ASSESSING THE IMPACT OF AIR PASSENGER TRAFFIC IN AIRPORTS ON COVID-19 RATES IN EUROPEAN REGIONS

Michael Stanton-Geddes

Airports Council International – Europe

ABSTRACT

This study provides the result of statistical analysis of weekly airport passenger traffic data and the rate of new COVID-19 cases (COVID-19 incidence rate) in Europe at both the country and sub national level during 2020, controlling for the prior incidence rate, the level of stringency of government measures, and the mobility of people. This paper focuses on the relationship between air travel and the COVID-19 incidence rate during the second half of the year because this addresses the real problem faced in Europe about whether to permit air travel after a novel virus was already highly present in the community. The paper does not seek to add to the literature about the role of air travel in the initial propagation of a novel virus, nor does it consider the impact of vaccine availability.

The analysis provides evidence that a 10% increase in airport passengers is correlated with a 0.14% increase in the COVID-19 incidence rate in Europe's subnational regions during the second half of 2020. As comparison, an increase of 10% in the index of the stringency of government measures is correlated with a 4.3% reduction in the COVID-19 rate, and 10% more mobility in the population with a 2.9% increase in the COVID-19 rate. The paper uses a model that is based in literature and applies ordinary least squares (OLS) regression techniques for fixed effects and pooled panel data.

The finding that increases in air transport traffic when a novel virus is already widely present suggests that efforts to restrict or control air travel are not likely to be efficient.

KEYWORDS

Airports and pandemics; Covid-19 and air transport; Airport passenger flows; Air transport statistics

1. INTRODUCTION

When COVID-19 is already present in the community and specifically subnational regions, to what extent does air travel contribute to the transmission of COVID-19?

This paper is an empirical study using recent health, traveller and population data to identify the relationship between air travellers and COVID-19 rates. The geographic scope is limited to countries using Nomenclature of Territorial Units for Statistics (NUTS) classifications at the first level, NUTS1. The data starts in week 15, which is when governments and agencies first were able to consistently report detection of COVID-19 cases, but the analysis focuses on the period after week 27, when COVID-19 was widely present in the community in Europe.

There is no doubt that air travel and especially long-haul inter-continental flights provided a pathway for the rapid global spread of COVID-19. Initial literature has already documented the role of aviation in spreading COVID-19 from where it originally emerged, as summarised in the literature review section of this paper. Some initial government measures were to limit flights from cities where the virus was detected, and by March 2020, one of the responses by European governments to the arrival of the COVID-19 pandemic was to place wide-spread limitations on air travel along with other measures to reduce person-to-person interaction, documented by EUROCONTROL's Network Operations Portal.

These restrictions, in the form of closure of airports to commercial air traffic, or prohibitions on travel for passengers without a recent negative test result or mandatory quarantines have persisted from the beginning of government responses in March 2020 well into 2021.

The timeframe studied in this paper is established to consider the relationship between air travellers in a region and COVID-19 rates when the virus is already present in the community. The period assess also ends at the end of 2020, before vaccines were administered in Europe. The first person in the world to receive a COVID-19 vaccine had the shot administered in the United Kingdom on 8 December (BBC, 2020. The European Medical Agency granted its first marketing authorisation for a COVID-19 vaccine on 20th December 2020, but the vaccination program only started reach the broad public at the start of the second quarter of 2021 (European Medicines Agency, 2020).

The geographical region focused on in this paper is the subnational level, because this analysis the data at the appropriate level of detail. During the COVID-19 lockdowns, with air services greatly reduced, people would travel from further away to the larger regional and hub airports which maintained flights. NUTS1 regions are socio-economic regions with population between 3 million and 7 million, usually covering a drive-time of a number of hours, and within those regions there may be multiple airports. The use of more detailed geographical levels would risk to distort the analysis by overly reducing the geographical areas to small regions, and potentially missing incidences of COVID-19 cases in adjoining geographic regions to the region where the airport is located.

An analysis at the country level was performed. This high level of geographic aggregation may not be appropriate for larger countries where COVID-19 rates varied significantly across regions. However, as European countries during the COVID-19 lock-downs made foreign travel and crossing borders difficult, the country-level of analysis is appropriate because it would isolate the impacts of cross-border travel.

2. LITERATURE REVIEW

COVID-19 has had a colossal impact on people, business and society. This paper complements studies that have identified the risk of transmission based on air circulation within aircraft cabins, and theoretical models for transmissions based on population data, by considering population-wide information and real data about factors in the community. Lau et al (2020) published one of the first papers about COVID-19 and travel, in March 2020, tracking the association between domestic and international air travel and the COVID-19 outbreak. They used data from Official Airline Guide (OAG) of flight schedules and the Civil Aviation Administration of China (CAAC). Moreover, Lau et al did not use actual traffic data, as it was not available. At the same time, Lau et all imputed traffic flows between Chinese regions and international routes based on historical travel flows. They performed a correlation analysis of confirmed COVID-19 cases with passenger flows. They found that domestic and international air travel was highly correlated with the growth in cases (Lau, 2020).

In October 2020, a meta-analysis of transmission events during flights by Freedman and Wilder-Smith (2020) pointed to little evidence that COVID-19 is readily transmitted from asymptomatic but infected passengers during flight. The meta-analysis considered factors such as if the passengers reported wearing masks, the cabin of travel, if the infected person was a passenger or crew, and if meals were served. The International Air Transport Association (IATA, 2020) published an analysis of data provided by airlines which also validated the argument that while it is likely there have been travellers with COVID-19, infection rates during air travel seem low.

Murphy et al (2020) published a micro-study of passengers on a 7-hour flight arriving in Ireland. Using detailed information about the passengers, the study found that 13 of 49 passengers on the flight had COVID-19 or tested positive shortly after the flight, and that these passengers led to a total of 59 infections. Some but not all of the passengers wore masks during the flight.

The European Centre for Disease Control published Guidelines in December 2020 based on the latest scientific evidence stating that travellers should not be considered as a high-risk population, nor treated as contacts of COVID-19 cases, unless they have been in known contact with a confirmed positive case. ECDC's review of scientific evidence and information concluded that the prevalence of COVID-19 among travellers is estimated to be lower than is the case for the general population.

Pana et al (2021) in a study that assessed country-level determinants of the severity of the first wave of the COVID-19 pandemic found that of the country-level parameters assessed, international travel was the main determinant of the severity of the first global wave of the COVID-19 pandemic. The paper suggested that international travel restrictions applied very early in the pandemic course should be considered to avoid rapidly increasing infection and death rates globally. The variable used by Pana et al for international travel was not the actual number, but a measure on international connectedness based on historical travel patterns. According to the paper, "one multivariable adjustment, international arrivals in 2018, as a marker of global connection, was significantly associated with an increase in the log mean mortality rate (0.033 (95% CI 0.012 to 0.054) per 1 million increase in international arrivals, p=0.003). This translates to an exp(B) of 1.034, equivalent to a 3.4% increase in the mean mortality rate for every 1 million." The paper does not consider time dependence and does not for instance use a lagged dependent variable to control for endogeneity.

Looking beyond air travel, other studies have also assessed impact of people's travel on the spread and increase of COVID-19 case rates.

Glaeser et al (2020) asked how does COVID-19 increase with mobility, and while finding significant heterogeneity between United States' east coast large cities, they found in general that during the initial stages of the pandemic, total cases per capita decrease 20% for every 20% decline in mobility. Glaeser used public transit passenger count data, cellular phone data to measure mobility, and state public health data, as well as descriptive data about the individuals within sub-regions based on employment and profession surveys, to measure reduced travel to workplaces by categories of people able to telework.

3. METHODOLOGY

The model identifies the relationship between air passengers and the COVID-19 test positivity rate in a country, during the second half of 2020 in Europe.

The model is inspired by the study of Glaeser et al cited in the literature review.

Model specification:

$$\ln(Rate\ of\ new\ COVID\text{–}19\ Cases)_{it} = B_1(\ln(Pax_{it-1})) + B_2\ (\ln\ (COVID\text{–}19\ Rate_{it-1})) + B_3(Stringency_{it-1}) + B_4\big(Google\ Mobility_{it-1)}\big) + \ \varepsilon_{it}$$

Where subscript ***i*** represents the geography and subscript ***t*** consists of ISO weeks in 2020, in the core model the weeks 27 to 53.

In the national-level model, ***i*** represents data from 35 European countries: Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and the United Kingdom.

In the subnational-level model, *i* represents the NUTS1 regions listed in the Annex. NUTS1 region data is only available for the EU27 member states, via the ECDC, and for the United Kingdom's regions, via the UK statistical agency.

In descriptive terms, the model assessed changes in the COVID-19 test positivity rate based on changes in air passengers in a *geographic region (country or NUTS1)* during a *certain week*, controlling for the COVID-19 test positivity rate in the prior week, the stringency of measures and changes in the movement of people.

The COVID-19 rates and all co-variates are transformed by taking the natural logarithm to normalise the data, ensuring that the data is linear in its parameters and appropriate for OLS regression, as shown in Figure 3 Residuals Plot for Subnational Panel Model **Σφάλμα! Δεν έχει οριστεί σελιδοδεικτης.**. The log-log specification allows interpretation based on percentage changes over the explanatory variables leading to percentage change of the response variable.

4. DATA

The dataset is a weekly panel for Countries and NUTS1 regions in Europe for each week of 2020 from week 15, when COVID-19 cases data is available through week 53, effectively the last 3 quarters of 2020.

The data was log-transformed, both the dependent and independent variables, to normalise the observations and is suitable for ordinary least squares (OLS) analysis based on the residuals presented in the analysis below.

The central time-period in focus is on the second half of 2020, to provide information about the role of air travel in the transmission of COVID-19 at a time when COVID-19 is already present in the community. The data is discussed in depth below. The dataset includes the sum of weekly arriving and departing passengers for 35 European countries, and 89 NUTS1 subnational regions in a panel data, and national and subnational rates of COVID-19 new cases per 100k population average over 14 days. All the data is averaged over International Organization for Standardization (ISO) weeks ("isoweek").

4.1 Response Variable: Covid-19 Incidence Rate

The response variable that this analysis measures is the incidence rate of new COVID-19 cases. The European Centre for Disease Prevention and Control has provided weekly updates on the 14-day notification rate of newly reported COVID-19 cases per 100 000 population. This is a core public health measure of the presence of COVID-19 in the

community (ECDC, 2021). The UK data is provided for a 7-day rolling average notification rate (Public Health England, 2021).
The data is available at the national level and subnational level. The subnational data required alignment with the taxonomy of NUTS1 regions (e.g., the data provided by Portugal uses a slightly different classification than the standard NUTS classifications) and also cleaning, for example there were 116 entries with a negative 14 day rate. These are clearly data entry errors and had to be removed.

The data was averaged for each isoweek.

4.2 Explanatory Variables

4.2.1 Airport passengers in the prior week

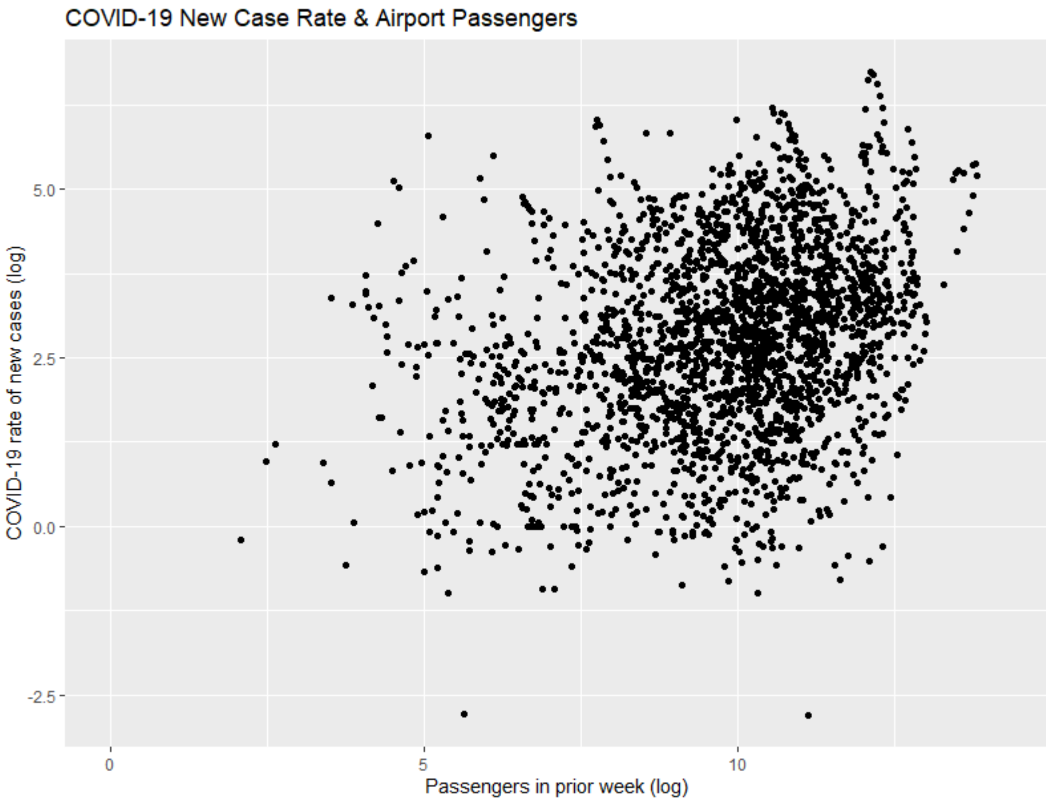
One challenge was obtaining airport level passenger data on a weekly basis. OAG's Traffic Database provides data aggregated at a monthly level. Data is not publicly available at a more detailed, weekly or daily, level. In normal times, the distribution of traffic throughout a month is fairly even, with some changes for holiday weeks or major events. During COVID, travel patterns were disrupted as passengers chose not to travel, and/or governments restricted travel, and airlines responded by limiting services.

Therefore, the total monthly data was appropriately allocated to each day of the month. This was done by using data published by ACI EUROPE (2021) regarding daily and weekly percent changes in airport traffic compared to 2019. The data provided by country served as an allocation key for the monthly passenger figures obtained from OAG, assuming relative homogeneity in the passenger traffic performance of each airport in a country. For countries where ACI EUROPE did not provide a country-level report, the regional level daily percent changes were applied to the country. The assumption of similar performance between airports can be seen as valid because airport traffic during the early months of the COVID-19 pandemic in Europe was resulting from government policy and measures, rather than differences in airport performance.

Each airport was identified by the country and NUTS1 region in which it is located, and the volume of passengers to/from the airport summed at both the national and sub-national level.

The relationship between primary explanatory variable of interest, passenger numbers, and the response variable, the COVID-19 rate is shown on Figure 1 National level; COVID-19 rate and air travellers in the prior week. It indicates that air traveller numbers have little to no impact on COVID-19 incidence rates. Examining the simple plotted relationship of COVID-19 incidence rate by subnational region and the number of air travellers to/from the region in the prior week offers an illustration of the general lack of a direct relationship.

Figure 1 National level; COVID-19 rate and air travellers in the prior week



4.2.2 Rate of COVID-19 new cases in the prior week

Naturally, the number of air travellers arriving to and departing from a region can only be one variable impacting the overall rate of new COVID-19 cases. In line with knowledge about virus transmission and epidemiology, an important factor will be the number of people who carry the COVID-19 virus, are contagious, and are in contact with other people. So, a lag of the response variable is included in the model as one of the explanatory variables (ECDC, 2021a).

Using the response variable as a lagged-predictor variable, may have issues of autocorrelation, at first reflection. The analysis of the data indicates that the autocorrelation is not a problem. While there is a link from COVID-19 rates in one week to another, it is not the same as the serial correlations seen in economic data, for example GDP. While GDP has a stickiness or inertia to remain at its current level across time periods, COVID-19 is a virus that left uncontrolled should grow exponentially, or fully controlled should decay to zero. In other words,, in a region with zero cases of COVID-19 and zero travel in or out of the reason, it is nearly certain that the following week the number of new cases will be zero. Alternatively, in a region with some COVID-19 cases, but zero person-to-person interaction and zero mobility resulting in COVID-19 virus aerosol droplets to be transmitted, the number of new cases in the following week would also be zero. The auto-correlation tables presented in *Figure 3 Residuals Plot for Subnational Panel Model, two-way fixed effects, weeks 27 to 53* show that the residuals in the model are not causes for concern.

4.2.3 Stringency of public measures

The level of the "stringency" of restrictions put in place by governments to control the transmission of the virus will also impact the COVID-19 rate. The Oxford COVID-19 Government Response Tracker is used to measure the strength of government measures (Oxford COVID-19 Government Response Tracker, 2021). This variable is a key control in the model, to account for reasons that travel and person-to-person interaction may be limited by public measures.

The Oxford COVID-19 Government Response Tracker dataset consists of 19 indicators, grouped by containment and closure policies, economic policies, health system policies and miscellaneous policies. The *Oxford COVID-19 Government Response Tracker* data is described on the University of Oxford Blavatnik School of Government site.

As the aim of this analysis is to consider the impact of air travellers on the incidence rate, the most useful variable is the overall "Stringency Index" figure for each country. It should be noted that there is one indicator [C.8] which specifically measures international travel controls. The overall Stringency Index is a simply average of the individual component indicator, described in detail on the project's webpage.

The data is provided on a daily level, and for this analysis it was aggregated to the mean value for each country by ISO week.

4.2.4 Community mobility

Google has provided Community Mobility data, which records changes against a baseline over time of people in retail & recreation areas, grocery & pharmacy, transit, parks, residential places and workplaces (Google, 2021). The Google Mobility data is described in depth on the Google Community Mobility Reports site.

This variable is an important part of the model. The World Health Organisation has studied the phenomenon of "pandemic fatigue" and reported that compliance with public rules seems to diminish over time (World Health Organisation, 2020). Therefore, the community mobility measure is important to capture actual levels of compliance of people with rules to limit interaction and reduce possibility for encounters during which COVID-19 spread.

The six Googly Mobility indices are all highly correlated. Therefore, in the analysis, only one of the variables was selected, to avoid specification problems in the model. The 'Retail & Recreation' index was selected. This index had the least gaps in data, it had low scores for multi-collinearity, and it had fewer outliers. For example, the Parks Index notably demonstrated large swings, presumably when lockdowns were lifted and coincided with good weather or weekends. The 'Retail & Recreation' index shows clear trends over weeks and days and does not have outliers which would reduce the effectiveness of the regression technique. To ensure validity of the results, regressions were performed with other indices and the conclusions remain valid.

Google provides this data at the national level and also sub-national level. While the subnational data is not coded using NUTS identifiers, the use of ISO 3166 and other sub-region identifiers allow for accurate and comprehensive correlation between the Google Mobility data subregion, and EU NUTS regions. This allowed the dataset to be completed at the subnational level with Community Mobility data for each NUTS1 region.

The data is provided on a daily level, and for this analysis it was aggregated to the mean value for each country by ISO week.

5. RESULTS AND DISCUSSION

As stated in the introduction, this paper is concerned about what the impact of air travel is when the virus is already widely present, so it focused on the second half of 2020 – weeks 27 to 53. In the annex *Regressions for additional time periods*, regression results are presented for additional time periods.

An ordinary least squares (OLS) regression was performed with fixed effects controlling for geography and week of the year for the data described above, and also for pooled data, at both the subnational and national level.

5.1 Subnational-level analysis

The ordinary least squares (OLS) regression analysis using fixed effects for region and week finds that during the second half of 2020 in Europe, the coefficient for airport passengers in Table 1 Impact of Air Passengers on COVID-19 rates during H2 2020 in European NUTS1 Regions is 0.014.

Since both the regressor and explanatory variables have been transformed by logs, the interpretation is that a 10% increase in the volume of airport passengers is related with a 0.14% increase in the COVID-19 rate per 100,000 population.

This finding between air passengers and COVID-19 new case rates is statistically significant at the $p<0.01$ level (p-value 0.009).

Other population and mobility variables have stronger influences on the COVID-19 incidence rate, as further discussed.

Table 1 Impact of Air Passengers on COVID-19 rates during H2 2020 in European NUTS1 Regions

Subnational: Relationship between Air Passengers & COVID-19 rates during H2 2020

	Dependent Variable: Rate of new COVID-19 cases per 100k population (log)	
	Fixed Effects H2	Pooled H2
Air Passengers in prior week (log)	0.014	0.015
	$p = 0.009^{**}$	$p = 0.00004^{***}$
COVID-19 rate in prior week (log)	1.003	0.985
	$p = 0.000^{***}$	$p = 0.000^{***}$
Stringency Index in prior week (log)	-0.431	-0.270
	$p = 0.000^{***}$	$p = 0.000^{***}$
Community Mobility in prior week (log)	0.294	0.211
	$p = 0.000^{***}$	$p = 0.000^{***}$
Constant		0.210
		p = 0.367
Observations	2,367	2,367
R^2	0.968	0.971
Adjusted R^2	0.967	0.971
F Statistic	$17{,}186.480^{***}$ (df = 4; 2252)	$20{,}049.160^{***}$ (df = 4; 2362)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

Conceptually, differences between regions and over time may be important. However, panel data is usually assessed over many years and not a 3-month period as in this study. Additionally, the geography is more homogeneous, and European Union member states coordinated measures via the European Commission, especially regarding travel to and from regions outside of the Schengen Zone. For this reason, an analysis was also conducted of the 'pooled' data.

The second column labelled *Pooled* assumes more homogeneity amongst regions and during the second half of 2020. The regression coefficient of 0.015 implies that the COVID-19 incidence rate would change by 0.15% for a 10% change in the number of air travellers. This has a stronger significance at a p-value of 0.00004.

The other variables report values as they would be expected to. The COVID-19 rate in the prior week is significant and has a high coefficient. As mentioned in the section *4.2.2 Rate of COVID-19 new cases in the prior* **week** the issue of autocorrelation of the lagged response variable must be examined. The residuals plot for the subnational panel below reveals that the variations and auto-correlation across time does not pose serious concerns for the validity of the model assumptions. *Figure 3 Residuals Plot for Subnational Panel Model* show that the data is stationary and residuals normally distributed.

The Stringency Index has a negative coefficient, meaning that as the Stringency Index increases, the COVID-19 rate in the subsequent week decreases, showing that *de jure* government stringency measures do impact the COVID-19 rate. This is expected, as it can be presumed that enough of the population will follow government rules, and limited person-to-person contact will reduce the rate, with a very strong negative relationship (as stringency levels increase, COVID-19 rates go down sharply).

The *de facto* measure of community mobility collected by Google Mobility mobile phone data shows that as mobility increases, the COVID-19 rate also increases. Again this coefficient (0.294) is in the right direction, and indicates a far stronger relationship.

Overall, the regression results provide the finding of statistically significant relationship between airport travellers and the COVID-19 rate, and also the other regressors. However, the impact of changes in number of air travellers is a fraction of the other variables.

5.2 National-level analysis

The national level data in *Table 2* shows that the variable for air travellers in the prior week does not have a statistically significant relationship with the weekly average of 14-day notification rate of newly reported COVID-19 cases per 100,000 population.

The OLS regressions for national level data lose statistical significance, and the coefficients also decease in magnitude. The general direction and relative influences of the variables remains constant.

Despite the weakness of the coefficient, it should be noted the level – 0.015 for the fixed effects regression and 0.005 for the pooled regression – confirm the more detailed analysis performed on subnational NUTS1 detail data.

Table 2 National-level regression results

National: Relationship between Air Passengers & COVID-19 rates during 2020 | Fixed Effects & Pooled

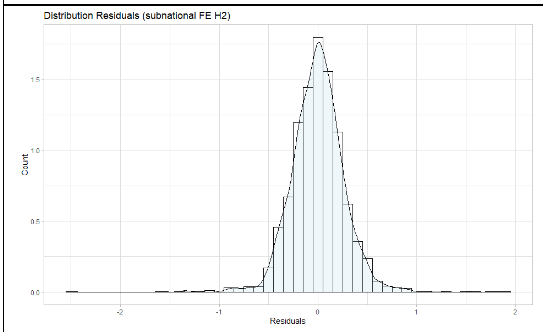
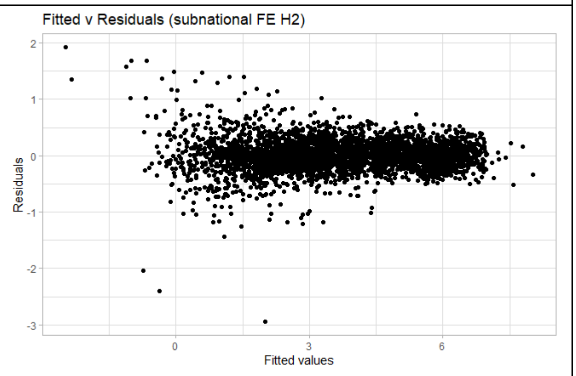
	Dependent Variable: Rate of new COVID-19 cases per 100k population (log)	
	Fixed Effects H2	Pooled H2
Air Passengers in prior week (log)	0.015	0.005
	p = 0.093	p = 0.390
COVID-19 rate in prior week (log)	1.015	0.992
	p = 0.000***	p = 0.000***
Stringency Index in prior week (log)	-0.380	-0.106
	p = 0.000***	p = 0.004**
Community Mobility in prior week (log)	0.302	0.287
	p = 0.00000***	p = 0.00000***
Constant		-0.735
		p = 0.028*
Observations	940	940
R^2	0.972	0.973
Adjusted R^2	0.970	0.973
F Statistic	7,549.417*** (df = 4; 867)	8,505.944*** (df = 4; 935)
Note:	*p<0.05; **p<0.01; ***p<0.001	

6. DIAGNOSTICS

The COVID-19 rates and Airport Passenger volumes are transformed by taking the natural logarithm to normalise the data, ensuring that the data is linear in its parameters and appropriate for OLS regression, as shown in *Figure 3 Residuals Plot for Subnational Panel Model* **Σφάλμα! Δεν έχει οριστεί σελιδοδείκτης.**. For consistency of interpretation, the other explanatory variables are also log transformed. The log-log specification allows interpretation based on percentage changes over the explanatory variables leading to percentage change of the response variable.

Panel effects are appropriate to apply. The F-test for individual (geography) and time effects rejected the null hypothesis. Supporting this test is the results of the Lagrange Multiplier Test for time effects that rejects the null hypothesis of no effects, so the fixed effect model controlling for time (weeks) is more appropriate.

Figure 2 Histogram and residual distribution

Histogram – Distribution of Residuals	Scatterplot of fitted and residual values
	

6.1 Endogeneity

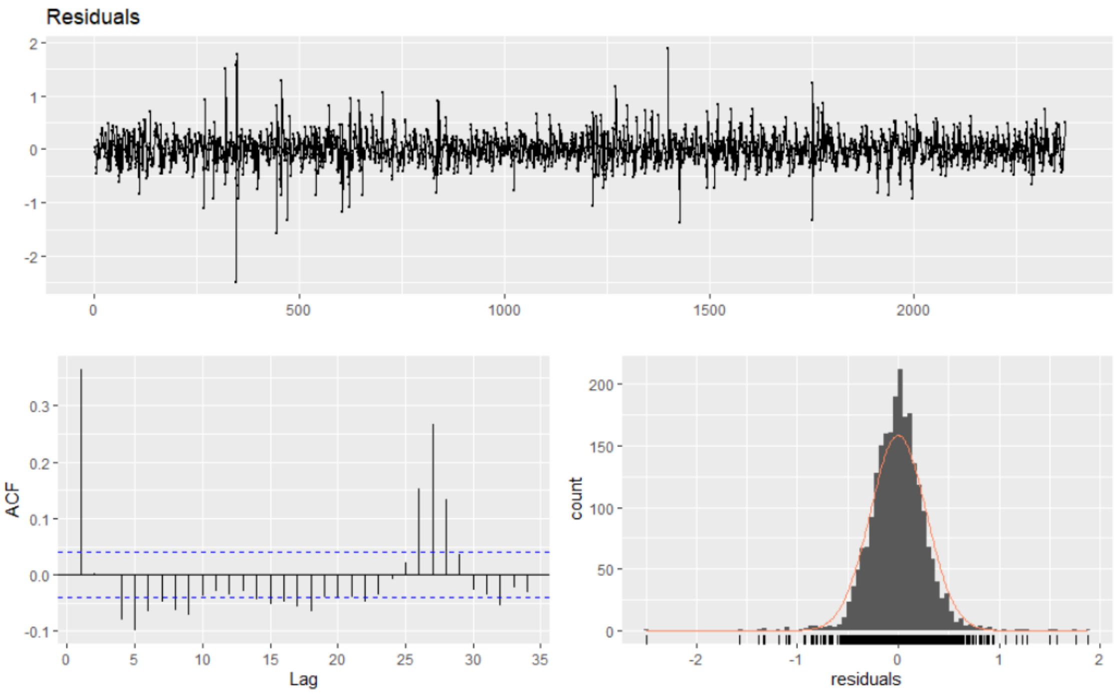
It was considered that the model may suffer from two issues; omitted variable bias and reverse causality.

Omitted variable bias

One question remained about the risk of a distortion to the regression coefficients because of an omitted variable. Specifically, the model only includes the number of airport passengers in the prior week. There is medical evidence that COVID-19 can present symptoms in some people more than one week after they were infected. This means that the COVID-19 rate in a certain week may be impacted the airport passengers two or more week previously.

However, adding additional lags of the airport passengers in the time periods two and three weeks before introduces a problem of multicollinearity. The weekly airport passengers in week -1 and airport passengers in week – 2 have a correlation rate of more than .90. In test regressions, the result of adding additional multiple lags highly distorted downwards the coefficients, even resulting in negative coefficients for the impact of airport passengers on the COVID-19 rate. This is clearly a spurious correlation, resulting from the multi-collinearity.

Figure 3 Residuals Plot for Subnational Panel Model, two-way fixed effects, weeks 27 to 53



Next, a first differences regression was calculated to control for possible bias of the lagged dependent variable, presented in Table 3 First Differences Regression. The regression results confirm that the most important covariate is the rate of COVID-19 in the prior week, which has a strongly positive relationship. The number of air passengers loses statistical significance, though the coefficient remains consistent with the pooled and fixed effects regressions. The stringency index covariate is not significant, and the covariate for community mobility is statistically significant at the 0.01 level with a positive coefficient.

Table 3 First Differences Regression

Subnational: Relationship between Air Passengers & COVID-19 rates during H2 2020

	Dependent Variable: Rate of new COVID-19 cases per 100k population (log)
	First Differences - Individual Effect
Air Passengers in prior week (log)	0.012
	p = 0.341
COVID-19 rate in prior week (log)	0.932
	p = 0.000***
Stringency Index in prior week (log)	-0.069
	p = 0.390
Community Mobility in prior week (log)	0.143
	p = 0.012*
Constant	0.003
	p = 0.695
Observations	2,295
R^2	0.857
Adjusted R^2	0.857
F Statistic	3,435.068*** (df = 4; 2290)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Lack of a meaningful instrumental variable

The incidence rate of COVID-19 cases in a country may be a strong determining factor for the number of air travellers arriving to or departing from that country or region. If the explanatory variable of airport passengers in the prior week is correlated with the error term, then the regression coefficients will be distorted.

To control for the problem of simultaneous causality, the use of an instrumental variable in the regression can correct for potential endogeneity. One potential variable which could be introduced as an instrument is the average air fare departing from airports in a country/region. The air fare data available in the OAG schedules analyser database was used to construct an instrument. The use of air fares as an instrument would be appropriate, as passenger demand is at any time elastic in response to the level of air fare; and it is clear as well that air fares meets the second criteria for a valid instrument of excludability, as air fares in and of themselves can have no impact on the COVID-19 incidence rate.

However, ultimately, the instrument variable method had to be discarded because of problems with the data. OAG air fare data is available only on a monthly basis, and the available data was too inconsistent to allow imputation of weekly levels. The use of an instrument that does not vary at the same period as the response variable would limit the interpretation of the results from an instrumental variable regression.

7. CONCLUSION

This paper has demonstrated that it is unlikely that the number of air travellers has a strong impact on the incidence rates of COVID-19 in countries in Europe, all other factors equal, during the second half of 2020. Statistical evidence indicates that a 10% change in the number of air passengers to/from a region is related to a change in the COVID-19 incidence rate of around 0.14%.

The paper relies on a large dataset, covering weekly airport arrivals and departures in the Europe region, during the second half of 2020, controlling for the stringency of government measures, community mobility, and ensuring direction of causality by using air ticket prices as instrument.

This finding re-affirms qualitative evidence and supports the scientific research on particle transmission in aircraft which finds that aircraft air recirculation rates, passenger seating, and air filters effectively limit the possibility for virus transmission, and theoretical models of virus transmission vectors by air travel.

Future research could expand the regions to look at countries outside of Europe and compare regional issues, especially where government measures were less strong. Additionally, the rapid vaccination of large portions of the population and the relaxation of travel restrictions, will offer the ability to compare differences in COVID-19 rates and understand the impact of air travel in the spread of a novel virus.

This paper does not say that air travel has no role in the transmission of COVID-19. It is abundantly clear, and research tracing passengers especially flying into strongly controlled countries with hotel-bound quarantine for all arriving passengers, has demonstrated that COVID-19 can be transmitted in aircraft. In countries and regions where closing the country to the rest of the world is not an option, this paper shows that measures targeting aviation alone are not productive.

8. REFERENCES

- ACI EUROPE (2021). COVID-19 Daily Airport Traffic Statistics. www.aci-europe.org.
- BBC News (8 December 2020). Covid-19 vaccine: First person receives Pfizer jab in UK. Available from: https://www.bbc.com/news/uk-55227325.
- David O Freedman, MD, Annelies Wilder-Smith, MD (2020), In-flight transmission of SARS-CoV-2: a review of the attack rates and available data on the efficacy of face masks, *Journal of Travel Medicine*, Volume 27, Issue 8, https://doi.org/10.1093/jtm/taaa178.
- Edward L. Glaeser, Caitlin Gorback, and Stephen J. Redding (2020); How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities; NBER Working Paper No. 27519; July 2020, JEL No. H12.
- European Centre for Disease Prevention and Control (2020). Guidelines for COVID-19 Testing and Quarantine of Air Travellers. Stockholm: ECDC. Available from:

https://www.easa.europa.eu/document-library/general-publications/guidelines-covid-19-testing-and-quarantine-air-travellers.

- European Centre for Disease Prevention and Control (2021). Data on 14-day notification rate of new COVID-19 cases and deaths & Data on testing for COVID-19 by week and country. Available from: https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19.
- European Centre for Disease Prevention and Control (2021a). Data on 14-day notification rate of new COVID-19 cases and deaths & Data on testing for COVID-19 by week and country (Accessed: 11 April 2020). Available from: https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19.
- European Medicines Agency (2020). EMA recommends first COVID-19 vaccine for authorisation in the EU. Available from: https://www.ema.europa.eu/en/news/ema-recommends-first-covid-19-vaccine-authorisation-eu.
- Google LLC (2020). Google COVID-19 Community Mobility Reports. https://www.google.com/covid19/mobility/ (Accessed: 11 April 2020)
- Hale, Thomas, Noam Angrist, Emily Cameron-Blake, Laura Hallas, Beatriz Kira, Saptarshi Majumdar, Anna Petherick, Toby Phillips, Helen Tatlow, Samuel Webster (2020). Oxford COVID-19 Government Response Tracker, Blavatnik School of Government. Available from: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
- IATA (2020). Extremely low risk of viral transmission inflight, Geneva, Switzerland. Available from: https://airlines.iata.org/analysis/extremely-low-risk-of-viral-transmission-inflight.
- Lau, et al (2020). The association between international and domestic air traffic and the coronavirus (COVID-19) outbreak. ScienceDirect. https://doi.org/10.1016/j.jmii.2020.03.026
- Murphy Nicola, Boland Máirín, Bambury Niamh, Fitzgerald Margaret, Comerford Liz, Dever Niamh, O'Sullivan Margaret B, Petty-Saphon Naomi, Kiernan Regina, Jensen Mette, O'Connor Lois (2020). A large national outbreak of COVID-19 linked to air

travel, Ireland, summer 2020. Eurosurveillance. 2020;25(42). Available from: https://doi.org/10.2807/1560-7917.ES.2020.25.42.2001624.

- Pana TA, Bhattacharya S, Gamble DT, et al (2021). Country-level determinants of the severity of the first global wave of the COVID-19 pandemic: an ecological study, BMJ Open; doi: 10.1136/bmjopen-2020-042034.
- Public Health England (2021). Coronavirus Data. Available from: https://coronavirus.data.gov.uk/.
- World Health Organisation (2020). "Pandemic fatigue - reinvigorating the public to prevent COVID-19: policy framework for supporting pandemic prevention and management: revised version". Available from: https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/publications-and-technical-guidance/2020/pandemic-fatigue-reinvigorating-the-public-to-prevent-covid-19,-september-2020-produced-by-whoeurope

9. ANNEXES

Data summary

	COVID-19 Incidence Rate	Weekly Air Passengers	Stringency Index	Retail Index
Minimum	0.061	0	21.56	7.143
1st quartile	6.293	3736	48.15	54.771
Median	16.710	18399	57.87	85.000
Mean	39.317	55658	58.81	77.174
3rd quartile	42.143	58706	69.91	97.910
Maximum	838.737	1523708	96.30	303.429

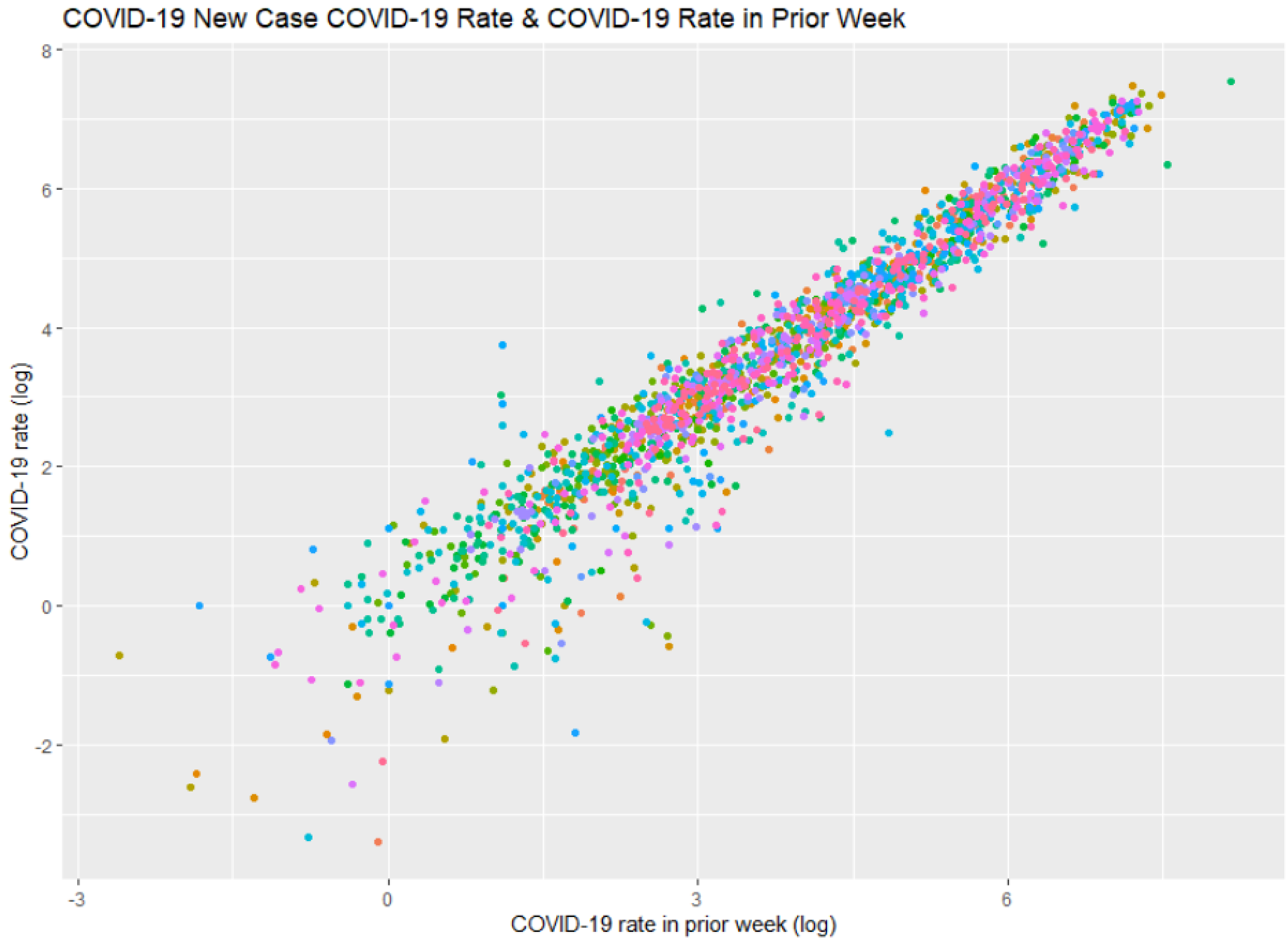
Countries in the dataset

Austria, Belarus, Belgium, Bosnia and Herzegovina Bulgaria Croatia, Czechia, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, and the United Kingdom

NUTS1 regions in the dataset

AT1, AT2, AT3, BE2, BE3, BG3, BG4, CY0, CZ0, DE1, DE2, DE3, DE4, DE5, DE6, DE7, DE8, DE9, DEA, DEB, DEC, DED, DEF, DEG, DK0, EL3, EL4, EL5, EL6, ES1, ES2, ES3, ES4, ES5, ES6, ES7, FI1, FI2, FR1, FRB, FRC, FRD, FRE, FRF, FRG, FRH, FRI, FRJ, FRK, FRL, FRM, HR0, HU1, HU3, IE0, IS0, ITC, ITF, ITG, ITH, ITI, LT0, LU0, LV0, MT0, NL1, NL3, NL4, PL2, PL4, PL5, PL6, PL7, PL8, PL9, PT1, PT2, PT3, RO1, RO2, RO3, RO4, RS1, RS2, SE1, SE2, SE3, SI0, SK0, UKC, UKD, UKE, UKF, UKG, UKH, UKI, UKJ, UKK, UKL, UKM, UKN.

A. Weekly air traveller traffic by Country



II. Regressions for additional time periods

The purpose of this paper was to specifically identify the current reality where COVID-19 is widely present in communities. The tables below present the results of OLS regressions (within models for panel data) for additional time periods. The regressions looking at longer periods would provide useful information for considerations about future novel viruses, but are not as relevant for COVID-19 given its wide-spread prevalence today.

It is unsurprising to see that the Air Passengers variable has a much higher coefficient during H1 2020 – as stated in the introduction this supports previous research findings.

Table 4 Subnational regressions other time periods

Subnational: Relationship between Air Passengers & COVID-19 rates during 2020 H1 and H2

	Dependent Variable: Rate of new COVID-19 cases per 100k population (log)		
	Fixed Effects	Fixed Effects H1	Fixed Effects H2
Air Passengers in prior week (log)	0.021	0.026	0.014
	$p = 0.057$	$p = 0.007^{**}$	$p = 0.009^{**}$
COVID-19 rate in prior week (log)	0.910	0.946	1.003
	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$
Stringency Index in prior week (log)	-0.298	-0.665	-0.431
	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$
Community Mobility in prior week (log)	0.227	-0.053	0.294
	$p = 0.000^{***}$	$p = 0.180$	$p = 0.000^{***}$
Observations	3,345	978	2,367

R^2	0.835	0.879	0.968
Adjusted R^2	0.828	0.869	0.967
F Statistic	4,063.922*** (df = 4; 3210)	1,642.033*** (df = 4; 903)	17,186.480*** (df = 4; 2252)
Note:			*p<0.05; **p<0.01; ***p<0.001

Table 5 National regressions other time periods

National: Relationship between Air Passengers & COVID-19 rates during 2020

	Dependent variable:		
	Fixed Effects	H1 Fixed Effects	H2 Fixed Effects
Air Passengers in prior week (log)	0.093	0.193	0.015
	p = 0.000***	p = 0.000***	p = 0.093
COVID-19 rate in prior week (log)	0.929	0.798	1.015
	p = 0.000***	p = 0.000***	p = 0.000***
Stringency Index in prior week (log)	-0.036	-0.024	-0.380
	p = 0.000***	p = 0.003**	p = 0.000***
Community Mobility in prior week (log)	0.225	0.029	0.302
	p = 0.000***	p = 0.709	p = 0.00000***
Observations	1,476	536	940
R^2	0.934	0.820	0.972
Adjusted R^2	0.930	0.796	0.970
F Statistic	4,923.760*** (df = 4; 1387)	538.359*** (df = 4; 473)	7,549.417*** (df = 4; 867)
Note:			*p<0.05; **p<0.01; ***p<0.001

ACKNOWLEDGEMENTS

Francesca Arduini, Benedikt Mandel, Benny Mantin, Hans-Martin Niemeier and Mike Tretheway provided valuable comments on drafts of this paper.

DISCLAIMER

All statements and opinions in this paper represent the view of the author alone and are not positions of organisations that the author is affiliated with.

DECLARATION OF INTEREST

The author is an employee of the trade association representing the European airports sector.

REPRODUCIBILITY

All of the data used in this paper is available from public sources.

AUTHOR'S BIO

Michael Stanton-Geddes is the Head of Economics for Airport Council International – EUROPE (ACI EUROPE), based in Brussels, Belgium where he works on airport economics, including monitoring market structure, connectivity, and the economic benefits of aviation and airports to their local communities. Prior to joining ACI EUROPE, Michael held a number of roles at Fleishman-Hillard, the International Trade Commission and GE and as an economic consultant at the World Bank. Michael has degrees from Tufts University and the Johns Hopkins University School of Advanced International Studies (SAIS).