Examining the American Airlines - US Airways Merger through a Structural Model

Sravan Ramaswamy

May 2021

1 Introduction and Goal

Throughout the 2000s and 2010s, the airline industry has been wracked with notable instability featuring bankruptcies, technological innovation, an increased number of competitive entrants in the form of low-cost carriers, and large-scale mergers. Steven Berry and Panle Jia describe in their paper, "Tracing the Woes: an Empirical Analysis of the Airline Industry," (2010) that all six of the major legacy carriers in the early 2000s (American Airlines, United, US Airways, Northwestern, Delta, and Continental Airlines) reported sharp reductions in profits with four of the six major carriers undergoing significant bankruptcy organization. Their structural model estimates the impact of the changes in demand and supply on legacy firm profitability.

Berry and Jia focus their analysis on data from 1999 and 2006. Since then, several significant industry events have occurred concerning the six legacy firms that impact the results of their research. First, in 2008, Delta completed its merger with Northwestern, and the latter ceased to exist independently in 2010. Also, in 2010, United merged with Continental. Lastly and most relevant to this paper, US Airways merged in 2013 with American Airlines and ceased to exist independently in 2015.

This project aims to partially replicate and extend the specifications Berry and Jia use with new data and counterfactual exercises to examine the impact of the American Airlines - US Airways merger on prices. Instead of reviewing 1999 and 2006, I will be focusing on the second quarter across four years, 2011, 2012, 2015, and 2016.

This analysis will consist of three main components. The first is a naive difference-in-difference analysis across all four years, with the former two acting as the pre-merger period and the latter two as the post-merger period. The second is a set of nested logit demand specifications using similar specifications to Berry and Jia (2010) using pre-merger data. The third is a counterfactual merger analysis between US Airways and American Airlines utilizing the estimated pre-merger demand and supply specifications.

2 Data

There are two primary sources of data that I will be using for this project, the Airline Origin and Destination Survey (DB1B) and the Airline On-Time Performance data. The DB1B is a random 10 percent sample of airline tickets from reporting US carriers, including itinerary, airline carrier, and fare information for each quarter.

The Airline On-Time Performance Data provides information on monthly flight-level on-time departure data for nonstop domestic flights. This data is used to construct the departure frequency and delay variable for each product. In addition, to determine the size of each market, I will also be using MSA Population data from the 2010 Census.

The market definition consists of a directional pair of airports (so "LAX-DCA" and "DCA-LAX" are two different markets). The market size is determined by the geometric mean of the MSA populations at the endpoint cities for a given route.

For this project's scope, I will only focus on one-way non-stop flights between the twenty busiest airports in 2016[1]. This is because a large number of unique products and markets

[1] https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/media/cy16-commercial-service-enplanements.pdf

substantially increases the time to estimate the results. In addition, none of the airports are geographically close together and thus do not need to be grouped into a single market.

A binned fare captures the idea that consumers may not see the difference between a \$300 fare and a \$320 fare for a given itinerary as substantially unique products for a given market and carrier.[2]. The product definition, similar to Berry and Jia, will consist of a unique combination of the origin airport, the destination airport, the ticketing carrier, and a binned fare, with the addition of year for the difference-in-difference analysis. The product share is given by the proportion of passengers in a given price "bin" for a given carrier in a market.

To evaluate the price changes as part of the AA-US merger, I will be using four cross-sections of data: the second quarter of 2011, 2012, 2015, and 2016. The second quarter is utilized to minimize the seasonal effect of travel seen in the summer and the holiday months.

3 Difference-in-Difference Model

To construct the naive difference-in-difference analysis, I examine the difference in fares on products sold by American Airlines and US Airways Pre- and Post-merger (treatment group) to the difference in fares for products not sold by those carriers (control group). I then construct weights using the market share of each product. Finally, all fares are deflated by the CPI index, with 2011 as the base year. To handle the coding for this section, I used R.

3.1 Summary Statistics

Table 1 describes the summary statistics for the difference in difference analysis. The sole focus of the analysis is on the change in prices that occurred after the AA-US merger

[2]Like Berry and Jia, I use the following set of bins: \$20 for all tickets under \$700 (so tickets between \$300 and \$320 with the same itinerary and ticketing carrier are aggregated as one product), \$50 for tickets between \$700 and \$1,000, and \$100 for tickets above \$1,000.

Variable	Mean	SD	Min	Max	P25	P50	P75
Premerger Weighted Fares (Control)	232.3747	234.5532	10	14200	120	170	270
Premerger Weighted Fares (Treatment)	276.4169	285.4944	10	9700	150	220	400
Postmerger Weighted Fares (Control)	185.8563	214.9882	10	15000	80	140	220
Postmerger Weighted Fares (Treatment)	241.7727	287.9493	10	59200	110	170	280

Table 1: Summary Statistics for Naive Difference-in-Difference Analysis

in 2013; the only relevant statistics here are those that concern the weighted price. There are 518,512 passengers in the control group and 202,559 passengers in the treatment group in the premerger period. There are 871,239 passengers in the control group and 313,176 passengers in the treatment group in the post-merger period.

3.2 Model

The model used for the naive difference-in-difference analysis is straightforward.

$$P_{jt} = \gamma * \text{ Treatment}_j + \lambda_t * \text{Year}_t + \delta * D_{jt} + \epsilon_{jt} \tag{1}$$

P_{jt} is the price of product j at year t. γ is the coefficient of the treatment variable. λ_t is the coefficient of the year variable, which is a dummy variable representing the years of data after 2013. The merger occurred in 2013, and so, 2015 and 2016 are the post-merger period which is denoted by the merger variable. D_{jt} is this merger variable multiplied by the treatment variable, and ϵ_{jt} is the error term. δ in the traditional DID framework represents the treatment effect on the treated products and is the central coefficient that we are examining in this model. Under a strict exogeneity assumption of the treatment variable D_{jt}, it can be shown that the DID estimator is the following:

$$\hat{\delta} = \left(\bar{P}_{\text{treatment},2} - \bar{P}_{\text{treatment},1}\right) - \left(\bar{P}_{\text{control},2} - \bar{P}_{\text{control},1}\right)$$

where

$$\bar{P}_{\text{treatment},t} = (\textstyle\sum_{j=1}^{J_{\text{treatment}}} P_{jt}w_j)/(\sum_{j=1}^{J_{\text{treatment}}} w_j) \quad \forall t = 0, 1$$
$$\bar{P}_{\text{control},t} = (\textstyle\sum_{j=1}^{J_{\text{control}}} P_{jt}w_j)/(\sum_{j=1}^{J_{\text{control}}} w_j) \quad \forall t = 0, 1. \tag{2}$$

$J_{\text{treatment}}$ is the number of products in the treatment group. J_{control} is the number of products in the control group. w_j is the market share of product j, which is used as a weight to calculate the average fare.

3.3 Implementation and Results

Table 2 describes the results of the naive difference-in-difference analysis. What it finds is a very statistically significant effect on D's coefficient. This result suggests that following the AA-US merger, the average price for products sold by American Airlines and US Airways rose over \$14 on average compared to products offered by other airlines post-merger. However, the causal reasons for this price change can be many-fold. The analysis lacks the proper randomization to determine whether the AA-US merger was the primary cause for the increase in prices.

3.4 Challenges

The primary difficulty in this analysis and the later sections was obtaining and cleaning the data, which took many person-hours due to the multiple different datasets used and the initial lack of clarity regarding many distinct variables. However, the data used for this particular model is limited to the binned fares of each product and the shares of each product within its market. Therefore, there is an additional possibility that some variables were improperly cleaned in this project, introducing measurement error into the estimation.

Earlier in this project, I had erred in establishing my difference-in-difference specification concerning the observable unit of the data. In addition, I struggled to specify precisely how Berry and Jia denote the market share of a product. In addition, I neglected to include

Table 2: Naive Difference-in-Difference

	Dependent variable:
	Fare
Treatment	43.769***
	(2.297)
2012	−12.878***
	(2.011)
2015	−22.474***
	(2.034)
2016	−75.124***
	(1.925)
D	14.590***
	(2.948)
Constant	239.053***
	(1.565)
Observations	148,963
R^2	0.024
Adjusted R^2	0.024
Residual Std. Error	0.359 (df = 148957)
F Statistic	724.871*** (df = 5; 148957)
Note:	*p<0.1; **p<0.05; ***p<0.01

year-fixed effects for the difference-in-difference specification in an earlier version of Table 2.

4 Premerger Nested Logit Model

4.1 Model

The model here is a vein of vertical nested logit models with differentiated products ala Berry (1994). In an oligopoly setting, US Airlines offer sets of differentiated products; the product differentiation arises from the directional airport pairs, airline brand, price point, frequency of departures, but not the number of connections since I only examine nonstop flights. The result is an endogenous substitution to the outside good since we do not observe flights with multiple segments that some individuals may elect to take in between flying nonstop and not flying at all. In addition, I do not utilize departure times which, like other product-unobservable characteristics correlated with the price, should be incorporated since aspects like the ability to refund can have drastic impacts on ticket price. For product j in market t, the utility of the consumer is given by

$$u_{jt} = x_{jt}\beta - \alpha p_{jt} + \sigma \ln\left(s_{jt|g}\right) + +\xi_{jt} \tag{3}$$

where $s_{jt|g}$ measures product j 's within-group (nest) market share, and σ is a parameter measuring how substitution patterns differ within and across nests. Since $s_{jt|g}$ is endogenous by construction in these models, I need at least two instruments to deal with the endogenous p_{jt} and $s_{j|g.}$ x_{jt} is a vector of product characteristics and ξ_{jt} is the unobserved (to researchers) characteristic of product j.

For instruments, I use route level characteristics of rival products, such as the average distance of rival routes and the number of ticket carriers. I also use a dummy variable to denote the destination as a hub airport for the carrier. Lastly, I use fitted values of the 25th and 75th quantile of fares for a given route, obtained through a quantile regression of fares on exogenous route characteristics.[3] For flight instruments, we regress departures on the characteristics of end-point cities and use the fitted departures as instruments[4]. Finally, I use the exogenous variables (such as distance) in the demand specification as instruments in the share equation and marginal cost function.

Berry and Jia assume that prices are set according to a static Nash equilibrium with multiproduct firms for the supply side. I also implement a simple marginal cost specification given by

$$mc_{jt} = w_{jt}\psi + \omega_{jt} \tag{4}$$

where w_{jt} is a vector of observed cost-shifters, ψ is a vector of cost parameters to be estimated, and ω_{it} is an unobserved cost shock.

I include the hub density effect outlined in BCS (2007) by including the hub variable in the marginal cost equation and assume, like Berry and Jia, that the hub structure does not change over time.

4.2 Implementation and Results

To estimate the model, I implemented a general method of moments routine which I solved using MATLAB's fminsearch function. The results of the estimated model are shown in Table 3. The results of the estimation vary in comparison to Berry and Jia. All of the airline dummy variables lack statistical significance in the supply specification. However, this is not an issue with the demand-side dummy variables. The hub variable in supply seems to denote a fall in marginal costs for flights coming to or from a hub airport for an airline. This outcome is in line with what Berry and Jia find. I also find that slot-controlled airports increase marginal costs and increase demand which differs from Berry and Jia. That

[3] The fitted fares are obtained from quantile regressions of fares on the following exogenous variables: carrier dummies, route level characteristics (distance, whether or not in tourist locations like Las Vegas, etc.), market size (given by the geometric mean of MSA population between origin and destination), number of competitors, and the carrier's share of cities connected via nonstop flights at both the origin and the destination airport.

[4] The regressors that predict departures are similar to those in the fare quantile regression, except that I (as well as Berry and Jia) also include the hub status of both end cities.

said, there are a great many quantitative and qualitative differences. Berry and Jia estimate connections and the fare coefficient for two types of passengers, while I only assume a single type here. The implementation can reason why the constant's magnitude is much larger as I aggregate these two consumer types. Next, the nested logit coefficient is larger in magnitude in Berry and Jia's model. I numerically constrain the coefficient, which may create the difference seen here. Qualitatively, some airline dummies differ from their counterparts in Berry and Jia, namely, Continental and Southwest; however, Southwest is close in magnitude to zero, making Continental's relatively significant difference in magnitude more notable.

4.3 Challenges

In attempting to replicate Berry and Jia, I immediately encountered an issue where data on departures and flight schedules among non-direct flights was not available to the public. As such, I could not recreate the set of variables used to derive their results on connecting flights. Instead, I relied only on the non-stop flight schedule data present in the Airline On-Time Performance data to construct the product data to the best of my ability.

Throughout this setup, many identification concerns may indicate why the above specification fails to replicate the qualitative outcomes that Berry and Jia find. One particularly problematic aspect is the choice to focus solely on nonstop flights when many types of flights are present in the industry and whose inclusion can induce a much richer substitution property for the model. There is also the potential matter of potentially weak instruments. These identification concerns should cause the use of the nested logit specification to fail to converge towards the true value. These limitations are on top of the issues that Berry and Jia already discuss, including endogenous revenue management by carriers, being unable to model missing products, which is even more problematic here. Consumers are also assumed to not have any choice in their purchase date and that the number of carriers in a market is temporally uncorrelated to demand shocks.

In the process of estimating the model, I struggled to narrow down appropriate estimates

Table 3: Nested Logit Parameter Estimates

Variables	Demand	Supply
Constant	-18.7^{***} $(1.03 * 10^{-6})$	-0.123 (1.83)
Fare	-1.05^{***} $(9.21 * 10^{-7})$	
Number of Departures	0.833^{***} $(1.63 * 10^{-6})$	
Number of Destinations	0.253^{***} $(1.76 * 10^{-7})$	
Distance	6.9^{***} $(1.248 * 10^{-6})$	
Distance^2	-4.11^{***} $(6.85 * 10^{-7})$	
Tourist	0.183^{***} $(5.98 * 10^{-7})$	
Slot	0.179^{***} $(3.51 * 10^{-6})$	0.213^{***} (0.006)
σ	0.521^{***} (1.00*10^{-6})	
Small Distance		-0.089^{***} (0.017)
Large Distance		0.478^{***} (0.005)
HubMC		-0.137^{***} (0.044)
Other Carriers	0.16^{***} $(1.54 * 10^{-5})$	0.653 (3.23)
American	-0.063^{***} $(4.45 * 10^{-6})$	1.25 (2.79)
Continental	1.46^{***} $(1.19 * 10^{-5})$	0.985 (8.81)
Delta	-0.574^{***} $(5.10 * 10^{-8})$	0.947 (3.06)
United	0.11^{***} $(8.81 * 10^{-6})$	1.43 (3.36)
US Airways	-0.137^{***} $(6.93 * 10^{-6})$	1.01 (3.16)
JetBlue	2.06^{***} $(5.40 * 10^{-6})$	0.0001 (4.15)
Southwest	0.0001^{***} $(9.23 * 10^{-6})$	-0.733 (4.27)
Observations	47321	47321

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

and standard errors. On BCApps versus my machine, I would find wildly different results for the standard errors, which could differ by a factor of tens of thousands. I was not able to identify the root cause of the problem. Still, I currently suspect it has to do with the differences in operating systems where fminsearch terminates earlier or later. This experience also cautioned me about the sensitivity of these functions in estimation.

Computationally, a large number of products in my estimation presented a more base concern: a lack of memory on my personal computer. I encountered memory limitations in calculating the price elasticity matrix since I worked in the 'double' numeric type. This is a 64-bit or 8-byte numerical data type which means each cell of this array stores 8- bytes of data. Across a matrix of size 47321 x 47321, this amounts to over ten gigabytes of memory. This matrix is inefficient to load and, on many non-dedicated computers, cannot be computed entirely. I typecast the price elasticity to the 'single' type, a 32-bit/4-byte data type, to resolve the problem.

5 Counterfactual Merger Analysis

5.1 Implementation and Results

The counterfactual merger was conducted only on the premerger period data and consisted of changing all US Airways-owned products to American Airlines products. In addition, I reassign the hub variables according to the new firm ownership and combine the number of departures between identical products of US Airways and American Airlines. I then estimate the shares using the demand function coefficients and recalculate the pricing equilibrium using the Bertrand oligopoly first-order conditions. Table 4 displays the estimated industry results of a counterfactual merger between American Airlines and US Airways on average prices, aggregate price elasticity, profits, etc. Like Table 1, averages were weighted by the market share of the product.

There is a predicted fall of around $25 in the weighted average price due to the industry

Table 4: Counterfactual Merger Comparison

Value	Merger	Initial
Aggregate Price Elasticity	-0.175	-0.227
Average Price	219.14	243.74
Lerner Index	0.0051	0.0039
Estimated Average Change in Profits	3,448.44	
Estimated Average Change in Revenue	-249,214	

merger. There also appears to be a rather substantial decrease in the amount of revenue and profits collected. The Lerner index, which rises in the merger, suggests that the marginal cost is close to equating the price in the industry pre or post-merger, which is likely an incorrect result. Berry and Jia find that the average price in the industry is over twice the marginal costs that they estimate. The aggregate price elasticity[5] decreases in magnitude from the merger. On average, the price of the products owned by American Airlines and US Airways fell around $11 due to the counterfactual merger. Other airlines only marginally adjusted their average prices from the premerger period, creating a slight contradiction with the results of the difference-in-difference analysis. However, the expected change in profits for the merged firm is around -$28,500. This result suggests that the merger should not be enacted during this period which aligns with reality as the merger occurred in 2013.

5.2 Challenges

The main difficulty was computational; recalculating the price equilibrium means inverting the price elasticity matrix often to pursue convergence towards the new price equilibrium. Given the large number of products in the premerger period, this is rather strenuous for my personal computer. I often ran out of memory while running the code, and this specific

[5] The aggregate price elasticity measures the percentage change in total demand when all products' prices increase by 1 percent

section of the code occasionally crashed the system outright. To resolve this issue, I did not run the code locally. Instead, I used the BCApps version of MATLAB to run my code and allotted more system memory towards its use. I also implemented the same memory management tools to compute the first price elasticity matrix for the nested logit demand and supply model. In addition to inverting the elasticity matrix, I also had to recalculate the new within-group shares for each product before solving the new price elasticity matrix.

In addition, there was some difficulty in debugging unusual results such as profits appearing as negative for some firms and creating specific cases to handle the occasionally awkward cases that could appear in the code, such as NaN values. In addition, at this stage, I recognized the importance to normalize covariates such as distance and fare since they had a much larger scale than other variables.[6]

6 Conclusion

In general, this partial replication task has revealed to me reasonable means of improving my general workflow when it comes to larger-scale projects that involve heavy aspects of coding. How much one should handle at a time is of the most vital importance. Dynamic questions of what can or cannot be replicated are filled with compromises that come with the study's exploratory nature. The different perspectives on the impact of the AA-US measure through Berry and Jia's model seem to struggle under the conditions I elected to estimate my model. As such, it might be worth considering other variables that can more finely describe the quality of specific seats on a plane to explain more of the variance in prices on airplane seats. Counterfactual simulation with many products is a computationally demanding aspect of estimation; thus, properly detailing the pseudocode before estimation is vital to maintaining a good workflow. While many of these lessons can seem basic in a vacuum, developing a keen sense of the observation structure in the data is a crucial element of creating more successful projects in the future.

[6] Project Code is at `https://github.com/SravanjR/bc-micro-methods/tree/main/Project/Code`

References

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25(2):242–262.

Berry, S. and Jia, P. (2010). Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43.