The impact of domestic travel bans on COVID-19 is nonlinear in their duration

Appendix: For online publication

Fiona Burlig, Anant Sudarshan, and Garrison Schlauch*

April 9, 2021

*Burlig: Harris School of Public Policy, University of Chicago, National Bureau of Economic Research, and Energy Policy Institute at the University of Chicago (EPIC), Chicago, IL 60637 (burlig@uchicago.edu); Sudarshan: Energy Policy Institute at the University of Chicago (EPIC), Chicago, IL 60637 (anants@uchicago.edu); Schlauch: Energy Policy Institute at the University of Chicago (EPIC), Chicago, IL 60637, (gschlauch@uchicago.edu). We thank Susanna Berkouwer, Eyal Frank, Amir Jina, Erin Kelley, and Gregory Lane for helpful comments, and the crowdsourcing teams at https://www.covid19india.org/ and https://github.com/GoogleCloudPlatform/covid-19-open-data for producing public information on COVID-19 outbreaks. All data and code required to reproduce the results in this paper are available from https://github.com/fburlig/BSS_COVID_migration.

1 Data

In this section, we describe our data acquisition and processing for the COVID-19 and migration data and travel ban information used in our analysis. The COVID-19 data are collated by open-source projects from official government sources and press reports. The migration data are gathered from official government census reports. The travel ban information is collected from government orders, official press releases, and news agencies.

1.A India

We use epidemiological data from COVID19-India (COVID-19 India Org Data Operations Group, 2020), an open-source project that scrapes data from various government agencies, official press releases, and news agencies. These sources include the Indian Council of Medical Research, the Ministry of Health and Family Welfare, the National Disaster Management Authority, and a variety of state-specific government agencies. Additional data come from state press bulletins, the Press Information Bureau, the Press Trust of India, Asia News International reports, and official government Twitter accounts.

Since only state-level COVID-19 data are available for Delhi, Telangana, Sikkim, Assam, Goa, and Manipur—all relatively small states—we aggregate their migration data to the state level and treat each state as its own administrative unit in the analysis. We also make three sample restrictions. First, we exclude the Andaman and Nicobar Islands and Lakshadweep. These are small island Union Territories comprising less than 1% of India's total population (Registrar General of India, 2011; WorldPop, 2020) and are difficult to reach from the mainland. Second, we drop districts with large outbreaks pre-existing the arrival of migrants, defined as districts whose cumulative cases before May 1st, when stranded migrants were first able to return home from some areas of the country (Ministry of Railways, Press Information Bureau, 2020), were in the 99th percentile. This helps to avoid confounding ongoing COVID-19 outbreaks with the effect of returning migrants. Lastly, we exclude 10 districts whose population is at least 80% urban according to the 2011 Census of India (Registrar General of India, 2011). These districts are predominately urban areas associated with large metropolitan cities and are not representative of migrant workers' more rural home districts.

The data are generally high quality but require a few cleaning steps, which we describe below:

Negative new daily cases, deaths, or tests. We observe negative cases, deaths, or tests in 0.6% of district-date observations from March 2nd through September 30th. Negatives may occur if more recent reports correct prior ones, such as if the number of previously reported cases was overstated. We set negative district-date observations to missing and evenly adjust prior data by the total reported reduction.

Inconsistent reporting of deaths. Due to a reporting change in the raw data, we observe an unrealistically large spike in deaths in Maharashtra and Delhi on June 16th, 2020, consistent with a known reporting backlog (The Hindue Net Desk, 2020). On that day, Maharashtra reported 1,409 new deaths and Delhi reported 437. The next highest daily deaths reports for the two states by the end of our sample are, 515 on August 15th and 129 on June 12th, respectively. We account for this reporting backlog by evenly distributing the spikes in deaths over the prior seven days. This does not impact our cumulative deaths estimates since every day in this range falls within the Phase 2 release period.

Large testing reports. The raw data occasionally report an unrealistically large number of tests on a given day. This commonly occurs in two instances. First, we find some large positive spikes in reported tests immediately followed by large negative spikes one or two days later, and vice versa. This likely represents a misreporting that is later corrected. We reduce the positive value by the negative one and set the latter to missing.

Second, we find some large positive spikes in tests on the first day a given district reports tests. For instance, 88% of districts report positive cases before tests, and 48.4% of the 83.9% of districts that report any tests report the greatest number by the end of our sample on the first day. This likely represents reporting of all tests conducted to that date on one day. To remove these spikes, we evenly distribute tests reported on the first day over all preceding days through the day prior to the first confirmed case. For example, if a district reports its first confirmed case on March 20th but does not report its first test until March 23rd, we evenly divide the tests reported on March 23rd between March 19th and March 23rd.

After making the above two corrections, we top-code tests above the 99th percentile in each district to remove any remaining outliers.

We collect migration data from the 2011 Census of India (Registrar General of India, 2011). We collect travel ban information from government orders (Indian Ministry of Home Affairs, 2020; Government of Maharashtra, Department of Revenue and Forest, Disaster Management, Relief and Rehabilitation, 2020), official press releases (News on AIR—News Services Division, All India Radio News, 2020), and news reports (Mehta, 2020).

1.B Indonesia

We use epidemiological data from COVID-19 Open-Data (Wahltinez et al., 2020), an open-source Github project that collates data from official and government sources and other open-source projects which collect similar data. For Indonesia, these data are scraped from the Indonesian Komite Penanganan COVID-19 dan Pemulihan Ekonomi Nasiona website (Komite Penanganan COVID-19 dan Pemulihan Ekonomi Nasional, Government of Indonesia, 2020).

Similarly to the data for India, we set negative province-date observations to missing and evenly adjust prior data by the total reported reduction.

We collect migration data from 2010 Census of Indonesia IPUMS International dataset (Minnesota Population Center, 2020), which acquires data from the country's national statistical agency, Statistics Indonesia, also known as BPS (Statistics Indonesia, Government of Indonesia, 2010). We collect travel ban information from news reports (Nurbaiti and Roidila, 2020; Beo Da Costa, 2020).

1.C South Africa

We use epidemiological data from COVID-19 Open-Data (Wahltinez et al., 2020). For South Africa, these data are obtained from Finmango (Finmango, 2020), another open-source project that collates official data sources, including government health departments and social media accounts.

Similarly to the data for India, we set negative district-date observations to missing and evenly adjust prior data by the total reported reduction.

We collect migration data from 2011 Census of South Africa IPUMS International dataset (Minnesota Population Center, 2020), which acquires data from Statistics South Africa, the country's national statistical agency (Statistics South Africa, Government of South Africa, 2011). We collect travel ban information from government orders (South African Government, 2020*a*) and official press releases (South African Government, 2020*b*).

1.D China

We use epidemiological data from COVID-19 Open-Data (Wahltinez et al., 2020). For China, these data are obtained from DXY COVID-19 Data (DXY COVID-19 Data, 2020), an open-source project that gathers data from Dingxiangyuan (Dingxiangyuan, 2020), an online community for physicians, pharmacies, facilities, and health care professionals.

Similarly to the data for India, we set negative province-date observations to missing and evenly adjust prior data by the total reported reduction.

We collect migration data from a report examining, among other things, the 2010 Census of China (Liu et al., 2014). We collect travel information from news reports (BBC News team, 2020; Zhong and Wang, 2020).

1.E Philippines

We use epidemiological data from COVID-19 Open-Data (Wahltinez et al., 2020). For the Philippines, these data are scraped from the Philippine Department of Health's website (Department of Health, The Philippine Government, 2020).

Similarly to the data for India, we set negative province-date observations to missing and evenly adjust prior data by the total reported reduction.

We collect migration data from the 2010 Philippine Census IPUMS International dataset (Minnesota Population Center, 2020), which collects data from the government's National Statistics Office (National Statistics Office, The Philippine Government, 2010). We collect travel ban information from official press releases (Abbey Gita-Carlos, 2020; Paculba, 2020) and news reports (CNN Philippines Staff, 2020; Marquez, 2020).

1.F Kenya

We use epidemiological data from COVID-19 Open-Data (Wahltinez et al., 2020). Although an explicit source is not provided for the Kenya data, these data most likely come from Finmango (Finmango, 2020), the same open-data organization that collected the South Africa epidemiological data.

Similarly to the data for India, we set negative county-date observations to missing and evenly adjust prior data by the total reported reduction. To align the level of aggregation of Kenya's COVID-19 data with that of its migration data, we interpolate daily COVID-19 cases from the county level to the district level (counties contain districts) by population. For example, if district d in county c comprises 20% of c's population, we assign 20% of c's daily COVID-19 cases to d.

We collect migration data from the 2009 Census of Kenya IPUMS International dataset (Minnesota Population Center, 2020), which collects data from country's National Bureau of Statistics (National Bureau of Statistics, Government of Kenya, 2009). We collect travel ban information from official press releases (Ministry of Health, 2020*a*,*b*).

1.G Census migration data

For Indonesia, the Philippines, South Africa, Kenya, and India, the census migration data include region specific numbers of domestic migrants who recently moved to COVID-19 hotspots. Aside from India, the migration stocks in each of these countries are at the same level of aggregation as the COVID-19 data used in our analysis.

The 2011 Census of India, on the other hand, only provides state-level information on recent movers to Mumbai (including other districts within Mumbai's state of Maharashtra), which is a higher level of aggregation than the districts in our sample. To allocate migrants across districts, we distribute state-level counts among districts within each state by population. Formally, we calculate the number of Mumbai migrants from each district d within state s by the equation

$$m_{ds} = \frac{m_s \times p_{ds}}{\sum_{d \in s} p_{ds}},$$

where m_s is the number of Mumbai migrants from state s and p_{ds} is the population of district d in state s.

For China, we only have data on the total number of migrants who recently moved provinces throughout the country. To approximate the number of inter-provincial migrants in Hubei—the province we designate as the country's hotspot—we multiply the ratio of Hubei's population to that of China as a whole by the total number of inter-provincial migrants in the country. Similarly to how we distribute migrants within each of India's states, we then distribute these migrants across China's other provinces by population. Although having more specific data on the distribution of China's migrants would be ideal, given the very low number of COVID-19 cases reported across the country after all travel ban releases from Hubei (Wahltinez et al., 2020), this is very unlikely to substantially affect our empirical estimate in Figure 3, Panel B.

Since the census reports for Indonesia, the Philippines, South Africa, and China include the number of movers in the five years prior to enumeration, and the reports for Kenya and India include the number of migrants in the year prior and four years prior to enumeration, respectively, we standardize to a central five year period prior to the year of enumeration. We do so by scaling the number of migrants in the latter two countries by a simple linear factor: five for Kenya, and five-fourths for India. Although the actual number of migrants may be nonlinear in duration from the time of enumeration, reassuringly, the number of migrants in India who moved states within four years prior to enumeration (13,521,489) is roughly four times the number who moved within one year prior to enumeration (3,531,612) (Registrar General of India, 2011).

2 Econometric analysis

Here, we provide sensitivity analyses to assess the robustness of our main empirical results shown in Figure 2, Panel D, and Figure 3, Panel B. These analyses are summarized in Figures A1, A2, and A4, and can be divided into four parts:

1. Sensitivity to different controls
2. Sensitivity to a linear trend adjustment
3. Sensitivity to the exclusion of migrants
4. Permutation inference test

We conclude by describing our main empirical results for India using COVID-19 deaths as the dependent variable rather than COVID-19 cases.

2.A Sensitivity to different controls

Mumbai natural experiment We first probe the robustness of the main result: the number of excess COVID-19 cases per 1,000 migrants in Phase 1, Phase 2, and Phase 3 districts 30 days after Mumbai's travel ban releases. Using our preferred specification, shown in the first row of Figure A1, we estimate that ban releases led to 31.53 excess cases per 1,000 migrants in Phase 1 districts (significant at the 1% level), 409.83 in Phase 2 districts (significant at the 1% level), and 12.96 in Phase 3 districts (not significant at conventional levels). The distribution of these estimates illustrates a parabolic relationship domestic travel bans in Mumbai and COVID-19 cases. We find relatively more cases per 1,000 Mumbai migrants after the intermediate length Phase 2 release than for the relatively short and long ban releases.

In our preferred empirical approach, we measure the number of Mumbai migrants in each district using data on recent movers reported in the 2011 Census of India (see Section

1.G). We also include two sets of fixed effects to control for unobserved factors that may influence the spread of COVID-19. First, we control for time invariant factors that vary across districts, such as demographics, income, and healthcare quality, through the inclusion of a full set of district fixed effects. Second, we include date-of-sample fixed effects to account for factors that are constant across districts but vary over time, such as under-detection of COVID-19 cases across the country, seasonality that may affect disease spread, and changes in disease containment measures.

We present several alternatives to this specification. In the second and third rows of Figure A1, we adjust the stringency of our time controls. In the second row, we remove the date-of-sample fixed effects. In the third row, we replace the date-of-sample fixed effects with a polynomial in sample date, which accounts for overall trends in COVID-19 cases over our sample period.

The results are generally robust to these varied controls. The point estimates and significance levels are similar without the date-of-sample fixed effects (Phase 1 = 26.33, significant at the 1% level; Phase 2 = 411.75, significant at the 1% level; Phase = 3.51, not significant at conventional levels) and—aside from the statistical significance of the Phase 1 estimate—are also similar after replacing the date-of-sample fixed effects with a polynomial in sample date (Phase 1 = 17.22, not significant at conventional levels; Phase 2 = 406.50, significant at the 1% level; Phase 3 = 5.78, not significant at conventional levels). Note that the date-of-sample fixed effects included in our main specification are more conservative than a polynomial in sample date, as they non-parametrically account for date-of-sample specific confounders.

In the fourth row of Figure A1, we include a district-specific testing control. This is defined as the sum of tests conducted over the previous 7 days in each district. We use this lagged measure to account for the delay between when tests are conducted and results become available. Due to the sporadic nature of reporting in some districts at certain points in time, we set this sum to missing if tests are not reported within the prior week, or if tests are not reported past a given point in the sample.

The inclusion of the testing control provides similar estimates as our main specification: Phase 1 = 34.65, Phase 2 = 414.30, and Phase 3 = -44.53. As expected, this control also somewhat increases the precision of our results, though produces the same general significance levels as before. The Phase 1 estimate is still significant at the 1% level, the Phase 2 estimate continues to be significant at the 1% level, and the Phase 3 estimate remains insignificant at conventional levels.

Cross-country comparisons Next, we examine the robustness of our cross-country comparisons: the number of excess COVID-19 cases per 1,000 migrants in Indonesia, South Africa, India (Phase 1), the Philippines, China, and Kenya. Using our preferred specification, we find a parabolic relationship between the duration of travel bans in hotspots and the number of excess COVID-19 cases per migrant in non-hotspots 30 days after travel ban releases (Figure 3, Panel B). These estimates are reproduced in Figure A2, where the parabolic curve drawn in blue is an unweighted quadratic function fit to the point estimates for our main specification.

As with our robustness checks for the Mumbai natural experiment, we adjust the stringency of our time controls. We plot unweighted quadratic functions fit to the point estimates from alternative specifications with gray curves in Figure A2. These curves display a similar parabolic relationship between travel ban durations and COVID-19 cases after removing the date-of-sample fixed effects (solid curve) or replacing them with a polynomial in sample date (short dashed curve).

2.B Sensitivity to a linear trend adjustment

Mumbai natural experiment Here, we examine the effect on our main results of removing the 'rotation' adjustment discussed in the main text. This adjustment accounts for linear trends in the pre-release coefficients, correctly removing the effect of any residual pre-trend that may have remained after fitting our main regression model.

Due to the relatively flat pre-release trends for Phases 1 and 2, removing this adjustment does not substantially affect their phases: Phase 1 = 28.07, significant at the 1% level, and Phase 2 = 596.20, significant at the 1% level. Failing to account for the relatively steep Phase 3 pre-release trend markedly increases the Phase 3 estimate to 312.15, which is significant at the 1% level. Still, we observe a parabolic relationship between travel ban durations and COVID-19 cases.

Looking at Figure 2, Panel C, it is clear that cases in Phase 3 districts do not exhibit a clear trend break (increase) after the travel ban release once we account for pre-release trends in our estimated coefficients. This lends credence to our decision to adjust for these trends. The actual Phase 3 treatment effect appears to be indistinguishable from zero.

Cross-country comparisons We also examine the effect of removing the 'rotation' adjustment from our cross-country comparisons. We find a similar parabolic relationship between bar duration and excess COVID-19 cases per 1,000 migrants. This relationship is shown by the long-short dashed curve in Figure A2.

2.C Sensitivity to the exclusion of migrant data

Lastly, we examine robustness to excluding migrant counts from our main specification. This is equivalent to running the same event study regression as our main specification without the migrants_d term:

$$I_{dt} = \sum_{s=-30}^{30} \beta^s 1[\text{days to release} = s]_{dt} + \alpha_d + \delta_t + \varepsilon_{dt}, \tag{1}$$

where the other terms in the equation are defined as before. As described in the main text, we then adjust the $\hat{\beta}^s$ coefficients for linear trends in the pre-release coefficients.

To produce estimates comparable to the specifications discussed above, for each phase we divide the sum of the adjusted post-release coefficients by the number of Mumbai migrants:

$$\text{excess cases per migrant} = \frac{\sum_{s=0}^{30} \tilde{\beta}^s}{N_{mig}}. \tag{2}$$

Consistent with classical measurement error, removing migrants from the specification significantly attenuates our results: Phase 1 = 0.17; Phase 2 = 130.10; and Phase 3 = 1.67. None of these estimates are significant at conventional levels. This helps to validate our empirical design and supports our hypothesis that incorporating information on migrants' likely travel destinations is highly relevant to empirical examinations of the effect of return migration on the spread of COVID-19.

2.D Permutation inference test

As an additional robustness check, we perform a permutation inference test on our main empirical model for India. This allows us to test whether our model is spuriously generating results when there is no true effect of travel ban releases on COVID-19 cases. These tests are also informative about our confidence intervals, serving as a non-parametric significance test.

To carry out this exercise, we hold the number of migrants and lockdown release phase constant. We then randomly assign each district's complete time path of COVID-19 cases to a different district without replacement while preserving the ordering of days within the data. This allows us to test whether trends might generate spurious correlations between travel ban releases and COVID-19 cases. We then re-estimate our main specification, including the linear pre-release trend adjustment, and store the point estimate from each draw. We repeat this process 10,000 times for each phase.

Figure A4 plots the resulting placebo estimates in comparison with our estimates that use the real data. Our real estimates of excess cases per 1,000 migrants are on the far right tail of the Phase 1 and Phase 2 distributions, and near the center of the Phase 3 distribution. For Phase 1, our real estimate is larger than all of the placebo draws; for Phase 2, it is larger than all but 8 placebo draws; and for Phase 3, it is only larger than 4,395 of the placebo draws. These results are consistent with the significance levels reported in the previous section and give us confidence that our estimates are not being spuriously generated. Furthermore, the implied permutation inference $p-$values on our estimates for Phases 1, 2, and 3 are < 0.0001, 0.0008, and 0.4395—if anything, smaller than our two-way clustered standard errors suggest.

2.E Empirical results using COVID-19 deaths as the dependent variable rather than COVID-19 cases

We focus on COVID-19 cases rather than deaths in the main text. We do this because, as described in Section 1.A, the deaths data exhibit much more significant reporting irregularities than the cases data. In addition, the registration of deaths in India is known to be incomplete, especially in rural areas. Even when deaths are registered, the cause of death is not always noted (Burgess et al., 2017). Recorded deaths from COVID-19 may therefore undercount true fatalities.

Nevertheless, we also report results from regression estimates using deaths as an outcome variable (Figure A3). We obtain results following a similar pattern to the per-migrant case outcomes presented in Figure 2. Using our preferred specification, we again find a parabolic relationship between the duration of travel restrictions and the per migrant effect on home district COVID-19 deaths within 30 days of travel ban releases. Perhaps in part due to our 30 day post-event window, our estimates are significantly less precise for COVID-19 deaths than for cases, making it difficult to definitely evaluate this relationship for the former.

References

Abbey Gita-Carlos, Ruth. 2020. "NCR-wide quarantine meant to 'save lives': DILG." *Press Release, Philippine News Agency*, https://www.pna.gov.ph/articles/1096687.

BBC News team. 2020. "China coronavirus: Lockdown measures rise across Hubei province." *BBC News*. https://www.bbc.com/news/world-asia-china-51217455.

Beo Da Costa, Agustinus. 2020. "Indonesia eases coronavirus travel ban despite fears it's too early." *Reuters*. https://www.reuters.com/article/health-coronavirus-indonesia-travel/indonesia-eases-coronavirus-travel-ban-despite-fears-its-too-early-idUSL4N2CQ2AN.

Burgess, Robin, Olivier Deschenes, Dave Donaldson, and Michael Greenstone. 2017. "Weather, climate change and death in India." *University of Chicago*.

Churches, Timothy, and Louisa Jorm. 2020. "Flexible, freely available stochastic individual contact model for exploring COVID-19 intervention and control strategies: Development and simulation." *JMIR Public Health and Surveillance*, 3(6): e18965.

CNN Philippines Staff. 2020. "LIST: Areas outside Metro Manila impose travel restrictions due to COVID-19." *CNN*. https://cnnphilippines.com/regional/2020/3/15/Areas-outside-Metro-Manila-travel-ban-flights-limited-COVID-19-threat.html?fbclid=IwAR39JQEfJNII8oNMuRtd3XsX3_yRHiYCWo0Pt73vRyDqNUGVcd7owyBEwuw.

COVID-19 India Org Data Operations Group. 2020. "Dataset for tracking COVID-19 spread in India." https://api.covid19india.org/.

Department of Health, The Philippine Government. 2020. "COVID-19 Tracker." https://doh.gov.ph/covid19tracker.

Dingxiangyuan. 2020. "COVID-19 global pandemic real-time report." https://ncov.dxy.cn/ncovh5/view/en_pneumonia?from=dxy&source=&link=&share=.

DXY COVID-19 Data. 2020. "2019 Novel Coronavirus Outbreak Time Series Data." https://github.com/BlankerL/DXY-COVID-19-Data.

Finmango. 2020. "Africa dataset." https://finmango.org/data.

Fitzgerald, Oisin, Mark Hanly, and Timothy Churches. 2020. "`covoid`: COVID-19 open-source infection dynamics."

Government of Maharashtra, Department of Revenue and Forest, Disaster Management, Relief and Rehabilitation. 2020. "Order No. DMU/2020/CR.92/DisM-1." http://bombaychamber.com/admin/uploaded/NEWS%20Block/Mah%20Govt%20-%204.6.2020%20-%20amendment%20to%20guidelines.pdf.

Indian Ministry of Home Affairs. 2020. "Order No.40-3/2020-DM-I (A)." https://www.mha.gov.in/sites/default/files/PR_Consolidated%20Guideline%20of%20MHA_28032020%20%281%29_1.PDF.

Kolthur-Seetharam, Ullas, Daksha Shah, Jayanthi Shastri, Sandeep Juneja, Gagandeep Kang, Anup Malani, Manoj Mohanan, Gayatri Nair Lobo, Gajanan Velhal, and Mangla Gomare. 2020. "SARS-CoV2 serological survey in Mumbai by NITI-BMC-TIFR." NITI-Aayog, Municipal Corporation of Greater Mumbai and the Tata Institute of Fundamental Research.

Komite Penanganan COVID-19 dan Pemulihan Ekonomi Nasional, Government of Indonesia. 2020. "Peta Sebaran." https://covid19.go.id/.

Liu, Ye, John Stillwell, Jianfa Shen, and Konstantinos Daras. 2014. "Interprovincial migration, regional development and state policy in China, 1985–2010." *Applied Spatial Analysis and Policy*, 7(1): 47–70.

Marimuthu, S, Melvin Joy, B Malavika, Ambily Nadaraj, Edwin Sam Asirvatham, and L Jeyaseelan. 2020. "Modelling of reproduction number for COVID-19 in India and high incidence states." *Clinical Epidemiology and Global Health.*

Marquez, Consuelo. 2020. "'Travel authority' only for persons crossing provincial, regional borders – Eleazar." *Inquirer.net.* https://newsinfo.inquirer.net/1285714/travel-authority-only-for-persons-crossing-provincial-regional-borders-eleazar.

Mehta, Manthan K. 2020. "First train from Mumbai leaves with Uttar Pradesh migrants." *The Times of India.* https://timesofindia.indiatimes.com/city/mumbai/first-train-from-mumbai-leaves-with-uttar-pradesh-migrants/articleshow/75632159.cms.

Ministry of Health. 2020*a*. "Govt announces extra measures to prevent spread of coronavirus." *Press Release*, https://www.health.go.ke/govt-announces-extra-measures-to-prevent-spread-of-coronavirus-nairobi-monday-april-6-2020/.

Ministry of Health. 2020*b*. "President Uhuru lifts movement ban in three counties." *Press Release*, https://www.health.go.ke/president-uhuru-lifts-movement-ban-in-three-counties-nairobi-monday-july-6-2020/.

Ministry of Railways, Press Information Bureau. 2020. "Railways start Shramik Special Trains to move migrant workers, pilgrims, tourists, students and other persons stranded at different places due to lock down." *Press Release*, https://pib.gov.in/PressReleasePage.aspx?PRID=1620027.

Minnesota Population Center. 2020. "Integrated Public Use Microdata Series, International: Version 7.2." *Minneapolis, MN: IPUMS*, https://doi.org/10.18128/D020.V7.2.

National Bureau of Statistics, Government of Kenya. 2009. "Census." https://www.google.com/search?q=National+Bureau+of+Statistics%2C+Government+of+Kenya&oq=National+Bureau+of+Statistics%2C+Government+of+Kenya&aqs=chrome.0.69i59.136j0j7&sourceid=chrome&ie=UTF-8.

National Statistics Office, The Philippine Government. 2010. "2010 Philippine Census." https://www.google.com/search?q=National+Statistics+Office%2C+The+Philippine+Government&oq=National+Statistics+Office%2C+The+Philippine+Government&aqs=chrome.0.69i59.112j0j7&sourceid=chrome&ie=UTF-8.

News on AIR—News Services Division, All India Radio News. 2020. "Maharashtra govt permits resumption of inter-district bus transport service." https://newsonair.gov.in/News?title=Maharashtra-govt-permits-resumption-of-inter-district-bus-transport-service&id=397699.

Nurbaiti, Alya, and Riza Roidila. 2020. "Indonesia suspends all passenger travel across cities to prevent Ramadan exodus." *The Jakarta Post.* https://www.thejakartapost.com/news/2020/04/23/govt-temporarily-bans-passenger-travel-to-prevent-mudik.html.

Paculba, Recthie. 2020. "DILG-10 bares guidelines for locally stranded individuals." *Press Release, Philippine Information Agency*, https://pia.gov.ph/news/articles/1042008.

Registrar General of India. 2011. "2011 Census of India." www.censusindia.gov.in/2011census/migration.html.

South African Government. 2020*a*. "Disaster Management Act: Regulations to address, prevent and combat the spread of Coronavirus COVID-19: Amendment." https://www.gov.za/documents/disaster-management-act-regulations-address-prevent-and-combat-spread-coronavirus-covid-19.

South African Government. 2020*b*. "Minister Nkosazana Dlamini Zuma on directions of once off movement of persons during Coronavirus COVID-19 alert level 4." *Press Release*, https://www.gov.za/speeches/minister-nkosazana-dlamini-zuma-directions-once-movement-persons-during-coronavirus-covid.

Statistics Indonesia, Government of Indonesia. 2010. "2010 Census of Indonesia." https://www.bps.go.id/.

Statistics South Africa, Government of South Africa. 2011. "2011 Census of South Africa." http://www.statssa.gov.za/.

The Hindue Net Desk. 2020. "Coronavirus live updates | June 16, 2020." *The Hindu*. https://en.wikipedia.org/wiki/The_Hindu.

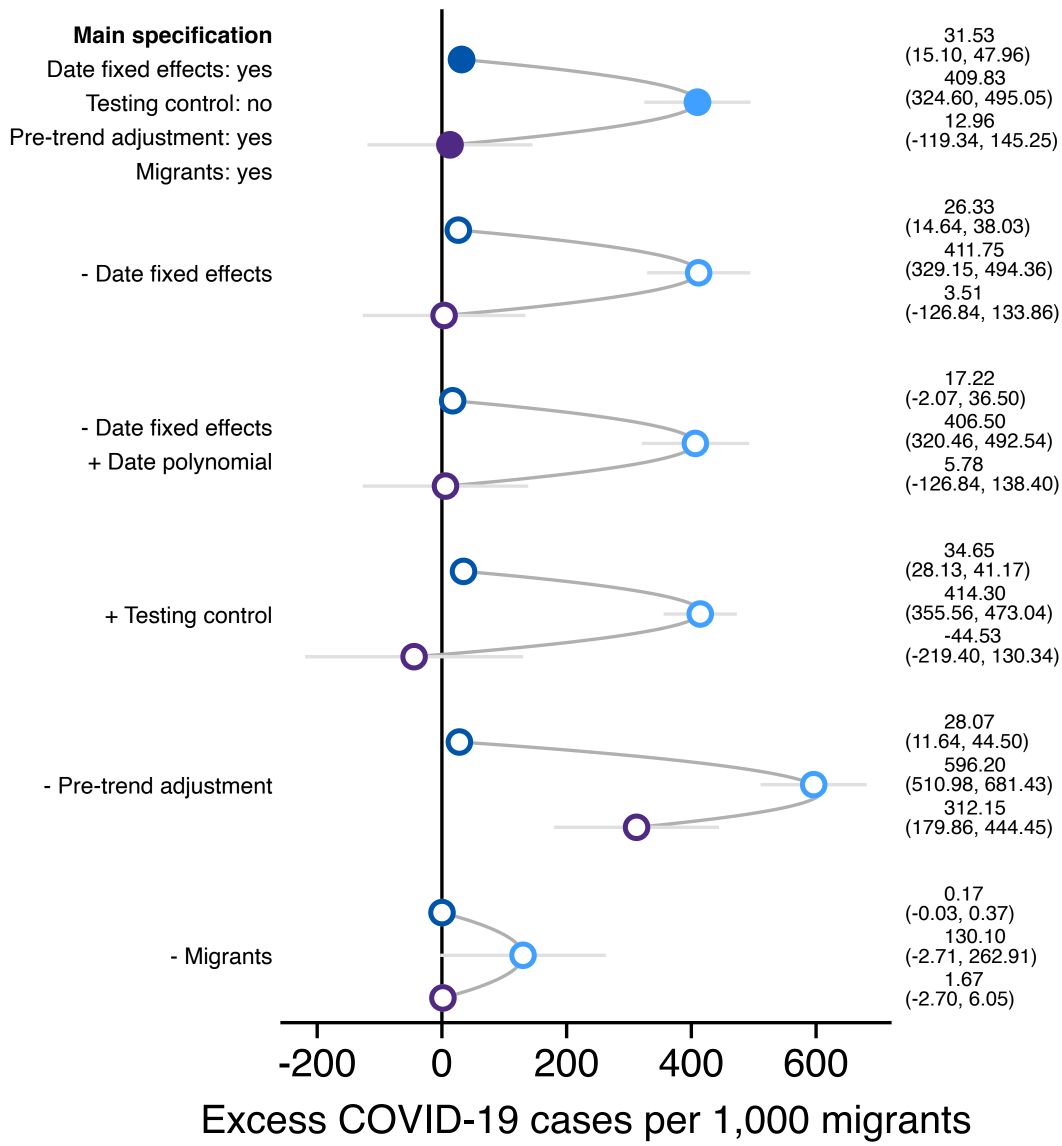
Wahltinez, O., et al. 2020. "COVID-19 Open-Data: curating a fine-grained, global-scale data repository for SARS-CoV-2." Work in progress.

WorldPop. 2020. "Constrained individual countries 2020 (100m resolution)." https://www.worldpop.org/geodata/listing?id=78.

Zhong, Raymond, and Vivian Wang. 2020. "China Ends Wuhan Lockdown, but Normal Life Is a Distant Dream." *New York Times*. hhttps://www.nytimes.com/2020/04/07/world/asia/wuhan-coronavirus.html.

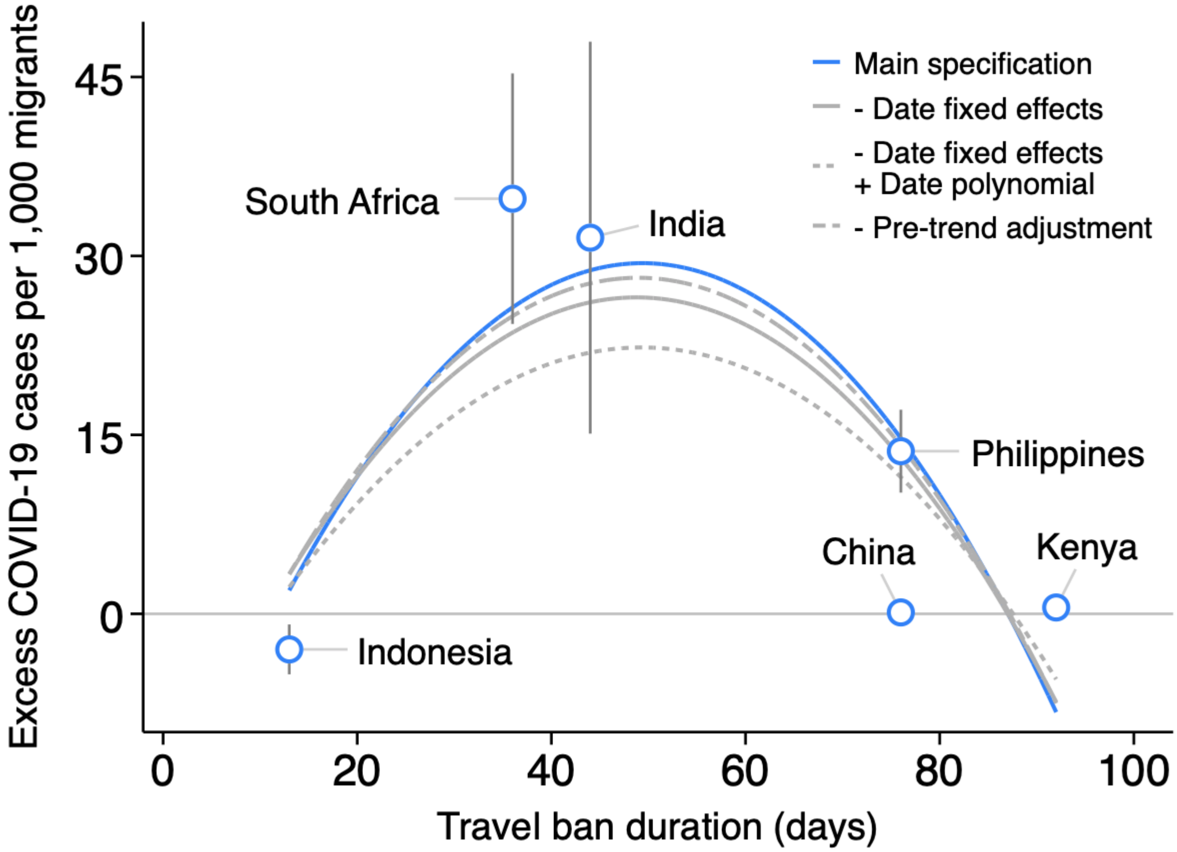
Figures

Figure A1: Econometric robustness for Mumbai natural experiment



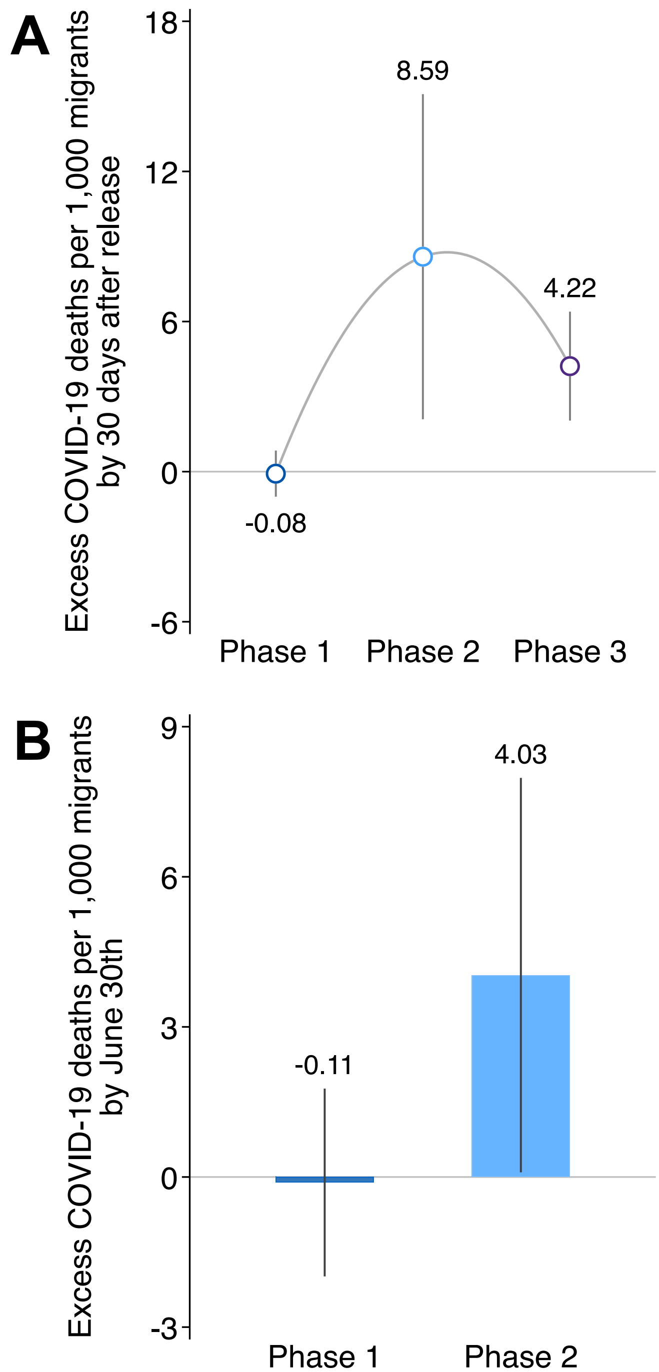
Notes: This figure shows excess COVID-19 cases per 1,000 Mumbai migrants in Phase 1 (short ban; dark blue), Phase 2 (intermediate ban; light blue), and Phase 3 (long ban; purple) districts 30 days after travel ban releases. The first row (solid dots) shows regression estimates from our main specification (Figure 2, Panel D). Each row presents robustness to an alternative specification, including adjusting the time controls (rows 2 and 3), adding a unit-specific testing control (row 4), removing the linear pre-release trend adjustment (row 5), and estimating a model without scaling by the number of migrants in each home district (row 6). 90% CIs are shown in gray horizontal bars. Parabolic curves are unweighted quadratic functions fit to the point estimates for each specification. See the main text for regression specifications and more detail

Figure A2: Econometric robustness for cross-country comparisons



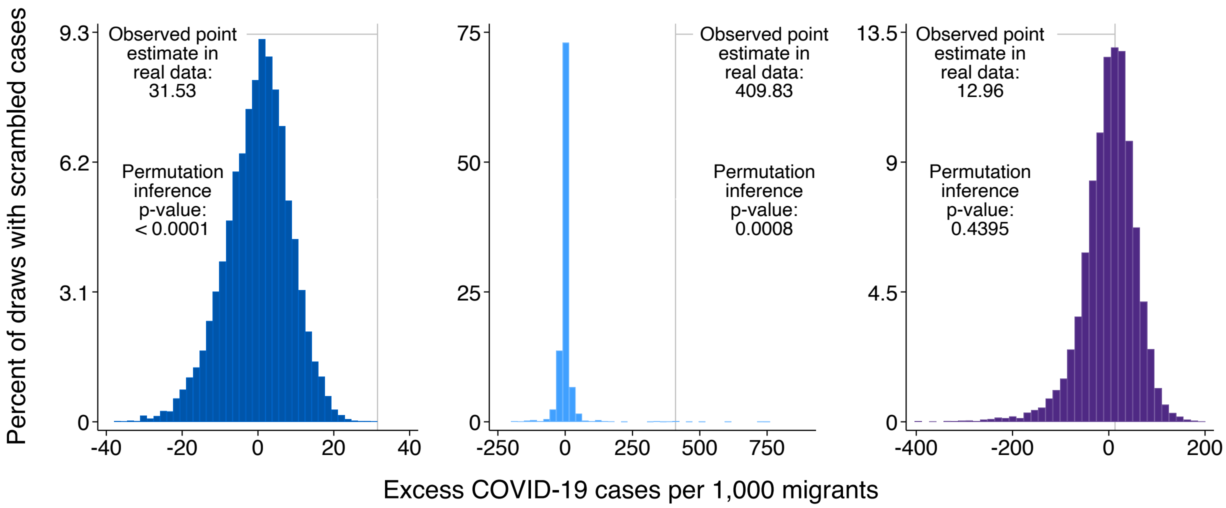
Notes: This figure shows excess COVID-19 cases per 1,000 hotspot migrants for each country in our sample 30 days after travel ban releases. Point estimates and 90% confidence intervals from our main specification (Figure 3, Panel B) are shown in dots and vertical gray bars, respectively. The blue parabolic curve is an unweighted quadratic functions fit to the point estimates for our main specification. The gray parabolic curves show robustness and are fit to point estimates from alternative specifications, including varying the time controls (solid and short dashed curves) and removing the linear pre-release trend adjustment (long-short dashed curve). See the main text for regression specifications and more detail.

Figure A3: COVID-19 deaths estimates



Notes: Panel A shows regression estimates of excess COVID-19 deaths per 1,000 Mumbai migrants in Phase 1, Phase 2, and Phase 3 districts 30 days after travel ban releases. Panel B shows regression estimates of excess COVID-19 deaths per 1,000 migrants in Phase 1 and Phase 2 districts by June 30th. Both panels display 90% confidence intervals as gray vertical bars. See the main text for regression specifications and more detail

Figure A4: Permutation inference, Mumbai natural experiment



Notes: This figure plots the results of a permutation inference exercise. We present the distribution of excess COVID-19 cases per 1,000 Mumbai migrants in Phase 1 (left), Phase 2 (middle), and Phase 3 (right) districts 30 days after travel ban releases by re-estimating our main specification on 10,000 randomly generated placebo datasets. To construct these datasets, we assign each district's full time series of COVID-19 cases to a randomly selected district without replacement. The vertical lines indicate the true point estimates for Phase 1 (31.53), Phase 2 (409.83), and Phase 3 (12.96) districts, generated using the real dataset (Figure 2, Panel D). The corresponding permutation inference $p-$values for Phases 1, 2, and 3 are < 0.0001, 0.0008, and 0.4395, respectively. See the main text for regression specifications and more detail.

Tables

Table A1: SEIR model parameters

Parameter	Value	Source
Source population	13,702,964	WorldPop (2020)
Sink population	16,642,462	WorldPop (2020)
Number of migrants	40,184 $\times$ 1.25	Registrar General of India (2011)
		WorldPop (2020)
Source positivity rate (July 15th)	32.7%	Kolthur-Seetharam et al. (2020)
Seed infections on March 25th	500	Assumed
R_0 source	3.75	Calibrated to match
		Kolthur-Seetharam et al. (2020)
R_0 sink	$0.5 \times R_0$ source	Similar to Marimuthu et al. (2020)
Latent period ($1/\omega$)	10 days	Churches and Jorm (2020)
		Fitzgerald, Hanly and Churches (2020)
Infectious period ($1/\lambda$)	10 days	Churches and Jorm (2020)
		Fitzgerald, Hanly and Churches (2020)

Notes: This table lists the parameter values used to calibrate our SEIR model, where source refers to Mumbai and sink refers to Phase 2 districts. Citations are included where applicable.

Table A2: Data sources

Country	COVID-19 data		Migration data		Population data	
	Admin level	Source	Year	Source	Year	Source
Indonesia	1	COVID-19 Open-Data[1]	2010	IPUMS[2]	2020	WorldPop[3]
		ID Komite COVID-19[4]		Statistics Indonesia[5]		
South Africa	2	COVID-19 Open-Data[1]	2011	IPUMS[2]	2020	WorldPop[3]
		Finmango[6]		Statistics South Africa[7]		
India	1 & 2	COVID19-India[8]	2011	ORGI[9,*]	2020	WorldPop[3]
					2011	ORGI[9,*]
China	1	COVID-19 Open-Data[1]	2010	Liu et al., 2014[10]	2020	WorldPop[3]
	1	DXY COVID-19 Data[11]				
Philippines	2	COVID-19 Open-Data[1]	2010	IPUMS[2]	2020	WorldPop[3]
		Philippine Dept. of Health[12]		Nat'l Statistics Office[13]		
Kenya	2[†]	COVID-19 Open-Data[1]	2009	IPUMS[2]	2020	Worldpop[3]
		Finmango[6]	2009	Nat'l Bureau of Statistics[14]		

Notes: For each country, this table lists the primary (1st row) and, where applicable, secondary (2nd row) data sources for the COVID-19, migration, and population data used in our analysis. [*]Office of the Registrar General & Census Commissioner, India. [†]Interpolated; see Section 1.F for details. *Sources:* [1]Wahltinez et al. (2020); [2]Minnesota Population Center (2020); [3]WorldPop (2020); [4]Komite Penanganan COVID-19 dan Pemulihan Ekonomi Nasional, Government of Indonesia (2020); [5]Statistics Indonesia, Government of Indonesia (2010); [6]Finmango (2020); [7]Statistics South Africa, Government of South Africa (2011); [8]COVID-19 India Org Data Operations Group (2020); [9]Registrar General of India (2011); [10]Liu et al. (2014); [11]DXY COVID-19 Data (2020); [12]Department of Health, The Philippine Government (2020); [13]National Statistics Office, The Philippine Government (2010); [14]National Bureau of Statistics, Government of Kenya (2009)

Table A3: Travel ban information, cross-country comparisons

Country	Hotspot(s)	Travel ban initiation	Travel ban relaxation	Travel ban duration
Indonesia	Jakarta[1]	Apr 24th[2]	May 7th[3]	13 days
South Africa	Cape Town[1]	Mar 26th[4]	May 1st[5]	36 days
India	Mumbai[6]	Mar 25th[7]	May 8th[8]	44 days
China	Wuhan (Hubei)[1,*]	Jan 23rd[9]	Apr 8th[10,†]	76 days
Philippines	NCR[1,*] & Cebu[1]	Mar 15th[11], 18th[12]	May 30th[13,14]	76 days
Kenya	Nairobi & Mombasa[1]	Apr 6th & 8th[15]	Jul 7th[16,‡]	92 days

Notes: For each country, this table lists the region(s) we designate as hotspots, the dates of travel ban initiations and relaxations, and the durations of travel bans, including citations where applicable. Where two travel ban initiation dates are listed for a single country, we use the earlier one to calculate travel ban duration. [*]Since the data are only available at the Admin 1 level, we use the Wuhan lockdown lift date rather than that for other regions of Hubei, which allowed travel two weeks earlier. [†]National Capital Region. [‡]Although the travel ban from Mandera was also relaxed at this time, Mandera does not report any COVID-19 cases in our data prior to July 7th.
Sources: [1]Wahltinez et al. (2020); [2]Nurbaiti and Roidila (2020); [3]Beo Da Costa (2020); [4]South African Government (2020*a*); [5]South African Government (2020*b*); [6]COVID-19 India Org Data Operations Group (2020); [7]Indian Ministry of Home Affairs (2020); [8]Mehta (2020); [9]BBC News team (2020); [10]Zhong and Wang (2020); [11]Abbey Gita-Carlos (2020); [12]CNN Philippines Staff (2020); [13]Paculba (2020); [14]Marquez (2020); [15]Ministry of Health (2020*a*); [16]Ministry of Health (2020*b*)