AIRFARE

By: Mohammed Yusuf Shaikh & Jiaqi Sun

Email: mohammedyusuf.shaikh@mail.utoronto.ca & jqsarah.sun@mail.utoronto.ca

1006695783

Eco 375: Applied Econometrics

University of Toronto

Department of Economics

Assignment 1

Date 23 November 2022

Abstract:

Our research topic is how market concentration and the average number of passengers per day of a U.S. domestic flight route affect the average one-way fare of that route between 1997-2000. We chose this research topic because around 91% of all U.S. airline passengers in 2021 were on domestic flights, so understanding factors affecting domestic airfare impacts many people. We will be adding fixed effects to multiple regression, which few studies have done.

---

1. Introduction

Our data project focuses on the data extracted from 1997-2000 based on US domestic airlines. According to recent existing articles there has been a dilemma on the relationship between airfare and market concentration. There are various factors that affect airfare like distance, average number of passengers travelling on domestic flights. Based on various factor we want to determine the factors affecting airfare and its relationship with the variables

2. The Context and Data

The data in this analysis were provided by Jiyoung Kwon, a former doctoral student in economics at MSU, who obtained the data from the Domestic Airline Fares Consumer Report by the U.S. Department of Transportation. The data follows a panel structure. Our data covers all 1149 U.S. domestic routes in the dataset, over 4 time periods (1-year intervals) from 1997 to 2000, hence a route-year unit of observation. The primary variables of interest are a measure of airfare, the natural log of the average one-way fare for each route measured in U.S. dollars, and market concentration, which can range from 0 to 1 and measures the share of business on an air route accounted for by its largest carrier — low market concentration means the competition is high among airline companies on that route and high market concentration means a monopolist or few oligopolists controls most of the flights available for that route. Table 1 provides summary statistics of these primary variables of interest and some other variables. The market concentration ranges from 0.1605 to 1 with a mean value of 0.6101 and a standard deviation of 0.1964, which indicates moderate variation in market concentration across route-year in the dataset. The average one-year fare ranges from 37 to 522 with a mean value of 178.80; however, over half of the route-year have fares that reside on the left half of the range. The sizable variability in fare data is also represented by the large standard deviation of 74.88. As well, we observe that the daily average number of passengers of all route-year ranges from 2 to 8497 with a mean of 637 and standard deviation of 812 as a result of few route-year having large number of passengers.

3. **Regression analysis**

Simple Linear Regression

Base Case

To begin our estimation of the relationship between average one-way fare and market concentration, we utilize a simple linear regression specification (Table 2 specification (1)) modeled by $fare_{i,t}=$

$$\beta_0+\beta_1$$

concentration $percent_{i,t}+$

$$e_{i,t}$$

for a route i in year t. Market concentration is multiplied by 100 to convert it from fraction to percent, which simplifies the coefficient estimator's interpretation. We use robust standard errors to account for any heteroskedasticity. This specification shows that a 1 percentage point increase in a route's market concentration is, on average, associated with 73 cent decrease in its average one-way fare. We expect 95 percent of other samples' measure of

$$\beta_1$$

to fall within the interval -.83 and -.63 (63 cents to 83 cents decrease in average one-way fare for a 1 percentage point market concentration increase). With a large *t*-statistic of -14.18, we reject the null (

$$H_0:\beta_1=0$$

) that there's no link between market concentration and airfare. However, this result is economically insignificant. For instance, the standard deviation of market concentration is 19.64 percentage points, so a 1 standard deviation change in market concentration would be a $14.33 decrease in average one way fare. Meanwhile, the standard deviation in average one-way fare is $74.88, so a 1 standard deviation difference in market concentration is associated with only 19% of a standard deviation in average one-way fare. Also, $15 isn't a salient difference when people compare airline tickets and decide which one to buy. However, as discussed in The Context and Data section, the distribution of average one-way fare is right-skewed, which means there are extremely high fares present in the dataset that heavily influence the estimate of

$$\beta_1$$

, making the coefficient estimator misleading.

Specification (2) takes the log of the average one-way fare to correct skewness and regresses it on market concentration. We again use the robust standard error. Specification (2) shows that a 1 percentage point increase in a route's market concentration is on average associated with a 0.49 percent decrease in its average one-way fare. We expect 95 percent of other samples' measure of

$$\beta_1$$

to fall within the interval -.0055 and -.0043 (0.43 percent to 0.55 percent decrease in average one-way fare for a 1 percentage point market concentration increase). The *t*-statistic of specification (2), -15.95, suggests market concentration becomes an even more statistically significant predictor after we corrected for skewness. Also, this result has moderate economic significance. For instance, a 1 standard deviation change in market concentration would be a 9.63 percent decrease in average one-way fare, which could result in a notable difference in the dollar value of air airline tickets, especially for more expensive tickets. However, this regression neglect to account for characteristics of different airlines that could explain differences in average one-way fare between different routes, such as whether the destination of the route is a popular travel region. This omitted variable bias suggests that (2), as well as (1), conflate variation in the average one-way fare driven by systematic differences between routes with market concentration, resulting in an overestimation of the effect of market concentration on average one-way fare and violation of the zero conditional mean assumption/

Multiple linear regression

Furthermore, there can be various other factors affecting avg on way fare other than market concentration percentage therefore which could be tested with multiple linear regression like distance, avg no. of passenger. The multiple regression runs for log of fare as dependent Y variable with market concentration as X1 variable and log of avg no. of passengers as X2 variable (control variable).
log(fare)=b0+b1(concenpct)+b2log(passen)+a+u
This specification shows that a 1 percentage point increase in a route's market concentration is associated with 51 cent decrease in its average one-way fare on average. This specification also shows that a 1 percentage point increase in average no. of passengers is associated with 12 cent decrease in its average one-way fare on average. The log of passengers increases the negative correlation of market concentration percentage and avg number of passengers. Also R^2 increases which shows the model is best fit. However, introducing more variable like distance and log of distance square. The R^2 increases to 42.3% from Table 3.

Limitations of results

The simple model suffers from omitted variable bias because we found correlation between concentration and passen, and a linear relationship between fare and passen. We also found nonlinear relationships between fare and concen, also fare and passenger, so we took log of fare and log of passenger in the extension model.
Since our data is panel, there might be unobserved factors that cause omitted variable bias.

Conclusion

Market concentration had economically insignificant but statistically significant negative effect on percentage change in fare and percentage change in average number of passengers has both economically and statistically significant negative effect on percentage change in fare, and fare became more expensive from 1997-2000.
In a nutshell the simple regression model did not include all the parameters required for a strong model. By adding more variables we proceeded with multiple regressions to get a clear understanding of the model analysed in our data project. The higher the market concentration on a route, the lower the

effect of the restriction on airfare. Based on finding we can conclude that relationship between market concentration avg. on way fare is significant however there relationship is slightly dependent

---

References:

Abdallah, Farid, and Walid Elhoss. 2019. “The Impact of Airport Congestion on Airfares: A Case Study on Lebanon.” In , 359–64. https://doi.org/10.1109/ICCIKE47802.2019.9004266.

- Capacity constraint / congestion is positively correlated with airfare

Gong, Gang. 2006. "Airfare, Competition, and Spatial Structure: New Evidence in the United States Airline Deregulation." Order No. 3186505, Boston University. http://myaccess.library.utoronto.ca/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fdissertations-theses%2Fairfare-competition-spatial-structure-new%2Fdocview%2F305360733%2Fse-2%3Faccountid%3D14771.

- show that: (1) route level airfare is significantly influenced by the hub status of the origin and destination airports; (2) the level of concentration and dominance at airport and route markets have important but variable impact on airfare; (3) the existence of low fare carriers has considerable impact on pricing in the market. Competition is effective in restraining the dominant carrier's market power; (4) there exists significant network autocorrelation among airfare at the level of route markets.

Zhang, Dapeng and Xiaokun Wang. 2016. "Investigating the Dynamic Spillover Effects of Low-Cost Airlines on Airport Airfare through Spatio-Temporal Regression Models." Networks and Spatial Economics 16 (3) (09): 821-836. doi:https://doi.org/10.1007/s11067-015-9300-z. http://myaccess.library.utoronto.ca/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fscholarly-journals%2Finvestigating-dynamic-spillover-effects-low-cost%2Fdocview%2F1810692309%2Fse-2%3Faccountid%3D14771.

- Correlation between fraction of market and airfare: A spatio-temporal regression model is used to analyze the relationship between airport airfare and low-cost airlines' market shares and show the impact of low-cost airline over time and space by the dynamic spillover-effect function. A case study of market share improvement of AirTran Airway at Albany International Airport (ALB) is used to illustrate the dynamic spillover effect. Results show that the average airfare of ALB will drop $18.64 right after AirTran takes 10 % market share, continue to drop in the following years, and stabilize at around 10 years. Airfare at other airports will also drop, and the magnitudes decrease as time and distance to ALB increase.

Table 1:

	Model 1 b/se	Model 2 b/se	Model 3 b/se
= bmktshr	-72.704***	69.431***	-77.976***
	(5.52)	(5.12)	(5.48)
avg. passengers pe-y		-0.005***	-0.014***
		(0.00)	(0.00)
distance, in miles		0.087***	
		(0.00)	
Constant	223.154***	52.995***	235.452***
	(3.54)	(4.49)	(3.68)
R-sqr	0.036	0.418	0.060
dfres	4594	4592	4593
BIC	52560.3	50257.3	52454.2

* p<0.05, ** p<0.01, *** p<0.001

Table 2:

```
. corr lfare concen lpassen
(obs=4,596)
```

	lfare	concen	lpassen
lfare	1.0000		
concen	-0.2213	1.0000	
lpassen	-0.2340	-0.0440	1.0000

Table 3

```
. do "C:\Users\LOCAL_~1\Temp\10\STD2830_000000.tmp"

. reg lfare concen lpassen, robust

Linear regression                               Number of obs     =      4,596
                                                F(2, 4593)        =     299.14
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1085
                                                Root MSE          =     .41214
```

lfare	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
concen	-.5154331	.0295652	-17.43	0.000	-.5733952	-.457471
lpassen	-.1205644	.007178	-16.80	0.000	-.1346368	-.106492
_cons	6.135513	.0451339	135.94	0.000	6.047029	6.223997

Table 4:

```
hausman fixed ., sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          .          Difference          S.E.
-------------+----------------------------------------------------------------
  concenpct  |    .0002373     .0009587       -.0007213        .0001335
     passen  |   -.0002941    -.0001419       -.0001522        .0000123
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      163.31
                Prob>chi2 =      0.0000
```

Table 5:

```
. xtreg lfare concenpct lpassen i.year, fe vce(robust)

Fixed-effects (within) regression               Number of obs     =      4,596
Group variable: id                              Number of groups  =      1,149

R-sq:                                           Obs per group:
     within  = 0.5070                                         min =          4
     between = 0.0392                                         avg =        4.0
     overall = 0.0497                                         max =          4

                                                F(5,1148)         =     254.31
corr(u_i, Xb)  = -0.4865                        Prob > F          =     0.0000

                                  (Std. Err. adjusted for 1,149 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lfare |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   concenpct |   .0015004   .0003421     4.39   0.000     .0008293    .0021715
     lpassen |  -.3695845    .024027   -15.38   0.000    -.4167263   -.3224427
             |
        year |
       1998  |   .0297916    .003244     9.18   0.000     .0234269    .0361564
       1999  |   .0590909   .0038642    15.29   0.000     .0515091    .0666726
       2000  |   .1285602   .0041662    30.86   0.000     .1203859    .1367345
             |
       _cons |   7.173499   .1497902    47.89   0.000     6.879606    7.467392
-------------+----------------------------------------------------------------
     sigma_u |  .48051657
     sigma_e |   .0804331
         rho |  .97274465   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```