In the absence of reliable, comparable testing data in developing countries, this paper develops a methodology to make indirect inferences about the spatial distribution of COVID-19 risk. Migration has played an outsized role in the spread of coronavirus across national borders, and across jurisdictions within countries. Consequently, pre-existing bilateral migration links with COVID-affected areas can be informative about disease risk in new locations. We construct an index of COVID-19 risk exposure for every country using data on bi-lateral migration and current COVID-19 cases. We validate our exposure index by comparing it to the number of confirmed COVID cases and deaths. Next we apply the same insight to create sub-national COVID-risk-exposure indices for Bangladesh and the Philippines. Using airport disembarkation card data collected from migrants who returned to Bangladesh, we show that the returnees’ districts of origin strongly predict subsequent quarantines and distress calls from the respective districts. Data on migration permits, and on migrants identified in nationally representative household surveys are also informative about recent airport returnees, and subsequent quarantines and distress calls. The COVID signature evident in all these datasets broadens the scope and applicability of our method, because at least one such dataset is available for large numbers of developing countries. We construct heat maps of COVID-19 risk at the district and sub-district levels to inform resource allocation. Our migration-based exposure index can also be combined with epidemiological modeling to improve predictions on the specific spatial patterns of disease spread within countries, and applied to internal-migration links to model the community spread of disease over time.

1 INTRODUCTION

The COVID-19 pandemic hit Western Europe and North America early, but it has begun to spread its tentacles into the global south at the time of our writing. The Bangladesh Government officially declared the entire country to be at risk of the pandemic on April 16, 2020. Most Sub Saharan countries remain relatively unscathed, although the number of cases increased by 40% between April 16 and 26.1 The virus is expected to pose a much larger danger to human health when it does reach low and middle income countries (LMICs). There are fewer than 2,000 working ventilators to serve the hundreds of millions of people across 41 African countries2, and only 27% of the population of developing societies had access to basic handwashing facilities in 2015.3

COVID-19 testing capacity is also very limited in LMICs, which renders sophisticated testing-based defensive strategies - like those implemented in South Korea or Taiwan - difficult to deploy. Widespread testing has allowed developed countries identify hotspots, and target resources accordingly. It was helpful to know, for example, that Louisiana or Lombardy needed more resources and attention than Kentucky or Venetia.4 LMICs need similar geographically disaggregated information to determine how to spatially target resources within each country. Given that widespread, nation-wide lockdowns are either too costly or infeasible in poorer countries (Barnett-Howell and Mobarak, 2020), accurate, sub-national targeting takes on even greater urgency.

ABSTRACT

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

Reshad Ahsan, University of Melbourne
Kazi Iqbal, Bangladesh Institute of Development Studies (BIDS)
Mahreen Khan, MIT Sloan School of Management
Ahmed Mushfiq Mobarak, Yale University, NBER, CEPR, IGC and Deakin University
Abu Shonchoy, Florida International University

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

2. https://www.nytimes.com/2020/04/18/world/africa/africa-coronavirus-ventilators.html. Even the 170,000 ventilators in the United States has been characterized as an acute shortage (https://tinyurl.com/ram86g5)
4. https://coronavirus.jhu.edu/us-map
Data deficiencies hamper resource allocation not only at the sub-national level, but also globally. International bodies like the World Health Organization need analogous comparative information across countries, to be able to spatially target resources and support to LMICs at greater risk. Again, the lack of uniformity is testing frequency and protocols across countries makes it difficult to identify relative disease risk and target support. There is large variation in testing even within sub-continents. As of April 26, 2020, 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 ten times as high in Uruguay compared to Bolivia (Worldometers.info).

In the absence of reliable, comparable testing data, this paper develops a methodology to make indirect inferences about the spatial distribution of disease risk, in order to inform both the global and subnational resource allocation questions. Our method relies on a simple insight on how the COVID-19 virus travels and reaches new locations: human-to-human transmission. Migration has played an outsized role in the spread of coronavirus across national borders, and across jurisdictions within countries. Pre-existing bilateral migration links with COVID-affected areas may therefore be informative about disease risk in new locations. Many countries in the global south have large emigrant populations residing in the high-income countries that were affected early by COVID-19. Those countries, and the regions within those countries that supply international migrants are more exposed to the risk of early disease spread, as large numbers of migrants are forced to return home in the wake of the financial crises initiated by country-wide lockdowns in those popular migration destinations. Consequently, returning migrants become important vectors driving the spread of the disease in these regions. For example, India with over 138,000 migrants in Italy in 2017 was far more exposed to the early spread of COVID-19 than Tanzania, which had a little more than 1,600 migrants in Italy in the same year.

Our analysis proceeds in the following steps, based on the insight that tracking both migration links and recent mobility have predictive value to inform COVID policy responses:

1. We combine the United Nations (2017) database of country-pair migration links with Johns Hopkins CSSE data on COVID outbreak intensity at each migration destination to construct an index of COVID-19 risk exposure for every country. The index value is determined not only by overall emigration rates, but each country’s migration links to specific destinations that were more affected by COVID, like Italy, United States or Spain.

2. We validate this COVID-19 exposure index by comparing it to the number of confirmed COVID cases through testing (especially in countries where reliable testing data are available), and also to the number of COVID-deaths, given the aforementioned limitations in 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 is retained even after controlling for a large set of country-characteristics that can proxy for within-country transmission of disease after COVID-importation via migrants, including population density, comorbidities, healthcare capacity, and policy responses like distancing and lockdowns.

3. Next we apply the same insight to create sub-national COVID-risk-exposure indices for Bangladesh and the Philippines, using data on district, sub-district, and municipalities of origin of emigrants to COVID-affected destinations. We start with airport disembarkation card data collected from migrants who returned to Bangladesh between January and March 2020 by the Civil Aviation Authority (CAAB), and show that the returnees’ districts of origin strongly predict subsequent quarantines (correlation +0.52, p-value<0.01) and distress calls to a government hotline from those districts (correlation +0.77).

4. Next we show that district origins of airport returnees are strongly correlated with the migration permits handed out to people from those districts by the government in the previous 5 years (correlation +0.73). This broadens the scope and applicability of our analytical approach, because airport returnee data may not be quickly accessible in digitized form in every LMIC, but migration permits data exist for many other countries. The permits data also typically carry more details, such as fine-grained addresses and family contact information. Those may be useful for contact tracing. Furthermore, it allows us to create a sub-district (upazila) level risk exposure index for Bangladesh, which is more useful for precise targeting of policy. That index also correlates well with distress calls at the sub-district level (+0.47).

5. 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 on the basis of HIES is also positively correlated with subsequent COVID-19 quarantines (+0.51, p-value<0.01). This broadens the scope of our approach even further, because HIES-style data is publicly accessible for many countries, including every country in which the World Bank has conducted Living Standards Measurement Study (LSMS) Surveys.6

6. Finally, we show another application in the Philippines, constructing the COVID risk exposure index at both the province and municipality levels. The index predicts the province-level COVID cases confirmed by the Filipino government (correlation +0.69).

We work with multiple sources of data in Bangladesh because our goal is to establish a “proof of concept” of an approach that can be applied to many other LMICs to make sub-national predictions, conditional on data availability and limitations, which are bound to vary across countries. The comparison of different measures also provides some insight on the relative advantages of - and proper use of - different data sources.

---

For example, many LMIC governments are collaborating with mobile service providers to collect information on call patterns (such as distress calls) to do contact tracing. The starting point of such datasets are decisions by individuals to make a call, which is very different from random sampling. Our analysis shows that these data are less correlated with data on migrants, airport returnees and quarantines, but have improved predictive power when we analyze within-region correlations.

Even if our migration analysis becomes less predictive of COVID occurrence over time as the disease-spread within LMICs switches from “externally imported through migrants” to “internal community-spread”, the sub-national heat maps that we create is still informative about the nature of economic stressors. Our heats maps also indicate places where remittances supply a larger share of income. Bangladeshi districts that sent many migrants to Italy, for example, are expected to experience larger adverse shocks to remittance income and may need greater public support at this time. Overall, remittances fell 34% (by US$ 500 million) year-on-year in April 2020.8

Our migration-based exposure index can also be combined with epidemiological modeling 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 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-19.9

2 COUNTRY-LEVEL COVID-19 EXPOSURE INDEX

We define a country’s COVID-19 risk exposure as follows:

\[
\text{EXP}_{it} = \sum_{d=1}^{D} M_{id} \left( \frac{\text{COV}_{dt}}{\text{POP}_{d}} \right)
\]

where \(i\) indexes migrant-sending developing countries and \(d=1,2,\ldots,D\) indexes migrant-receiving destination countries. \(M_{id}\) is the stock of migrants from source country \(i\) in destination \(d\) in mid-2017 and is taken from the United Nations World Population Prospects: The 2017 Revision (United Nations, 2017). \(\text{POP}_{d}\) is the total population in \(d\) in mid-2017. \(\text{COV}_{dt}\) is a measure of COVID-19 outbreak intensity in destination \(d\) on day \(t\). We proxy outbreak intensity with the number of confirmed cases reported by Johns Hopkins CSSE (Dong, Du, and Gardner, 2020), and in some specifications, with the number of deaths as a result of COVID-19.10 Our exposure measure is the product of country \(i\)’s stock of out-migrants and the outbreak intensity in the respective destination. Thus, it can be interpreted as the expected number of returning migrants to \(i\) who have been infected by COVID-19.11,12

In Figure 1 below, we illustrate migration-related COVID-19 exposure for a select set of countries. The first graph includes two countries, India and the Philippines, that were exposed early and intensively. Philippines’ early exposure is explained by its migration links to China, which was the destination for approximately 75 thousand Filipino migrants in mid-2017. In contrast, India’s early exposure is explained by the sheer number of its emigrants. For instance, 3.3 million+ Indian migrants in the UAE raises its exposure, even if the UAE only had a handful of confirmed cases in late January.

In contrast to India and the Philippines, Mexico was exposed much later. This is explained by its close migration links to the U.S., where 98% of Mexican emigrants reside. Mexico’s exposure closely mirrors the U.S.’s own outbreak trend, which was characterized by a relatively late start but a rapid growth in confirmed cases thereafter.

Lastly, 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, and their exposure to countries where the outbreak occurred later, e.g. the U.K. and the U.S. Afghanistan was exposed early through its migration links to Iran.

**FIGURE 1:** Cumulative COVID-19 exposure over time.

---

7. China has been using contact tracing applications since February while India launched the Aarogya Setu on April 2, 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/


9. Examples include: Facebook social connectivity data to show that areas with stronger social ties to two early COVID-19 “hotspots” (Westchester County, NY in the U.S. and Lodí province in Italy) generally have more confirmed COVID-19 cases as of March 30, 2020 (Kuchler, Russell and Stroebel, 2020); Google’s newly released mobility report used to track human mobility and COVID-19 spread (Chan, Skal, and Torgler, 2020); migrant shares to predict COVID-19 spread in developing countries (Lee et al, 2020).

10. These data were downloaded from https://github.com/CSSEGISandData/COVID-19 on April 19, 2020.

11. While we do not expect all out-migrants to return to country \(i\), we show below that out-migration stocks are highly correlated with incoming traveler data in Bangladesh during January–March 2020.

12. We have experimented with other versions of our exposure measure where we normalized \(M_{id}\) with country \(i\)’s population, used the number of confirmed cases in the destination country instead of per capita cases etc. These results are presented in the Appendix.
2.1 Validation
To explore whether our migration-based exposure measure 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. Specifically, we compare the exposure index defined for each country on April 19 with the cumulative number of confirmed cases in that country on April 19. Figure 2(a) plots the comparison, and the correlation coefficient between the two measures is +0.67.

The actual number of cases will be affected by other factors such as each country’s social-distancing regime. We control for this in panel (b) of Figure 2 using data for the subset of countries available from Hale et al. (2020). The correlation coefficient increases to +0.72.

Panel (c) adds controls for the log of a country’s GDP per capita, the number of years of compulsory schooling it requires, an indicator of its health access and quality, the trade share of GDP, and the share of tourism earnings to total export earnings in panel (c). The idea is to account for tourism and trade, which are other mobility channels through which the virus can spread. Encouragingly, the correlation between actual cases and our exposure measure remains stable at +0.65.

2.3 Validation with Data on Deaths
A concern with our validation exercise is measurement error in the confirmed case data. Testing capacity varies across countries, which makes confirmed cases a possibly mis-measured proxy for virus intensity. In Figure 3a, we restrict our validation sample to the 32 (of 146) countries that publicly share data on the number of tests they have conducted. This is useful secondary validation on the assumption that countries that publicly share their testing data have better testing coverage. The unconditional correlation between actual cases and our exposure measure is +0.53 in this sub-sample.
Lastly, in panel (b) of Figure 3b we reconstruct our exposure index replace confirmed cases with using COVID-related deaths in destinations to proxy \( \text{COV}_{d} \) in equation (1), and compare it to the actual number of COVID-related deaths in each migrant-sending country. We do so under the assumption that deaths are less likely to be measured with error than the measurement error introduced through cross-country variation in testing capacity and frequency. We compare this death-based exposure measure with the actual number of COVID-related deaths in each country. The unconditional correlation between these two measures is 0.66.

2.3 Highly Exposed Countries According to our Measure

The COVID risk exposure index we have constructed based on the insight that the disease spreads across borders through migrant-sending country. We can therefore use the index to identify countries that are likely more exposed to disease risk, given their migration links to other COVID-affected countries. We provide this list below, restricting our attention to the sub-sample of countries where the virus is not yet widespread (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.

We apply the same methodology to construct sub-national COVID-19 risk exposure indices for Bangladeshi districts and sub-districts, and Filipino provinces and municipalities. We make use of multiple datasets on migrants to evaluate how well different sources of data perform, because that is useful for replication in other countries with their unique data availability and limitations. We validate these sub-national indices with public health data on COVID-19 cases, quarantines, and distress calls to government hotlines.

3.1 Bangladesh District-Level Risk Exposure Index Using Airport Returnees

We first calculate a district-level risk exposure index using disembarkation card data collected from migrants returning to Bangladesh in February and March, 2020 by the Civil Aviation Authority of Bangladesh (CAAB). The cards record the destination country migrant return from as well as their home addresses, and districts of origin can be reliably extracted. Since CAAB intercepts migrants as they return home, it is the most accurate. It is likely the most reliable depiction of return migration pattern during this period of high exposure risk enabling us to track the spread of COVID-19 initiated by returnee migrants. Consequently, we consider this data as the ‘gold standard’ for the returnee migrant data needed for making the most accurate predictions of COVID-19 exposure risk at the district level.

We calculate the exposure index using CAAB data at the district level to show the inter-district difference in exposure to COVID-19 (see FIGURE 4). The heat map shows that district with larger share of returnee migrants from countries with high COVID-19 cases are more likely to be at higher risk of exposure.

3.2 Multi-Source Indices

We apply the same methodology to construct sub-national COVID-19 exposure indices using multiple datasets on migrants during this period of high exposure risk enabling us to track the spread of COVID-19 initiated by returnee migrants. Consequently, we consider this data as the ‘gold standard’ for the returnee migrant data needed for making the most accurate predictions of COVID-19 exposure risk at the district level.
We conduct a validation exercise, in line with our international analysis, to determine the credibility of our indices at the sub-national level. Due to data limitations on actual cases and deaths, we use data on the number of people quarantined and the number of distress calls made to the Bangladesh government’s COVID-19 national hotline number for validation purposes. Results indicate that the risk exposure calculated using returnee’s districts of origin strongly predict subsequent quarantines (correlation if +0.52, p-value<0.01; Figure 5(a)) as well as the distress calls to a Bangladeshi government hotline from those districts (correlation if +0.77, p-value<0.01; Figure 5(b))

These results are further supported by mobile-survey findings in the Cox’s Bazar district where returning migration and short trips made by the respondent are the strongest predictors of COVID-19 risk

**FIGURE 5: VALIDATION OF CAAB DATA**

![Graph](image)

**NOTES:** In panel (a), we illustrate the correlation between the exposure index and the number of quarantines. In panel (b), we illustrate the correlation between the exposure index and the distress calls within the seven divisions in Bangladesh. Divisional fixed effects are used to proxy for the regional variation in mobile-phone connectivity that may affect the incoming distress calls placed.

**3.1.1 Administrative Data from Migration Permits**

The airport returnee data described above may not be readily accessible in digitized form in many LMIC contexts and here we present some alternative administrative and survey data sources on migrants that might be available at various sub-national levels, thus broadening the scope of our analytical approach.

Firstly, we look at administrative data from registration of migration permits issued by the Bureau of Manpower, Employment and Training (BMET) in Bangladesh. This type of permit data provides more fine-grained detail on addresses and family contact information, which can facilitate contact tracing of returnee migrants. We find that the district origins of airport returnees from CAAB data is strongly correlated (correlation of +0.73, p-value<0.01) with the average migration permits issued by the BMET in the past five years (panel (2) in Table 2). The Overseas Employment Development Board (OEDB) in Philippines and the Ministry of Manpower in Indonesia are examples of similar agencies in other LMICs that can provide this type of data.

A similar validation exercise as before shows that the exposure risk calculated using migrants’ districts of origin strongly predict subsequent quarantines (correlation if +0.59, p-value<0.01; Figure 6(a)) as well as the distress calls to a Bangladeshi government hotline from those districts (correlation if +0.71, p-value<0.01; Figure 6(b)). This further validates the use of migrant permit data in this approach in the absence of returnee data.

**FIGURE 6: VALIDATION OF CAAB DATA**

![Graph](image)

**NOTES:** In panel (a), we illustrate correlation between the exposure index and the number of quarantines. In panel (b), we illustrate the correlation between the index and distress calls within the seven divisions in Bangladesh. Divisional fixed effects are used to proxy for the regional variation in mobile-phone connectivity that may affect the incoming distress calls placed.

While the prior analysis helps us to identify the districts in Bangladesh with the greatest exposure to COVID-19 risk, the applicability of this information for policy makers to implement lockdowns is constrained by the area and the population exposed. On average, the top ten high risk districts in Bangladesh have a population of 542,000 and a density of 2,925 per sq km. Consequently, we further break down the exposure analysis at the sub-district (upazila) level for more precise targeting of
policy (Figure 7). The detail in the BMET data enables us to get to this more fine grained level of analysis. The following heat map identifies upazilas that are at greater risk of exposure based on their exposure to migrants.

**FIGURE 7**

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

Validation of the sub-distinct indices result in pair-wise correlation of +0.47 (p-value <0.01; Figure 8(a)) between the log of exposure index and the number of distress calls received at the sub-distinct level.

**FIGURE 8**: Validation of BMET Sub-District-Level Exposure

In Figure 9, we illustrate cumulative migration-related COVID-19 exposure for a select set of sub-districts. Panel (a) focuses on the major metropolitan areas of Bangladesh.

**FIGURE 9**: Migrant Induced Cumulative COVID-19 Exposure Over Time (Sub-District Level)

Bangladesh first officially reported one COVID-19 infection on March 8 and over 140 Bangladeshi returnees returned from Italy on March 14.22 returned to their villages without adequate quarantine measure. We observe that the exposure risk of Dhaka, the capital city, starts rising in line with the announcement of the first reported case. Dhaka not only has a high exposure to migrants in general but more specifically to migrants from Malaysia, where the cases increased significantly in the second week of March 2020 that explains the early increase. Dhaka is closely followed by Comilla and Narayanganj, which have high exposure to international migrants. While Narayanganj is exposed to migrants from Malaysia, they only have a third as many migrants compared to Dhaka, which explains a later take off in exposure. It was also exposed to Singapore, where cases have largely been contained until very mid-April. Comilla also has a similar profile thus showing a similar trend. Also, both these cities are heavily exposed to migrants in Saudi Arabia, where COVID-19 cases increased at a later date. Gazipur, Chittagong and Sylhet follow close behind with high increases in cases expected in the following week where returnee migrants from Singapore and the countries in the Gulf Cooperation Council delays the cumulative exposure. Returnees from Italy to most of these metropolitan areas were also confirmed to have been COVID-19 positive.23

3.1.2 Survey Data from HIES

Returnees are also predicted by migrants identified in the nationally representative Household Income and Expenditure Survey (HIES) 2016.24 We find that there is also a strong correlation (correlation of +0.46, p-value<0.01) between returnees and migrants identified in the HIES at the district level (panel (1) of Table 2). This helps to further broaden the scope of our approach since national surveys in the HIES-style is publicly accessible for many countries in the LMIC category. For example, this includes every country where the World Bank has conducted Living Standards Measurement Study (LSMS) surveys or research bodies such as International Food Research Institute (IFPRI) may have conducted national level surveys.

Validation of the exposure risk calculated using migrants’ districts identified by HIES data show that there is significant correlation with subsequent quarantines (correlation of +0.51, p-value<0.01; Figure 10(a)) as well as the distress calls to a Bangladeshi government hotline from those districts (correlation if +0.54, p-value<0.01; Figure 10(b)). This further validates the use of widely available HIES data in this approach in the absence of returnee data. This is well illustrated in recent applications of using survey data on migrant distributions available for

---

24. A full description of the data is provided in the Online Appendix: Link Here
India, Bangladesh and Pakistan (Lee et al. 2020) to predict COVID-19 outbreaks.

The results vary depending upon the level of analysis. The log of the average of 2015 to 2019 migrants from BMET performs the best in predicting log of CAAB arrivals at the district level in both simple and multiple regression models. Results are statistically significant at the conventional level. However, when the analysis is carried out at destination and district-destination levels, HIES migrant data performs better than the BMET 5-year average.

The HIES survey data is better able to reflect the migrant stock while BMET, being an administrative registration of outgoing migrants, reveals the flow of migrants to a specific destination from a district in a given year. While averaging the data from 2015 to 2019 corrects for some of that, it is unable to fully compensate for the stock measure, especially at smaller units of observation such as destination and district-destination.

### 3.2 Case Study: Philippines

In the following case study of Philippines, we apply our index to calculate the risk exposure of the Philippines at the province and municipality level using administrative data on international migrants. Data for this analysis comes primarily from the Overseas Worker Welfare Administration (OWWA), the government agency tasked with ensuring the well-being of overseas migrants and their families while actual case data comes from the Government of the Philippines26.

The following heat maps (Figures 11(a) and (b)) show the risk exposures at the province and municipality level along with the number of actual cases reported. The municipality level analysis is able to provide us with a more fine-grained identification of the areas that are at highest risk of exposure to COVID-19.

#### Notes

26. We thank Dean Yang and Caroline Theoharides, who have worked extensively on migration patterns from the Philippines for sharing this data with us to facilitate the analysis.

---

**Table 2: Comparing CAAB Returnee Data with HIES and BMET Migrant Data**

<table>
<thead>
<tr>
<th>(Log) Number of Arrivals from CAAB Data (At the District Level)</th>
<th>District Level</th>
<th>Destination Level</th>
<th>District-Destination Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Log) Migrants (HIES 2016)</strong></td>
<td>0.464***</td>
<td>-0.0113**</td>
<td>0.702***</td>
</tr>
<tr>
<td><strong>(Log) Migrants (BMET, 2015-2019 Avg)</strong></td>
<td>-5.54</td>
<td>-0.0113</td>
<td>-0.27</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-3.71</td>
<td>-0.744***</td>
<td>-0.535***</td>
</tr>
<tr>
<td><strong>R-Squared</strong></td>
<td>0.081</td>
<td>4.51</td>
<td>-7.13</td>
</tr>
</tbody>
</table>

**Notes:** Number of observations at the district level is 64 for all datasets, thus corresponding with the total number of districts in Bangladesh. Number of observations at the destination and district-destination level varies depending on the dataset used.

---

25. A full description of the data is provided in the Online Appendix: Link Here.
As in the prior analysis, we use actual case data (where it is available) to validate the credibility of the index. The index significantly predicts the province-level COVID-19 cases confirmed by the Filipino government (correlation +0.73, p-level<0.01; Figure 13(a)) as well as municipality-level COVID-19 cases (correlation +0.65, p-level<0.01; Figure 13(b)).

4 Summing Up/ Way Forward (TBU)

We expect the analysis and validation checks to aid policy makers—especially those who are constrained by adequate testing capacity—to prioritize areas for rapid action. This identified high risk and susceptible areas need to enhance hospital and screening capacity as well as direct the flow of medical resources—in addition to imposing stringent lockdown measures to reduce the disease spread, known as “flattening the curve.” Moreover, these vulnerable areas need immediate social protection measures and other support for specific income groups. The global policy response 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.

FIGURE 13: Validation of Province-Level Exposure

NOTES: In panel (a), using province level data from the Philippines, we validate our index using actual case data. We get a correlation coefficient of 0.69 and exclude provinces for which we have no actual case data.

REFERENCES


Contact: ahmed.mobarak@yale.edu. We are grateful to excellent research assistance from Salim Benhachmi, Avinno Faruk, Matt Krupoff, Coyan Mobarak, Neeti Zaman, and to Caroline Theocharides, Siu Yuat Wong and Dean Yang for sharing data from the Philippines, and a2i (Government of Bangladesh) for sharing distress call data.