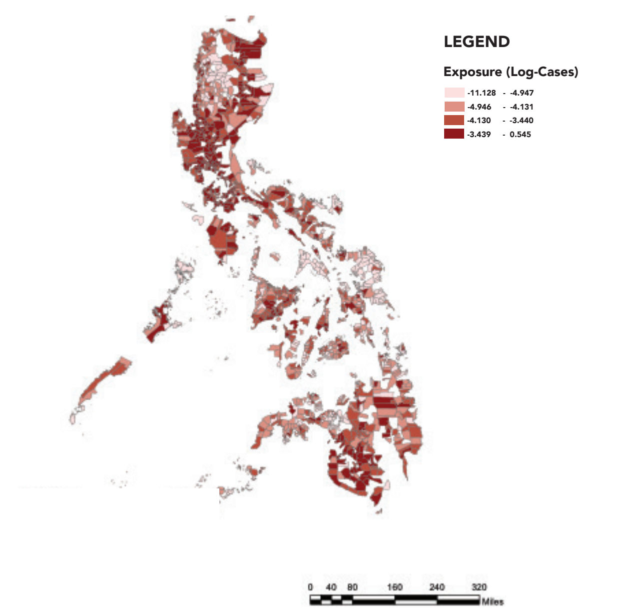
FIGURE 4: Heat Maps for COVID-19 Risk Exposure in the Philippines

THE PHILIPPINES COVID-19 RISK EXPOSURE (BY CASES), BY PROVINCE



N.B. Administrative Data is used for calculating the COVID-19 Risk Exposure Index at the Municipality Level. Names are provided only for the highest risk category.

THE PHILIPPINES COVID-19 RISK EXPOSURE (BY CASES), BY MUNICIPALITY



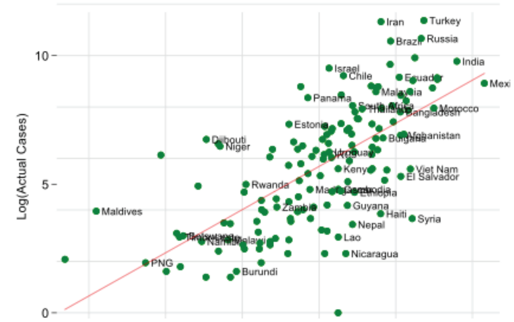
N.B. Administrative Data is used for calculating the COVID-19 Risk Exposure Index at the Municipality Level. Names are provided only for the highest risk category.

3 VALIDATION

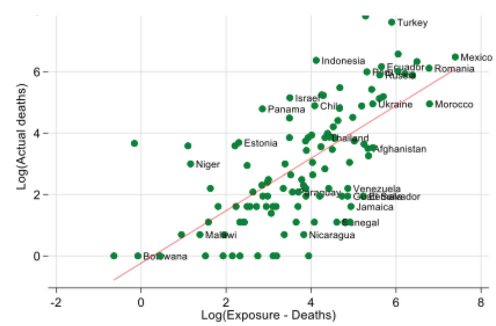
To explore whether our migration-based risk exposure index is informative about the likelihood of COVID-19 presence in the migrant-sending countries, we compare our index to the actual number of confirmed cases and confirmed deaths. Our country-level index is strongly correlated with both confirmed cases (+0.67, *p*-value < 0.001; Figure 5a), and with confirmed deaths (correlation +0.66, *p*-value < 0.001; Figure 5b).

FIGURE 5: Comparing Predicted COVID-19 Exposure with Confirmed Cases and Deaths

(a) CONFIRMED CASES



(b) DEATHS

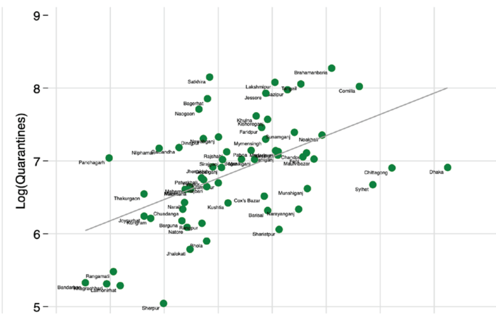


We validate the district-level exposure index for Bangladesh based on CAAB data by first comparing it to the number of people quarantined. The index value for a district is a strong predictor of subsequent quarantines in that district (correlation +0.52, *p*-value < 0.001; Figure 6a). Since quarantine decisions may be partly driven by migrant returnee presence in that district, we also validate using the sub-district origins of distress calls to a government hotline between March 22 and April 12, 2020. The district-level correlation between our CAAB returnee exposure index and distress calls is +0.77 (*p*-value < 0.001; Figure 6b). The indices constructed using BMET data are also strongly correlated with distress calls at the district (correlation +0.71, *p*-value <0 .001) and sub-district levels (+0.47, *p*-value < 0.001, Figure 6d).

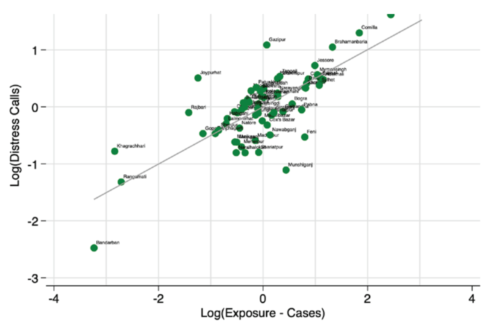
Finally, we validate the province- (correlation +0.71, *p*-level < 0.001; Figure 6d) and municipality-level (correlation +0.64, *p*-level < 0.001) Filipino index using COVID-19 cases reported by the Filipino government.

FIGURE 6: Validating Sub-National Indices in Bangladesh and Philippines Using Quarantines, Distress Calls to Government Hotline and Testing Data

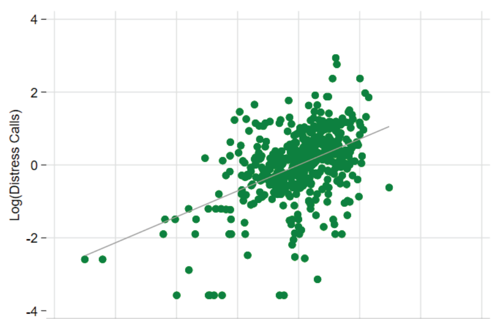
(a) COMPARING CAAB DATA WITH QUARANTIES



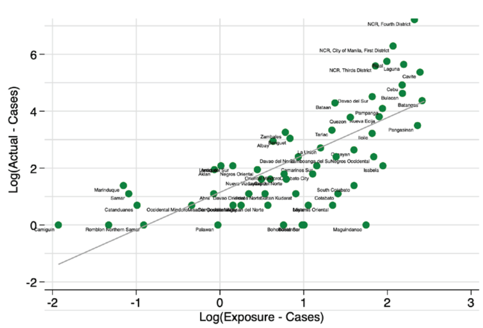
(b) COMPARING CAAB DATA WITH DISTRESS CALLS



(c) COMPARING BMET DATA WITH DISTRESS CALLS (SUB-DISTRICT OR *UPAZILA* LEVEL)



(d) COMPARING FILIPINO INDEX WITH CONFIRMED CASES (PROVINCE LEVEL)



4 RETURNEES PREDICT EMERGENCE OF COVID-19 SYMPTOMS IN SURVEY DATA

We conducted phone surveys of a representative sample of households in one district in Bangladesh to assess the drivers and impacts of COVID-19. The survey contained a module on symptoms designed (with the help of public health experts) to indirectly identify the likelihood of COVID-19. Consistent with the logic of our risk exposure index, human movement is the strongest predictor of COVID-19 symptoms in this survey. Respondents in communities where at least one migrant returned in the 2 weeks prior to the survey are significantly more likely to report COVID-19 symptoms (odds ratio 2.57, CI: 1.34-4.96). Spending at least one day away from home in the same period was also strongly positively correlated with showing symptoms (odds ratio 2.20, CI: 1.28-3.79). This supports the insight underlying our approach: human mobility is critical to the geographic spread of COVID-19.

5 DATA AND METHODS

To construct the indices, we define a country's COVID-19 risk exposure as follows:

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

where *i* indexes migrant-sending developing countries, *d* indexes migrant-receiving destination countries, M_{id} is the stock of migrants from source country *i* in destination *d*, POP_d is the total population in destination country. COV_{dt} is a measure of COVID-19 outbreak intensity in destination country *d* on day *t*. The M_{id} data comes from the United Nations World Population Prospects (United Nations, 2017). COVID-19 outbreak intensity (COV_{dt}) is measured by the number of confirmed cases reported by Johns Hopkins CSSE (Dong, Du, and Gardner, *Lancet* 2020), or in some specifications, the number of COVID-19 related deaths. These data were downloaded from https://github.com/CSSEGISandData/COVID-19 on April 19, 2020.

To conduct sub-national analysis for Bangladesh, we used disembarkation card data collected from incoming travelers to Bangladesh (December 2019 and March 2020) by the Civil Aviation Authority of Bangladesh (CAAB). The distress call data, aggregated at the sub- district level, was shared by the Government of Bangladesh, who set-up multiple COVID-19 national hotline numbers that citizens can call into. The phone survey data came from a representative sample of households in one district in Bangladesh to assess the drivers and impacts of COVID-19.[2] We received government administrative records on migrant permits from the Bureau of Manpower, Employment and Training (BMET) in Bangladesh. Household Income and Expenditure Survey (HIES) is a national representative household survey conducted by the Bangladesh Bureau of Statistics. Municipality level risk exposure index for Philippines is constructed using administrative data on international

2. The sampling frame for this phone survey was the Cox's Bazar Panel Survey (CBPS), a longitudinal study tracking a representative sample of 5,020 refugee and host community households in Cox's bazar district of Bangladesh, living near and far from Rohingya refugee camps. We successfully contacted 909 of 1,255 households in April 2020, and 99% of contacted households consented to participate.

migrants from the Overseas Worker Welfare Administration (OWWA), which Dean Yang and Caroline Theoharides kindly shared.

6 POLICY IMPLICATIONS AND NEXT STEPS

We conducted this analysis to help policymakers in Bangladesh spatially target their COVID- 19 response policies. Whether it's targeting financial support, public health measures, or lockdowns and quarantines, policymakers need information at the sub-national level to prioritize and allocate limited resources more cost-effectively. The methods can be applied in other countries, and we stand ready to help implement if that would be useful. We expect that our method and validation checks will be helpful for decision makers who are currently operating in environments constrained by inadequate testing capacity. International bodies need to identify countries while national- and regional-level decision makers need to prioritize specific locations that require a rapid response in terms of enhancing hospital and screening capacity, flow of medical resources, or imposing more stringent social distancing and lockdown measures that are spatially targeted. Vulnerable areas may also need immediate social protection support and targeted relief for those at greatest risk of food insecurity. Our international analysis in Section 2 can also be useful for the global policy response. Global coordination by international bodies such as the World Health Organization remains critical, as the recent re-emergence of disease in China and Singapore after initial containment makes clear that it is difficult for countries to succeed in isolation without paying attention to disease progression in other regions.