UBC SAUDER
SCHOOL OF BUSINESS

Prepared By

Bahareh Aghababaei

MSc student in Business Administration, Transportation and Logistics

Instructor

Prof. David Green

A referee report on an econometric analysis of domestic air traffic demand in regional airports: Evidence from India

April 13, 2023

1. Article information

Title: An econometric analysis of domestic air traffic demand in regional airports: Evidence from India

Authors: K. Chandrashekhar Iyer, Nivea Thomas

Year: 2021

Published in: Journal of Air Transport Management

2. Paper summary

The article provides an overview of the Indian aviation industry, which is experiencing significant growth due to factors such as low-cost carriers, modern airports, and advanced information technology interventions. The authors also offer background information on the current state of air transport in India, highlighting that the country has 106 operational airports, with plans to construct 100 more by 2024. However, despite this growth, the average number of air trips per person in India is just 0.08, indicating untapped potential in the aviation market, particularly within the domestic segment.

To fill a research gap, the authors focus on modelling domestic air traffic demand in regional airports, which are non-hub airports that generate insufficient demand to attract airlines and have fewer than three million passengers annually. Although often located in challenging geographic areas, they may not necessarily be far from towns or cities. The authors chose to study regional airports because there is a dearth of literature on assessing the value of regional aviation. In fact, previous studies on Indian air transport demand have concentrated more on hubs and large international airports. In addition, as domestic traffic continues to grow in India, the authors assert that the contribution of regional airports cannot be overlooked, necessitating an investigation of the factors influencing air traffic demand in these airports.

3. Literature review

Some studies have been conducted on air transport in regional and remote areas, mainly on the relationship between air transport activity and economic development. Baker et al. (2015) provided the first empirical evidence of a short- and long-term causality between regional aviation and economic growth. The authors analysed 88 regional airports in Australia from 1985–86 to 2010–11 to determine the catalytic impacts of regional air transport on regional economic growth. The results showed the economic significance of regional air transport confirms the importance of the airport as infrastructure for regional councils and the need for them to maintain and develop local airports.

In addition, the study by Zhang et al. (2017) investigated the market development patterns of regional airports in Australia. The authors found that a rise in commodity prices led to increased traffic volume in markets where the local economy heavily relied on mineral resources, while an appreciation of the Australian dollar resulted in decreased passenger flow in tourism-dependent areas. Additionally, the study identified the presence of leading airlines and low-cost carriers and the availability of international services as positive factors contributing to market growth.

Dziedzic et al. (2020) investigated the factors influencing airline customers' choice of airport, explicitly focusing on small airports and their catchment areas, to explain the variation in traffic volumes across 146 regional gateways in 21 EU countries in 2016. The study utilized multiple linear regression and correlation analysis and found that population size, airport charges, and the need for capacity coordination were the most significant factors related to the number of passengers using these airports. The research also explored the correlation between the proportion of low-cost, full-service, and charter carriers at small regional airports and airport traffic, revealing various correlations with moderate statistical significance and highlighting the need for deeper analysis of smaller samples of airports.

4. Data and methods

To ascertain the impact of diverse factors on air passenger volumes at regional airports, the authors utilized a multiple linear regression model applied to fifty-seven regional airports across India. The variables used in the model and their corresponding data sources are provided in Table 1

Table 1. list of variables considered in this study

Variables	Definition	Data source
Passengers (PAX)	Total number of domestic passengers at the airport	Airport Authority of India (AAI)
Population density (DEN)	Population density of the state	Indiastat

Per capita income	Per capita income of the state	Economic and Statistical Organization, Government of Punjab
Distance to the nearest airport (AIRCOMP)	Distance to the nearest airport in km	Flightradar24 database
Distance from the city centre (CITYCENTRE)	Distance from the city centre in km	Airport websites, Google Maps
Accessibility dummy (ACCESS)	Airports having public transport connectivity to the city centre with a value 1 and 0 for others	Airport websites, Travel websites
Hub dummy (HUB)	States with hub airports are given a value 1 and 0 for others	Manual calculations
Railways competition dummy (RAILCOMP)	Airports facing competition from the railway sector given a value 1 and 0 for others	Indian Railway Catering and Tourism Corporation (IRCTC)

The research investigates the determinants of air traffic demand in regional airports, focusing on the total number of domestic air passengers as the dependent variable. The researchers considered state population density and per capita income to capture the economic status of the state and the demand for air travel. The variable "Distance to the nearest airport" was also included in the model to account for competition from other airports, as passengers may choose to travel to other airports that offer more extensive air services. Another factor considered is the distance from the city centre, as accessibility is crucial for attracting passengers, and the coefficient of the variable is predicted to be negative if the city centre is far from the airport.

The authors also included three dummy variables to account for different factors that could affect the competitiveness of airports. Firstly, they used an "accessibility dummy" to indicate whether a particular airport has good public transport connections to and from the city centre, which was considered an essential factor in attracting passengers. Secondly, they used a "hub dummy" to show whether an airport is located in a state with a major hub airport, such as Delhi, Mumbai, Chennai, Kolkata, Bangalore, and Hyderabad, which collectively handle over 60% of the total passenger traffic in India. Finally, the authors also considered the extensive railway network in India as a critical competitor to domestic airlines, so they included a "railways competition dummy" to indicate whether the city where the regional airport is located has a railway station.

5. Results

Table 2 presents the outcomes of the regression model, which examines the influence of various variables on air passenger volumes in regional airports. The model's R2 value is 0.3979, and to identify the multicollinearity issue, the researchers checked the Variance Inflation Factor (VIF). The VIF assesses the degree of multicollinearity in the regression variables, and all values are below the recommended maximum of 10. The findings indicate that the distance to the nearest airport has a statistically significant impact on domestic air traffic volumes in regional airports. Additionally, the Accessibility dummy is statistically significant in the model, implying that public transport connectivity from the airport to the city center and vice versa positively affects the number of passengers at the airport. The other relationships obtained from this study are comparatively weaker.

Table 2. Multiple Linear Regression results

Variables	coefficients	P-Value	VIF
Population density	75.0324	0.292	1.32
Per capita income	2.196623	0.389	1.91
Distance to the nearest airport	4881.809	0.014*	1.22
Distance from the city centre	- 693.8043	0.956	1.06
Accessibility dummy	694460.8	0.000**	1.07
Hub dummy	- 370170	0.132	1.47
Railways competition dummy	- 512861.6	0.432	1.09
constant	8008.577	0.991	
R2 0.3979			
Adjusted R2 0.3119			

Note: *p < 0.05, **p < 0.01, these represent significant p values

6. Identification issues

The paper outlined above has several identification issues, likely stemming from simplifying assumptions made due to a lack of available data. In my opinion, the model for estimating the determinants of air traffic demand in regional airports lacks sufficient strength to identify the causal effects of coefficients on air traffic demand. The coefficients themselves are not consistent estimators. The high proportion of unexplained variance in the dependent variable, as indicated by the low R^2 value, further reinforces this issue. As such, this report will focus on identifying the specific identification problems present in this paper.

6.1. Inadequate Explanations and Lack of Comprehensive Reporting

The paper's explanation of the proposed model is insufficient and lacks depth. The authors only presented the coefficients and p-values without providing the standard deviation of coefficients and residual values. This incomplete presentation makes it challenging to interpret the model adequately. R-squared value alone cannot completely assess the model's performance. To evaluate the model's goodness-of-fit accurately, additional measures such as residual plots, standard errors of coefficients, and other relevant statistics should be considered. The lack of information on residuals and standard deviations of coefficients makes it challenging to evaluate the model's performance and draw conclusions about the accuracy of the estimated coefficients.

The identification problems discovered in the paper are listed as follows:

6.2. Endogeneity

Although the authors checked the Variance Inflation Factor to see if there is multicollinearity between independent variables, they did not evaluate the correlation between the error term and independent and dependent variables. This correlation causes the estimated coefficients to be biased and inconsistent, leading to incorrect statistical inference and incorrect conclusions about the causal relationships between the independent and dependent variables. Endogeneity here arises from omitted variable bias. Different factors affecting airport demand were not considered in this model.

- The tourist potential of the state has not been considered in the model. India is a huge tourism market, and regional airports are sometimes the only way to access tourist destinations. Domestic trips in India are for various purposes like business, holidaying, leisure, recreation, social, pilgrimage and religious, education and training, health and medical, shopping and others. The proximity of the airports to the city centre was considered in the model. However, the proximity to major cities, tourist destinations, and business centres was not included, while they can significantly impact the number of passengers it attracts. Airports in highly populated areas or near popular tourist attractions will likely attract more passengers.
- One of the crucial factors determining airports' air demand is the extent of the airline network in an airport. The number and types of airlines that operate at an airport can also

affect passenger traffic. Airports with a diverse range of airlines and a range of destinations are more likely to attract a more significant number of passengers.

- It should be noted that seasonal migration for work is a common practice in India, with individuals moving from rural areas to urban labour markets and industries. However, this particular study did not account for migration as a variable. With the introduction of the Regional Connectivity Scheme in 2016, air travel has become more accessible and affordable, which may lead to an increase in the number of migrant workers using flights to reach their destinations.
- Passenger traffic can be affected by the cost of flights, airport charges, and other associated expenses, indicating price competitiveness is an essential factor. Airports that provide affordable flight fares and services like parking fees are more likely to draw in higher passenger numbers (Dziedzic et al., 2020).
- Passenger traffic can also be influenced by the quality and capacity of an airport's infrastructure, including runways, terminals, and parking facilities. Airports with modern and well-maintained infrastructure are more likely to attract passengers (Dziedzic et al., 2020).

In the presence of endogeneity caused by omitted variable bias, the coefficient estimates can become biased and inconsistent, resulting in incorrect conclusions about the relationship between the independent and dependent variables. The direction of the bias can vary based on the estimators and the covariance between the regressors and omitted variables. For example, when there is a positive covariance between an omitted variable (such as proximity to tourist destinations, airline networks, or price competitiveness) and both the dependent variable and a regressor, the estimated coefficient of the included regressor is overestimated. This happens because the effect of the omitted variable is partially captured by the included variables, making it seem like their effects are more significant than they actually are.

To address this endogeneity caused by omitted variable bias, the author should include additional measurable relevant variables. For those for which there is no way to measure it directly, like the airport's service quality, they can either use a proxy variable or a control function. For the proxy variable, they need to choose a variable correlated with service quality, e.g., customer satisfaction ratings, and include it as an additional independent variable in the model. However, this model's effectiveness depends on the correlation's strengths between an omitted variable and a proxy.

6.3. Selection bias

The data collection process in this paper is affected by selection bias, as another issue that has been identified. Selection bias occurs when the sampling method used to select the observations for the study is not random or representative of the population being studied. It can arise in several ways, such as when individuals or units self-select into the study, or when the researchers intentionally or unintentionally select specific groups or individuals for the sample. In this particular study, the authors only included individuals above a certain income threshold as potential candidates for air travel, creating a selection bias. This bias is a major obstacle to valid statistical and causal inferences. It leads to inconsistent estimators as the sample does not represent the population, making the sample statistics deviate from the population parameters. Therefore, the estimates obtained from the biased sample do not accurately reflect the average causal effects of covariates on airports' air demand.

6.4. Insufficient Data variations

The data used by the authors was restricted to only one year, namely 2018, resulting in a limited number of observations of 57. This sample size is insufficient for conducting advanced data analysis techniques like regression. A small dataset may not adequately represent the entire population, and it may also have limited variation, which can constrain the effectiveness of the regression analysis. Furthermore, a small dataset can make the model more vulnerable to overfitting, whereby it fits the data too closely and captures the noise instead of the underlying patterns. Consequently, it may lead to a less accurate model and poor generalization of new data.

On the other hand, a larger sample size can enhance the precision of the estimated coefficients and lead to more statistically significant results with more minor standard errors. This will enable more reliable inference about the relationship among the variables in the model.

6.5. Measurement error

Furthermore, the authors implied that the data utilized in their study were subject to imprecise measurements. This means the model is tainted with measurement error which could reduce the reliability of the regression model. Measurement error refers to a circumstance in which the true empirical value of a variable cannot be observed or measured precisely. While measurement error to right-hand side explanatory variables will result in biased and inconsistent estimates in a multiple regression framework, the direction of bias is less clear and will depend on the correlations between the measurement errors of the various regressors.

Measurement error significantly affects the validity and accuracy of the estimated coefficients and the overall inference. If the measurement error is random, it can reduce the precision of the estimated coefficients, resulting in a decrease in statistical power and an increased risk of Type II errors. On the other hand, if the measurement error is systematic, it can lead to biased estimates, causing an altered perception of the relationship between the variables. Systematic measurement errors can also generate spurious relationships between the variables, leading to misleading conclusions.

7. Reference

Baker, D., Merkert, R., & Kamruzzaman, M. (2015). Regional aviation and economic growth: cointegration and causality analysis in Australia. Journal of Transport Geography, 43, 140-150.

Dziedzic, M., Njoya, E. T., Warnock-Smith, D., & Hubbard, N. (2020). Determinants of air traffic volumes and structure at small European airports. Research in Transportation Economics, 79, 100749.

Zhang, Y., Wang, K., & Fu, X. (2017). Air transport services in regional Australia: Demand pattern, frequency choice and airport entry. Transportation Research Part A: Policy and Practice, 103, 472-489.