Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The effects of air transport mobility and global connectivity on viral transmission: Lessons learned from Covid-19 and its variants

Youngran Choi a, b, Li Zou a, Martin Dresner b

a David B. O’Maley College of Business, Embry-Riddle Aeronautical University, 1 Aerospace Boulevard, Daytona Beach, FL, 32114, USA
b Logistics, Business & Public Policy, R.H. Smith School of Business, University of Maryland, College Park, MD, 20742, USA

1. Introduction

On November 25th, 2021, scientists and health officials in South Africa announced the discovery of a new variant of COVID-19 that had more than 50 modifications from the original coronavirus. The new variant, named Omicron by the World Health Organization (WHO), was the fifth major COVID-19 variant declared since the initial outbreak of the coronavirus in December 2019 in Wuhan, China, following Beta, Alpha, Delta, and Gamma. To constrain the transmission of the new variant, travel restrictions were immediately imposed by countries around the world. For example, on November 26th, the U.S. announced limitations on foreigners entering the country from South Africa and neighboring states in Southern Africa. Other countries (e.g., Israel) banned the entry of all foreigners and/or prohibited their citizens from traveling to Southern Africa. Despite uncertainty about Omicron’s lethality, containment measures from governments were generally much faster and more aggressive than responses to the initial COVID-19 virus two years previously. However, by December 1st, 2021, Omicron had still been detected in more than 20 countries, including some of the countries that acted quickly with foreign restrictions.

Many public health researchers view border-control measures as essential to slowing viral spread (Mallapaty, 2020a). Reducing the speed of transmission gives scientists and health authorities time to prepare for the arrival of the virus, while determining anti-contagion measures, such as social distancing, mask wearing, and vaccination requirements. Other researchers are suspicious about the effectiveness of travel restrictions, suggesting that they tend to be imposed too late to prevent the global circulation of viruses and, therefore, have limited effectiveness (Tian et al., 2020; Russell et al., 2021). Moreover, travel restrictions targeting select countries may be viewed as unfair discrimination against people living in these countries. The threat of travel restrictions can impede the open flow of information about the discovery of viral variants; that is governments may choose not to report viral outbreaks to escape the imposition of travel bans against their citizens (Devi, 2020).

Several studies have estimated the efficacy of travel restrictions on mitigating or delaying viral transmission (e.g., Hurford et al., 2021; Chang et al., 2020; Yang et al., 2021). In analyzing 300 papers on viral transmissions, Grépin et al. (2021) categorize travel restrictions into four categories, including transportation suspension, border restrictions, entry/exit screening, and entry quarantines. In assessing outcomes of these restrictions, several measures are used, including number of observed cases, date of epidemic peak, speed and risk of transmissions, case doubling time, arrival time of the virus, reproductive number, and cumulative cases. Grépin et al. (2021) note that while some evidence shows the effectiveness of travel-related restrictions for impeding the spread of COVID-19, especially the Chinese domestic travel ban...
implemented in Wuhan, China during the early phases of the pandemic, it is still imperative to investigate the complex, dynamic effects of international air travel on COVID-19 transmission in the global context (Mallapaty, 2020b).

With Grépin et al.’s (2021) call for action in mind, we model factors impacting the air travel-induced international transmission of the COVID-19 virus and its variants. We collect data on international flights and seat capacities prior to the declaration of the pandemic by the WHO, and construct measures of the connectivity of the air transport networks for countries around the world. We then estimate the impact of global air transportation networks on the speed of viral transmission. Moreover, we adopt measurements developed by Brockmann and Helbing (2013) on effective distance for modeling an important factor impacting the speed of transmission. Finally, we test the effectiveness of a key policy variable, the reduction in total inbound seat capacity to a country, on slowing the time of viral transmission and its detection.

Our analysis builds on previous studies by using a large-scale database measuring viral transmission to more than 200 countries and territories around the world from the country where the virus was first discovered. Moreover, we measure transmission time for not only COVID-19, but four of the major variants that followed the initial virus discovery. Finally, we build our measure for effective distance based not only on direct air connections, but also on indirect connections from the country of viral origin.

Our findings provide evidence that effective distance can be a significant determinant influencing viral transmission. Specifically, after controlling for country-specific levels of healthcare infrastructure, we find countries that are poorly networked with the country where the virus is first discovered prior to the first viral discovery (i.e., countries that have greater effective distance from the point of viral outbreak), generally experience delays in the transmission and announcement of their first viral outbreak. Thus, effective distance buys countries time to plan their responses to the eventual arrival of the virus, supporting previous results from Sun et al. (2021a). However, policy initiatives to reduce international air capacity taken after the virus is first discovered to slow transmission into a country are largely ineffective, a finding that is consistent with Sun et al. (2021b).

Our study contributes to the growing literature that empirically investigates the impacts of air travel on viral spread and the effectiveness of travel restrictions on preventing transmission. The implications from our findings can help improve public policies with respect to travel restrictions and other measures to mitigate viral impacts from future pandemics.

The rest of the paper is organized as follows: Section 2 reviews the relevant literature on the role of air transport on viral transmissions and the effectiveness of travel restrictions in containing viral transmission. Section 3 describes the data collected for analysis and presents our variable development. Section 4 presents our empirical model. Section 5 provides estimation results and conducts supplementary analyses for robustness checks. Finally, Section 6 concludes the paper with a discussion of its limitations, policy implications, and future research.

2. Literature review

2.1. The effect of air travel mobility on viral transmission

Global aviation networks enable fast and convenient air travel, requiring less than 48 h to connect any two commercial airports worldwide. With air travel times shorter than the incubation periods for infectious diseases, the risk of viral transmission through undetected infected passengers is present on all flights. As a result, the connectivity enabled by air transport networks has facilitated the spread of viruses, including COVID-19 (e.g., Christidis and Christodoulou, 2020; Sun et al., 2020; Farzaneh et al., 2021), Ebola in 2014 (Bogoch et al., 2015), A/H1N1 influenza in 2009 (Hosseini et al., 2010), and SARS in 2003 (Bowen and Laroe, 2006).

The dissemination of infections through global airlines has been noted by Budd et al. (2009). Air transport systems contribute not only to viral spread through the mobility of people, but also to the spread of infectious diseases through the mobility of insects. Huang et al. (2012) note that the transmission risks from vector-borne diseases are moderated by climate and environmental factors. By combining spatial distribution data from four major vector-borne diseases with global air travel schedule data and location-specific climate data, Huang et al. (2012) employ a boosted regression tree mapping tool to estimate air travel-induced risks of imported infections. The authors then use their risk-assessment model to develop a geographic information system (GIS) to assess vector-borne disease risks.

Bowen and Laroe (2006) investigate the role of airline network accessibility in the spatial-temporal transmission of SARS. The initial outbreak of SARS was in Guangdong, China in February 2003. The virus spread to 25 countries or regions but was contained within five months. Using data for a sample of 207 large cities with scheduled air passenger services in March 2003, the authors find that air network accessibility, measured by the relative travel time for a one-stop connecting flight to a non-stop, direct flight, was a determining factor influencing the arrival time of SARS infections. Furthermore, the impact from this factor was found to diminish during the latter stages of the disease diffusion.

Hosseini et al. (2010) consider three variables predicting viral spread, including the availability of direct, non-stop air travel from Mexico, indirect, one-stop air travel from Mexico to a destination country, and total healthcare spending per capita at the destination country, to estimate the time-to-reporting of the first confirmed A/H1N1 case in 130 countries in 2009. Using a log-logistic survival model, their findings suggest that the ability of a country to rapidly detect, diagnose, and report viral infections is a critical element in constraining the global transmission of an emerging virus.

Brockmann and Helbing (2013) combine a susceptible-infected-recovered (SIR) dynamic model, an epidemiological framework for analyzing the contagion of an infectious disease, with a multi-node, multi-level global mobility network (i.e., air transport network), to assess viral transmission. The authors develop a unique measure, called “effective distance”, based on the most probable path between two nodes in a mobility network. Effective distance enables the calculation of the speed and arrival times of a contagion without the need for parameters, such as the viral reproduction rate and recovery rate. Brockmann and Helbing (2013) apply their method to simulate the diffusion of the 2009 H1N1 influenza virus and 2003 SARS virus. They find that effective distance successfully predicted the speed and arrival time of these two contagious diseases in the context of the global air transport mobility network.

2.2. The effectiveness of travel restrictions in containing viral transmission

There have been several studies that assess the efficacy of travel-related restrictions in mitigating and slowing the spread of outbreaks of infectious diseases, including Ebola (Bogoch et al., 2015), A/H1N1 (Hosseini et al., 2010), SARS (Bowen and Laroe, 2006), and MERS (Poletto et al., 2016). Bogoch et al. (2015) estimate the impact of travel restriction measures on reducing the diffusion of the Ebola virus. They find that exit screening of outbound travelers is relatively more effective than entry screening of inbound travelers. Hosseini et al. (2010) analyze the spread of the A/H1N1 influenza and validate the importance of implementing rapid detection, diagnoses, and reporting during the early stages of a pandemic in the country where the epidemic emerged, as well as in other countries that have strong travel connections to originating country.

Hosseini et al. (2010) provide support for the WHO’s recommendation against air travel restrictions that narrowly target the location of the initial viral outbreak. Instead, the authors suggest international health authorities target countries with both poor health care infrastructure and high air travel flows from the outbreak location. While air travel was
found to play a critical role in the risks of transmission of MERS, infection control measures could have reduced imported infection cases (Poletto et al., 2016). These measures include prompt identification of the infection, the effective isolation of patients, and public awareness of the virus (Poletto et al., 2016).

Chinazzi et al. (2020) conduct a comprehensive study investigating the effectiveness of the travel ban implemented in Wuhan, China on January 23, 2020, to impede COVID-19 transmission, as well as other restrictive measures taken by countries and airlines beginning in February 2020. The authors simulate several possible viral transmission scenarios using a Bayesian method and find that while the travel ban from Wuhan helped to slow the spread of COVID-19 within mainland China by 3-5 days, it had a greater impact on the viral transmission from mainland China to the rest of the world, reducing the outbound cases by up to 80% until mid-February. The authors then provide three important policy implications. First, travel restrictions imposed in Wuhan would have been more effective in slowing viral spread to other Chinese cities if implemented earlier. Second, the impact of travel restrictions imposed in Wuhan diminished over time. Third, the viral mitigation effect from travel restrictions, alone, is limited, and therefore should be combined with other anti-contagion public health policies and behavioral changes to impede transmission.

The effects of travel restrictions, social distancing, quarantines, lockdowns, and other anti-contagion policies on the growth rate of COVID-19 cases are empirically investigated by Hsiang et al. (2020). Using data at the city/province/state/regional levels for six countries (China, South Korea, Italy, Iran, France, and the U.S.) during the early stage of the pandemic (from January to April 2020), Hsiang et al. (2020) study the mitigation effects of various policies on the COVID-19 case growth rate. These effects vary across the different policies and on different populations, but the overall effects are statistically significant in reducing the COVID-19 growth rate. Along the same lines, Hurford et al. (2021) find that the implementation of both travel restrictions and physical distancing measures contributed to a 92% reduction of COVID-19 cases in the Canadian province of Newfoundland and Labrador over a nine-week period in spring, 2020.

Yang et al. (2021) simulate viral importation risks under various travel control measures for travel to Hong Kong. Their findings support a moderate relaxation of travel restrictions from locations with a low prevalence of infections (e.g., travel bubbles). They further suggest that increased importation risks from a travel relaxation may need to be offset by proactive measures, such as contact tracing, post-arrival testing, and the separation of quarantined travelers from the local population. Sun et al. (2021a) linked the effectiveness of flight restrictions with newly emerged variants, suggesting that bans were implemented too late to contain viruses at an emerging location and prevent viruses from traveling via air transport. Ding et al. (2021) formally identified an optimization model, proposing a framework for an early-stage transportation lockdown and quarantine (TLQP) problem and determining what transportation needs to be restricted during the early stages of a pandemic.

In summary, there have been several papers that model international viral transmission based on air transport networks, as well as the mitigation efforts taken to reduce viral spread. We contribute to this literature by modeling the spread of COVID-19 and four of its variants to determine the impact of aviation networks on the spread of these viruses. Following Sun et al. (2021a), we expand the scope of investigation from COVID-19 to its variants, while including a large sample of countries around the world in our analysis, including countries that only have indirect flight connections from the country of viral origin. Finally, we measure how a key policy variable, the decrease in air capacity flows into a country, impacts the speed and detection of viral spread.

3. Data and variable development

Four primary data sources are employed in our study. First, we retrieve country-specific data on newly confirmed COVID-19 cases from the Oxford COVID-19 Government Response Tracker (Hale et al., 2021). Using newly confirmed cases of COVID-19, we develop an arrival time variable for 224 countries to represent the elapsed days for the first confirmed case reported after January 1, 2020. Second, we compile the arrival time by country for four COVID-19 variants - Beta, Alpha, Delta, and Gamma from GISAID (Shu and McCauley, 2017). (See Table 1 for a list of the four variants analyzed in our study.) Third, we retrieve a stringency index compiled by Oxford University (OxCGRT). The index is used to represent lockdown measures enacted by countries to prevent viral transmissions. Lastly, we collect monthly schedule data for all the non-stop international passenger flights covering all countries worldwide from December 2019 to October 2020 from Diio Mi-Market Intelligence for the Aviation Industry (Diio, 2020). The flight schedule database is used to develop measures for air mobility and connectivity. To supplement the flight schedule data, we also collect passenger traffic data from the SABRE passenger booking system, using Sabre Market Intelligence 6.5 data portal.

3.1. Transmission speed and effective distance

Following Brockmann and Helbing (2013), we develop an effective distance measure to represent the most likely path for COVID-19 viral transmission through international air travel from China to other countries (and then from the country where each of the four major COVID variants was originally reported to other countries). Although COVID-19 was first reported in Wuhan, China, we consider all international outbound flights originating from China when calculating this variable, since Wuhan is well-connected internally to other cities in China (Zhang et al., 2020). We calculate effective distance measures based on both outbound and non-stop flights from China to the rest of the world during the pre-pandemic base case months of December 2019 and January 2020. In the following section, we first present the calculation of effective distance for those destinations that had direct flights from China, and its potential relationship with the arrival time of COVID-19. Then we present the calculations for effective distance for all other countries that had no direct flights but relied on connecting flights from China. Finally, we illustrate the arrival time of COVID-19 for our sample countries with direct and (only) indirect flights from China in Fig. 2.

For a destination country (denoted as Country j) that has direct, non-stop flights from China (denoted as Country i), its effective distance $d_{ij}$ from China in month $t$ is defined as the following:

$$d_{ij} = 1 - \ln PS_{ij}$$

Eq. 1

| Table 1 | COVID-19 variants (WHO, 2021) |
| --- | --- | --- | --- |
| WHO label | Country of first detection | Date of first detection | Date of designation as Variant of Concern (VOC) by WHO |
| Beta | South Africa | May 2020 | 18 December 2020 |
| Alpha | United Kingdom | September 2020 | 18 December 2020 |
| Delta | India | October 2020 | 11 May 2021 |
| Gamma | Brazil/Japan\(^a\) | November 2020 | 11 January 2021 |

Note: Omicron was not included in our study as it was designated as one of variants of concern (VOC) on November 24, 2021, after our study period.

\(^a\) The Gamma variant was first detected in people returning to Japan from Brazil’s Amazonas state. Therefore, we include Japan as a possible originating country along with Brazil (Callaway, 2021).

\(^1\) Includes self-governed territories, as well as countries.
where $PS_{ijt}$ represents the share of outbound seats (flights) from China to the destination country $j$ of the total outbound seats (flights) scheduled from China in month $t$.

In December 2019, there were a total of 8,678,191 outbound seats and 40,889 outbound flights from China to 70 countries worldwide, according to Dido’s passenger flight schedule data. Among those 70 destinations, the effective distance $d_{ij}$ by seats ranged from 2.94 (i.e., Japan) to 10.78 (i.e., Afghanistan). Countries with the smallest (largest) effective distances from China account for the largest (smallest) share of total outbound seats or flights.

Table 2 ranks the top ten countries by effective distance. The table also provides data on arrival time of the virus (number of days since January 1, 2020) and the ratio of effective distance (based on seat capacity) relative to arrival time, namely effective speed (Brockmann and Helbing 2013). Our analysis shows that for countries that had direct flights from China in December 2019, the correlation between the effective distance from China and COVID-19 arrival time at the destination country is 0.68 when effective distance is calculated based on seat share, or 0.65 when effective distance is calculated based on flight share. Similar correlation values can be calculated based on scheduled outbound flights or seats in January 2020.

Fig. 1 presents the total scheduled outbound capacity from China by number of flights and seats from December 2019 to April 2020. The figure illustrates the dramatic drop in outbound capacity from a peak in January 2020 to a trough in April 2020. For example, flights per month declined from over 40,000 to under 5,000. The figure also indicates the number of countries served by direct flights from China. In this case, the decline was not as dramatic, from 71 in January 2020 to 49 in April 2020, providing an indication that frequencies were sharply trimmed but not entirely cut to many countries by April 2020.

In addition to the 70 countries that had direct flights from China in December 2019,147 countries were connected to China via one-stop connections, and 6 additional countries via two-stop connections, based on our flight schedule data. We calculate effective distance for these countries following Brockmann and Helbing (2013); that is, we first calculate effective distance between all possible pairwise countries (i.e., nodes) and then calculate the sum of effective distances across all two-flight segments that connect China to destination countries that do not have direct connections. Effective distance on the first flight segment from China (Country $i$) to the connecting country $k$ is calculated using Eq. (2) and denoted as $d_{ikt}$ in month $t$. Effective distance is calculated for the second flight segment of the connecting route, from country $k$ to the destination country $j$ in month $t$ by Eq. (3):

\[
d_{ikt} = 1 - \ln PS_{ikt} \tag{2}
\]

\[
d_{ijkt} = 1 - \ln PS_{ijkt} \tag{3}
\]

where $PS_{ijkt}$ represents the share of the outbound seats scheduled from country $k$ to the destination country $j$ of the total outbound seats scheduled from country $k$ to the rest of the world in month $t$.

We then calculate the sum of the effective distances for all one-stop connecting routes from China (Country $i$) to the destination country $j$ (i.e., $\sum_{k=1}^{N} d_{ikt} + d_{ijkt}$) and rank the sums across all possible one-stop connecting routes. Following Brockmann and Helbing (2013), effective distance from China to the destination country $j$ is defined as the connecting route with the shortest path; that is, the smallest sum of the effective distances across the two flight segments from China to country $k$ and from country $k$ to country $j$. Using a similar method, effective distance is also calculated for destination countries that require at least two-stop connections from China.

Based on the flight schedule data from December 2019, we find that effective distance for the 147 countries that can be reached from China via one-stop connection ranges from 8.53 (i.e., Dominican Republic via the U.S.) to 17.76 (i.e., Sao Tome and Principe via Portugal) (see Fig. 2).

We develop similar effective distance measures for outbound seats scheduled from South Africa using April 2020 data, from the U.K. using August 2020 data, from India using September 2020 data, and from Brazil and Japan using October 2020 data; that is, the countries where the four first major variants of COVID-19 were initially detected. For each of the variant calculations, we use seat schedule data for one month prior to the date when a variant was first detected to compute effective distance (see Table 1 for date and country of first detection).

### 3.2. Transmission speed and travel restriction policy

To account for travel restriction policy initiatives, we compute the percentage reduction in total inbound international seats from the pre-pandemic base period, December 2019, to the period prior to when the infection was first reported. For COVID-19, we use the percentage seat reduction from the base period to February 2020, since most governments did not require cuts in aviation capacity until after February 1. Inbound seat capacity reductions from the base period to April 2020, August 2020, September 2020, and October 2020 are used for variants Beta, Alpha, Delta, and Gamma, respectively.

Even though stringent lockdowns, quarantines and travel restrictions were implemented in Wuhan, China, beginning January 23, 2020, and then extended to other cities in mainland China, the network-based characteristics of global air travel facilitated viral spread. Countries with hub airports in the global air transport network have greater connectivity and higher intensity (i.e., number of flights and seats) of

---

1 The ratio values based on flight capacity are similar to the results using seat capacity as the basis for calculating effective distance.

2 We use seat shares to calculate effective distance in the remainder of the paper.

3 In January 2020, there were 71 countries that had direct flights from China, 146 countries connected with China via one-stop connection, and six countries via two-stop connections, according to the flight schedule data.

4 There were 6 country destinations that needed more than a one-stop connection from China, including Saint Martin, Wallis and Futuna Islands, Tuvalu, Liberia, Montserrat, and Falkland Islands, with effective distances ranging from 14.16 to 17.49.
international traffic, and thus higher risks for importation of infected cases. The speed of the global spread of COVID-19 is striking. By March 11, 2020, when the WHO (2020) declared the outbreak of COVID-19 as a pandemic, the virus already had spread to 114 countries (by announced cases).

In the model for COVID-19, we exclude 13 countries that first reported a confirmed case later than 120 days after January 1, 2020, assuming these countries did not initially have timely and reliable reporting systems. We also rescale days to confirmed case, our dependent variable, to range from 0 to 100, to provide comparisons in

6 The 13 countries are Palau, Cook Islands, Kiribati, Micronesia, Saint Barthélemy, Western Samoa, Vanuatu, Marshall Islands, Solomon Islands, Curacao, Lesotho, Comoros, and Tajikistan.
transmission time between COVID-19 and its variants. Table 3 provides descriptive statistics for the variables in our model. Table 4 presents correlations among variables. Summary statistics and correlations for the variant datasets are available upon request.

4. Empirical model

We explore the determinants of the arrival time of COVID-19 and its variants into countries around the world. Our model examines determinants that contribute to the viral arrival time using origin-destination country pairs, with the outbreak locations of COVID-19 and each of its variants as the origin country and all other countries as destinations. One of the challenges in estimating the effect of determinants related to air travel intensity is controlling for potential endogeneity. Importantly, our policy variable assessing a reduction in air travel capacity to stem viral transmission, may be determined by endogeneity. Importantly, our policy variable assessing a reduction in air travel capacity, may be determined by non-exogeneity of the policy variable, we adopt a two-stage regression approach.

In the first stage, we regress the percentage reduction in seat capacity on three instruments: the stringency index that measures the level of lockdowns in a country, the growth rate of confirmed virus cases in a country, and the cumulative number of confirmed viral cases in a country. The first stage regression is specified as follows:

$$\text{SeatReduction}_i = \alpha + \beta_1 \text{IDX}_i + \beta_2 \text{Growth}_i + \beta_3 \text{C}_i + \epsilon_i$$

The dependent variable, SeatReduction, measures the percentage reduction in the total number of inbound seats into country $j$ measured a month prior to when one of the variants (depending on the estimating model) was first detected, compared to the pre-pandemic level of inbound seats in country $j$ (December 2019, used as the base case month). IDX is the stringency index, accounting for country-specific lockdown measures that may be associated with seat reductions. Growth refers to the growth rate of confirmed viral cases, and $C$ denotes cumulative confirmed cases per capita (i.e., per million population) in country $j$. The instruments for the variants are measured one month prior to when the infection was first reported. For COVID-19, we use February 2020 data as a period to measure the instruments. The vector $\chi$ denotes the exogenous controls that are included in the second-stage regression. $\epsilon_j$ is a robust error term clustered at a regional level.

In the second stage regression, we use the fitted values for seat reductions and effective distance as independent variables to determine the detection times of COVID-19 and its variants. The equation is specified as:

$$\text{Days}_j = \alpha + \beta_1 d_{ij} + \beta_2 \text{SeatReduction}_i + \beta_3 \text{WEF}_i + \epsilon_j$$

where Days$_j$ our dependent variable, denotes the number of days to the first confirmed case of COVID-19 (or one of the variants) reported by a destination country $j$, counted from January 1, 2020. Direct connectivity, $d_{ij}$, a key explanatory variable in our model, is measured using effective distance (based on seat shares) from the initial outbreak location, country $i$, to the destination country, $j$.

Our model also includes a measure for the health infrastructure of a country since countries with poor infrastructures may take longer to report viral cases. WEF refers to a country’s health index, as reported by the World Economic Forum (WEF, 2019). For countries without WEF health infrastructure scores, we use life expectancy obtained from the World Bank (2019), as this measure is highly correlated at 0.93, with WEF, $\epsilon_j$ is an error component clustered at a regional level.

To strengthen our empirical strategy, following Stock and Yogo (2005), we use the Cragg-Donald Wald F-statistic to jointly test for weakness of the instruments in the presence of endogenous variables. The F-statistics of the Cragg-Donald Wald test for the estimations of COVID-19 and its variants are greater than 10 (except Alpha and Gamma), a conventional cut-off level, supporting the finding that our instruments are generally strong. In addition, the Cragg-Donald Wald F-statistic of Alpha is 9.11, close to 10, a generally accepted threshold. We select our instruments based on the critical values using test statistics for the COVID-19 estimation. In addition, we conduct a test for over-identification restrictions. The $p$-values of the Hansen’s J-test across all our estimations are greater than 0.10 showing that our instruments are appropriately independent from the error process, supporting the validity of our instruments for the estimations.

5. Estimation results and discussions

We report our two-stage regression results in Table 5 and Table 6 using the origin-destination pairs for COVID-19 and the variants. The pairs include routes with China as the origin country for COVID-19; South Africa for Beta; the U.K. for Alpha; India for Delta; Brazil and Japan for Gamma.

The first stage equation findings in Table 5 show that the stringency index is significant in predicting seat capacity reductions in response to COVID-19 and its variant detection (except for Beta) with greater stringency levels associated with greater capacity reductions, as might be expected. The growth rate in confirmed cases is a significant variable at predicting COVID-19 and Alpha capacity reductions, with a positive coefficient indicating that higher growth rates are associated with greater reductions in capacity. The marginally significant coefficient in the COVID-19 regression has a negative sign, the opposite to what might be expected. However, in the same regression, the coefficient for the confirmed case rate per million population is a positive and highly significant predictor of seat capacity reductions.

Table 6 reports the results of our second stage equation. We use the fitted values from the reported regressions in Table 5 for our policy variable, seat reductions from pre-covid levels. Our results show that network connectivity, as measured by effective distance, significantly predicts transmission time for COVID-19 and the most of the variants (except Alpha). The results indicate that the greater the effective distance (i.e., the lower the intensity of flights) to a country, the greater the elapsed transmission and reporting time for the virus. This finding adds empirical support to the optimization framework developed by Ding et al. (2021), in which effective distance paths are modeled as key simulators for determining the epidemic situations in countries.

The fitted value for our policy variable, the percent reduction in inbound seat capacity compared to pre-covid levels, is insignificant across estimations. The insignificant results for predicting transmission time for COVID-19 and its variants suggest that the policy responses were largely ineffective; that is, the reduction in inbound seat capacities were likely cases of “too little and too late”. This finding is consistent with Sun et al. (2021a). It is notable that effective distance is generally more significant than the policy variable at predicting the timing of the
COVID-19 and Gamma. The implication is that countries with better mission time than policy moves to reduce inbound seat capacities after - are reported in parentheses. *p
Note: p-values based on the robust standard errors clustered at the regional level
Second-stage regression for detection times of COVID-19 and the variants.
Table 6
Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p
First-stage regression for percent seat reduction as dependent variable.
Table 5
Correlation matrix for the variables in the COVID-19 model.
Table 4

| (1) Days to first confirmed case | (2) (3) (4) (5) (6) (7) (8) (9) (10) |
|-------------------------------|-----------------------------------|
| 1.00                          | 1.00                              |
| (2) Days to first confirmed case (rescaled) | 1.00                              |
| 0.67                          | 0.67                              |
| (4) Effective distance (passenger numbers from China) | 0.71                              |
| 0.66                          | 0.66                              |
| (6) Seat reduction (%), adjusted | 0.14                              |
| (7) Stringency index | -0.31                             |
| (8) Growth rate of confirmed cases | -0.41                             |
| (9) Cumulative confirmed cases per million capita | -0.22                             |
| (10) WEF health index | 0.43                              |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p

| (1) Stringency index | (2) Growth rate of confirmed cases | (3) Cumulative confirmed cases per million capita | (4) Effective distance (seat capacity from Wuhan) | (5) Effective distance (passenger numbers from China) |
|----------------------|----------------------------------|-----------------------------------------------|-----------------------------------------------|-------------------------------------------------|
| 0.121**              | -1.537*                          | 4.077***                                      | -0.66                                         |
| (3.309)              | (2.342)                          | (11.383)                                     | 0.66                                          |
| 0.00417**            | 2.965                            | (4.599)                                      | 0.67                                          |
| (0.090)              | (5.007)                          | (1.090)                                      | 0.67                                          |
| 0.0406               | 0.0406                           | 0.0406                                       | 0.67                                          |
| (4.716)              | (4.716)                          | (4.716)                                      | 0.67                                          |
| 0.0406               | 0.0406                           | 0.0406                                       | 0.67                                          |
| (4.716)              | (4.716)                          | (4.716)                                      | 0.67                                          |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p

| (1) Route-based seat reduction (%) | (2) WEF health index | (3) Effective distance | (4) Effective distance (passenger numbers from China) | (5) Cumulative confirmed cases per million capita |
|----------------------------------|----------------------|------------------------|---------------------------------------------------|-----------------------------------------------|
| -0.166                           | -0.797               | -0.707**               | -0.109**                                          |
| (0.0179)                         | (1.090)              | (1.090)                | (1.090)                                           |
| 0.0289                           | 0.147                | 0.147                  | 0.147                                             |
| (0.090)                          | (5.007)              | (5.007)                | (5.007)                                           |
| 0.0398                           | 7.373**              | 7.373**                | 7.373**                                           |
| (0.3393)                         | (1.099)              | (1.099)                | (1.099)                                           |
| 0.0386                           | 1.854                | 1.854                  | 1.854                                             |
| (0.1507)                         | (0.090)              | (0.090)                | (0.090)                                           |
| 0.0386                           | -1.247               | -1.247                 | -1.247                                            |
| (0.1507)                         | (0.090)              | (0.090)                | (0.090)                                           |
| 0.0386                           | 4.829                | 4.829                  | 4.829                                             |
| (0.1507)                         | (0.090)              | (0.090)                | (0.090)                                           |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p

| (1) Constant | (2) R-squared | (3) Sample Size | (4) WEF health index | (5) Stringency index |
|-------------|--------------|----------------|----------------------|----------------------|
| 15.087**    | 0.333        | 165            | 0.121**              |
| (3.221)     | 165          | 0.121**        | (3.309)              |
| 94.002***   | 0.0917       | 161            | 0.121**              |
| (7.082)     | 161          | 0.121**        | (2.342)              |
| -117.922**  | 0.204        | 162            | 0.121**              |
| (-4.562)    | 162          | 0.121**        | (-0.342)             |
| 35.39       | 0.0729       | 332            | 0.121**              |
| (3.025)     | 332          | 0.121**        | (0.360)              |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p

| (1) Effective distance | (2) Route-based seat reduction (%) | (3) WEF health index | (4) Constant | (5) R-squared |
|------------------------|-----------------------------------|----------------------|-------------|--------------|
| 5.178***               | -0.166                            | -20.199**            | 47.252***   |
| (0.0179)               | (0.0179)                          | (0.042)              | (0.003)     |
| 3.309                  | -0.0502                           | -18.978              | 76.866*     |
| (0.167)                | (0.0179)                          | (-0.273)             | (0.064)     |
| 2.478*                 | 0.0259                            | -6.630               | 17.802      |
| (0.077)                | (0.0179)                          | (-0.814)             | (0.768)     |
| 2.638***               | 0.398                             | -1.055               | 13.616      |
| (0.0398)               | (0.0179)                          | (-0.927)             | (0.709)     |
| 0.0385                 |                                    | -61.869***           | 137.750***   |
| (0.935)                |                                    | (0.000)              | (0.000)     |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p

| (1) Root MSE | (2) Sample Size | (3) Gragg Donald Wald F statistics | (4) Hansen J statistics (p-value) |
|-------------|----------------|-------------------------------|-------------------------------|
| 14.086      | 165            | 503.14                        | 0.219                        |
| 18.159      | 69             | 2.91                           | 0.234                        |
| 16.552      | 79             | 35.64                          | 0.537                        |
| 22.973      | 79             | 9.11                           | 0.191                        |
| 14.488      | 104            | 9.11                           | 0.206                        |

Note: p-values based on the robust standard errors clustered at the regional level are reported in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

first announced viral case in a country. This implies that direct mobility from the outbreak location before the virus is first detected (effective distance), is more likely to be a key determinant in predicting transmission time than policy moves to reduce inbound seat capacities after the virus is first detected (SeatReduction).

Our WEF health index measure, used as a proxy for the sophistication of the health care sector in a country, is significant in the prediction of COVID-19 and Gamma. The implication is that countries with better health infrastructures tend to report their first viral case faster than those with lower levels of health infrastructure. However, differences in the level of health infrastructure did not significantly impact the timing of the discovery of first viral cases for other variants. Therefore, an infrastructure advantage in reporting viral cases is not evident in reducing time to viral discovery.

In general, the fit for our model is best for the prediction of COVID-19. These results suggest that variables other than those included in our model; that is, variables other than effective distance, the reduction in air capacity levels, and the level of a country’s health infrastructure, became more critical at determining viral speed and detection times as the virus progressed from initial discovery into its variants.

For robustness checks, we present additional estimation results. For our base case estimations, we calculated effective distance using the intensity of seat capacity on routes from viral origin to other countries. It is noteworthy that seat capacity may not reflect actual passenger traffic, especially given changing load factors once covid was discovered. Therefore, we re-estimate the viral transmission time for COVID-19 using measures for effective distance based on actual passenger data collected from SABRE. We calculate effective distance in two ways - using passenger numbers based on booked tickets on outbound flights from China to destination countries and then using seat capacity data on outbound flights only from Wuhan, the city where COVID-19 was initially discovered.

The results of the robustness estimations are shown in Table 7 (stage 1 of the 2SLS estimations) and Table 8 (stage 2 of the 2SLS estimations). The Cragg-Donald Wald F statistics are greater than 10 for both the China-based calculations and the Wuhan-based calculations. The p-values for the Hansen’s J-test also support the validity of our selected
The level of health infrastructure, as measured by screening capacity at airports, and climate and local environmental conditions. Omitting these variables may reduce the power of our model, as noted in Christidis and Christodoulou (2020). Importantly, we find that travel restrictions, as measured by inbound seat capacity reductions, generally are not effective at increasing viral transmission times, supporting previous findings (Sun et al., 2021a). We further find that the level of medical infrastructure, as represented by the WEF healthcare infrastructure index, was a significant predictor of COVID-19 and Gamma’s arrival time but was insignificant at predicting the arrival of other variants. Finally, we find that our model provides poorer predictions for the arrival time of the variants than for COVID-19, suggesting that factors other than those included in our model became more important predictors of the speed of viral transmission as the pandemic progressed.

The findings from our paper contribute to the growing literature on viral transmissions that can offer insights into designing public policies with respect to travel restrictions during future pandemics. The intensity of seats between the country where a virus first breaks out and destination countries (i.e., effective distance), measured prior to the viral outbreak, is generally associated with shorter viral transmission times, suggesting that air transport network connectivity is an important determinant of the timing of viral discovery in a country. However, policy efforts taken once the virus is first discovered to reduce air network connectivity seem to be ineffective; the measures are likely too little or too late. Therefore, governments may look to other ways to reduce the impact of the virus; for example, internal measures related to social distancing.

There are a number of caveats and limitations to our analysis. Importantly, we do not consider other measures that could impact viral transmission time and discovery. In addition to considering epidemiological factors, such as viral incubation time, reproduction rates, and pathogen transmission media, future models could include variables measuring socio-economic and demographic characteristics of travelers, screening capacity at airports, and climate and local environmental conditions. Omitting these variables may reduce the power of our models since these variables could be important at predicting the diffusion of COVID-19 and its variants, as suggested in Christidis and Christodoulou (2020).

### Author contributions

Conceptualization: LZ, MD 25 Methodology: YC, LZ, MD. Investigation: YC, LZ, MD. Visualization: YC, LZ Project administration: YC, LZ. Supervision: LZ, MD 30. Writing – original draft: YC, LZ, MD. Writing – review & editing: LZ, MD.

### Declaration of competing interest

Authors declare that they have no competing interests.

### Table 7

First-stage regression for percent seat reduction as dependent variable (for robustness).

|                        | (A)                      | (B)                      |
|------------------------|--------------------------|--------------------------|
| Stringency index       | 0.155** (4.54)           | 0.126** (2.81)           |
| Growth rate of confirmed cases | −1.543*** (−6.29)       | −1.583 (1.92)            |
| Cumulative confirmed cases per million capita | 4.429*** (38.75)         | 4.512*** (9.46)          |
| Effective distance     | −1.008* (−2.39)          | −0.754** (−2.96)         |
| WEF health index       | −8.505 (−1.12)           | −1.847 (−1.15)           |
| Constant               | 31.05 (1.99)             | 20.72*** (5.15)          |
| R-squared              | 0.583                    | 0.335                    |
| Sample Size            | 66                       | 165                      |

Note: t-statistics, based on the robust standard errors clustered at the regional level are reported in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

### Table 8

Second-stage regression (for robustness).

|                        | (A)                      | (B)                      |
|------------------------|--------------------------|--------------------------|
| Effective distance     | 9.229*** (13.52)         | 5.491*** (22.46)         |
| Route-based seat reduction (%) | 0.128 (2.58)            | −0.067 (0.77)            |
| WEF health index       | −62.36*** (−3.93)        | −13.28* (−1.15)          |
| Constant               | 95.18** (2.93)           | 6.707 (0.38)             |
| Pseudo R-squared       | 0.536                    | 0.564                    |
| Root MSE               | 15.156                   | 14.087                   |
| Sample Size            | 66                       | 165                      |
| Cragg-Donald Wald F statistics | 544.3 (170.9)        | 0.324 (0.168)           |

Note: z-statistics based on the robust standard errors clustered at the regional level are reported in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01.

In summary, our results suggest that air traffic intensity from the country where a virus is first reported to destinations countries, measured prior to the viral outbreak, is generally a significant contributor to viral transmission and discovery time. Countries that are historically well connected to the location of viral origin tend to have shorter time to the first reported viral discovery. Efforts made to reduce capacity levels in response to the virus detection are largely ineffective at increasing the time to viral discovery, suggesting that these measures are likely too little or too late.

6. Conclusion, policy implications, limitations and future research

In response to the COVID-19 pandemic, travel restrictions have been adopted by countries to slow down or mitigate viral spread. To better understand the effectiveness of travel-related measures in containing or slowing viral transmission, we construct models to estimate the arrival time of COVID-19 and four of its variants in countries around the world. We find, not surprisingly, that the air travel intensity from the country where the initial virus outbreak occurred is a primary factor in determining the arrival and discovery time of a virus in the destination country, suggesting that historical air travel activity is a key to predicting viral transmission risk, as noted in Christidis and Christodoulou (2020). Importantly, we find that travel restrictions, as measured by inbound seat capacity reductions, generally are not effective at increasing viral transmission times, supporting previous findings (Sun et al., 2021a). We further find that the level of medical infrastructure, as represented by the WEF healthcare infrastructure index, was a significant predictor of COVID-19 and Gamma’s arrival time but was insignificant at predicting the arrival of other variants. Finally, we find that our model provides poorer predictions for the arrival time of the variants than for COVID-19, suggesting that factors other than those included in our model became more important predictors of the speed of viral transmission as the pandemic progressed.

The findings from our paper contribute to the growing literature on viral transmissions that can offer insights into designing public policies with respect to travel restrictions during future pandemics. The intensity of seats between the country where a virus first breaks out and destination countries (i.e., effective distance), measured prior to the viral outbreak, is generally associated with shorter viral transmission times, suggesting that air transport network connectivity is an important determinant of the timing of viral discovery in a country. However, policy efforts taken once the virus is first discovered to reduce air network connectivity seem to be ineffective; the measures are likely too little and too late. Therefore, governments may look to other ways to reduce the impact of the virus; for example, internal measures related to social distancing.

There are a number of caveats and limitations to our analysis. Importantly, we do not consider other measures that could impact viral transmission time and discovery. In addition to considering epidemiological factors, such as viral incubation time, reproduction rates, and pathogen transmission media, future models could include variables measuring socio-economic and demographic characteristics of travelers, screening capacity at airports, and climate and local environmental conditions. Omitting these variables may reduce the power of our models since these variables could be important at predicting the diffusion of COVID-19 and its variants, as suggested in Christidis and Christodoulou (2020).
Data availability

Data will be made available on request.

References

Bank, World. 2019. “Life Expectancy at Birth, Total (years).” The World Bank, Washington, D.C.. https://data.worldbank.org/indicator/SP.DYN.LE00.IN
Bogoch, I.I., Crepaz, M.I., Cetron, M.S., Brownstein, J.S., Pesik, N., Miniota, J., Tom, T., Hu, W., Nicolucci, A., Ahmed, S., Yoon, J.W., Berry, I., Hay, S.I., Anema, A., Tatem, A.J., MacFadden, D., German, M., Khan, K., 2015. Assessment of the potential for international dissemination of Ebola virus via commercial air travel during the 2014 west African outbreak. Lancet (London, England) 385 (9962), 29–35. https://doi.org/10.1016/S0140-6736(14)61828-6.
Bown, J.T., Larsee, C. 2006. Airline networks and the international diffusion of severe acute respiratory syndrome (SARS). Geogr. J. 172 (2), 130-144. https://doi.org/10.1111/j.1475-4959.2006.00196.x.
Brockmann, D., Helbing, D. 2013. The hidden geometry of complex, network-driven contagion phenomena. Science (New York, N.Y. https://doi.org/10.1126/science.1245290, 342(6164), 1337-1342.
Budd, L., Bell, M., Brown, T., 2009. Of plagues, planes and politics: controlling the global spread of infectious diseases by air. Polit. Geogr. 28 (7), 426–435.
Callaway, E., 2021. ‘The perfect storm: lax social distancing fuelled a coronavirus variant’s brazilian surge. Nature. https://www.nature.com/articles/s41586-021-01480-3. (Accessed 22 December 2021).
Chan, S.L., Harding, N., Zachreson, C., Cliff, O.M., Prokopenko, M., 2020. Modelling quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Christidis, P., Christodoulou, A., 2020. The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic. PLoS One 5 (9), e12763. https://doi.org/10.1371/journal.pone.0127633.
Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore Y, Vespignani, A., 2013. The hidden geometry of complex, network-driven randomness. Int. J. Environ. Res. Publ. Health 17 (10), 3356. https://doi.org/10.3390/ijerph17103356.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Christidis, P., Christodoulou, A., 2020. The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic. PLoS One 5 (9), e12763. https://doi.org/10.1371/journal.pone.0127633.
Christidis, P., Christodoulou, A., 2020. The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic. PLoS One 5 (9), e12763. https://doi.org/10.1371/journal.pone.0127633.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Technol. 129, 103218 . https://doi.org/10.1016/j.trc.2019.103218.
Choi, Y., Wandelt, S., Sun, X., 2021. TLQP: early-stage transportation lock-down and quarantine problem. Transport. Res. C Emerg. Techno