Supplement to "Ex-post evaluation of the American Airlines-US Airways merger: a structural approach"

by Christian Bontemps, Kevin Remmy and Jiangyu Wei

B Robustness checks

In this section, we document some robustness checks. We discuss the choice of instruments then, the scenario for imputing the values in the counterfactual analysis. In the main text, we have selected our set of instruments in order to have a minimum degree of overidentification to get optimal and stable estimates, while not rejecting the overidentification test. In this robustness analysis, we report the **demand** side estimates and some estimates for other choices of sets of instruments. We use the following sets:

IV 1: Instrument set for 2011 used in the main specification

IV 2: Instrument set for 2011 without interaction variables

IV 3: The three instruments which are used in common for both 2011 and 2016

IV 4: Instrument set for 2016 used in the main specification

IV 5: Instrument set for 2016 without interaction variables

We build these five instrument sets for both year's data. The results are reported in Tables 16 and 17. We also display the estimated elasticities as well as the overall price changes calculated from the counterfactual analysis of Scenario 1.

We also add a column of results where we omit the distance squared in the product attributes, combining this specification with the optimal set of instruments chosen in the paper (IV1 for 2011 and IV4 for 2016). Indeed, we observe that the utility to fly is an increasing function of the distance for most of the products of our sample. It seems natural to check whether introducing the distance linearly in the utility function changes the results.

The first column (OLS) holds estimates from a nested logit model where we do not instrument for price nor the within-group share of products. For these non consistent estimates, we notice that the price coefficient is biased towards zero, which is expected as the true coefficient is negative and the correlation between the unobserved product attribute and the price positive (more demanded products are higher priced).

For 2011, we see that the estimates are relatively robust for the first three columns. Then, the estimates become very unprecise, especially with IV3 (only one degree of overidentification). IV4 is rejected by the J-test. IV5 is not but gives very large confidence intervals for the parameters. Overall, the estimates of elasticities and counterfactual are relatively robust.

For 2016, the estimates are also sensitive to changes in the instrument sets. Note that here, the column "IV 4" is the default specification used in the main part of the paper and provide the more accurate estimates. However, the estimates using the sets "IV 1" and "IV 2" does not seem to be be admissible from an economic point of view. Also, we reject the J test for these two cases.

We see that for 2011 and 2016, we recover a U-shaped dependence of utility on distance and that not including the square slightly changes the estimates but not the outcomes of the model.

C Details on the estimation procedure

C.1 Details on the supply-side + scenario 2

The first-order conditions with respect to the prices of the products offered in market t are given by equation (8), for all $j \in \mathcal{J}_t$. We can stack the first order conditions to get the following matrix/vector equality:

$$s_t + \Delta_t[p_t - mc_t] = 0_{\mathcal{J}_t}, \tag{C.1}$$

where s_t and mc_t are vectors collecting the market shares and marginal costs, respectively, of products offered in market t. The matrix Δ_t is of dimension $\mathcal{J}_t \times \mathcal{J}_t$ and holds own- and cross-price derivatives, with

$$\Delta_{kl,t} = \begin{cases} \frac{\partial s_{lt}}{\partial p_{kt}} \text{ if } k, l \in J_{ft}, \\ \\ 0 \text{ otherwise.} \end{cases} \tag{C.2}$$

Observe that s_t, Δ_t depend on the product attributes of all products proposed in the market, including the prices and the marginal cost of each product depends on its attributes.

When one knows the demand function and the marginal cost, the price can be derived from looking at the fixed point of equation (C.1). For example, assume that in market t, one firm offers products 1 and 2 and another firm offers product 3, the prices of these three products solve the following FOC system (omitting the product attributes):

$$\begin{bmatrix} s_1(p_1,p_2,p_3) \\ s_2(p_1,p_2,p_3) \\ s_3(p_1,p_2,p_3) \end{bmatrix} + \begin{bmatrix} \frac{\partial s_1}{\partial p_1}(p_1,p_2,p_3) & \frac{\partial s_2}{\partial p_1}(p_1,p_2,p_3) & 0 \\ \frac{\partial s_1}{\partial p_2}(p_1,p_2,p_3) & \frac{\partial s_2}{\partial p_2}(p_1,p_2,p_3) & 0 \\ 0 & 0 & \frac{\partial s_3}{\partial p_3}(p_1,p_2,p_3) \end{bmatrix} \begin{bmatrix} p_1 - mc_1 \\ p_2 - mc_2 \\ p_3 - mc_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

When setting the prices of its products, firm 1 takes p_3 as "given" and endogeneizes the impact of a change in p_1 on both s_1 and s_2. It does not take into account the impact of p_1 on s_3. When setting the price of product 3, firm 2 only solves one equation:

$$s_3 + \frac{\partial s_3}{\partial p_3}(p_3 - mc_3) = 0.$$

Table 16: Demand-side variables with different instruments: year 2011

	OLS	IV 1	IV 1	IV 2	IV 3	IV 4	IV 5
Mean utility							
Intercept	-5.137	-3.332	-4.184	-3.306	-3.004	-3.357	-2.597
	(0.037)	(0.187)	(0.117)	(0.190)	(0.685)	(0.196)	(0.320)
Price	-0.299	-1.723	-1.484	-1.742	-1.938	-1.703	-2.203
	(0.008)	(0.120)	(0.094)	(0.121)	(0.469)	(0.124)	(0.211)
Stops	-0.959	-1.397	-1.620	-1.369	-1.218	-1.320	-1.015
	(0.023)	(0.056)	(0.047)	(0.056)	(0.163)	(0.056)	(0.082)
OriginConn	2.206	3.721	3.154	3.807	4.379	3.878	5.152
	(0.062)	(0.289)	(0.238)	(0.292)	(1.024)	(0.296)	(0.478)
Distance	-0.189	-0.293	0.434	-0.291	-0.301	-0.307	-0.315
	(0.028)	(0.060)	(0.048)	(0.060)	(0.069)	(0.059)	(0.072)
Distance2	0.057	0.224		0.226	0.258	0.228	0.300
	(0.007)	(0.024)		(0.024)	(0.069)	(0.024)	(0.037)
Nesting	0.637	0.711	0.734	0.704	0.675	0.686	0.635
Parameter (λ)	(0.003)	(0.014)	(0.012)	(0.014)	(0.071)	(0.024)	(0.035)
Carrier FEs							
AA	0.305	0.356	0.313	0.358	0.383	0.364	0.416
	(0.024)	(0.047)	(0.043)	(0.048)	(0.063)	(0.047)	(0.058)
DL	0.125	0.150	0.144	0.154	0.156	0.154	0.159
	(0.020)	(0.039)	(0.036)	(0.040)	(0.043)	(0.039)	(0.048)
US	0.203	0.620	0.591	0.623	0.661	0.605	0.711
	(0.023)	(0.049)	(0.044)	(0.049)	(0.110)	(0.049)	(0.066)
WN	-0.342	-0.896	-0.802	-0.910	-1.030	-0.917	-1.190
	(0.021)	(0.073)	(0.063)	(0.074)	(0.246)	(0.075)	(0.118)
LCC	0.662	-0.030	0.040	-0.026	-0.058	0.020	-0.098
	(0.026)	(0.056)	(0.049)	(0.056)	(0.141)	(0.058)	(0.084)
Statistics							
J-statistic	N/A	4.433	8.8314	0.410	0.0379	45.480	0.359
Degree of overidentification	N/A	5	5	3	1	4	2
Ftest Price	N/A	54.8	66.8	76.7	105.0	84.7	93.5
Ftest Nest share	N/A	847.12	853.5	1098.8	938.1	505.9	715.5
Elasticities							
Own-price elasticity		4.17	3.510	4.24	4.86	4.22	5.78
95% CI		(3.56,4.80)	(3.07,3.98)	(3.63,4.87)	(2.19,8.42)	(3.45,4.99)	(4.41,7.39)
Con. semi-elasticity		0.615	0.656	0.608	0.573	0.597	0.517
95% CI		(0.584,0.639)	(0.634,0.674)	(0.577,0.634)	(0.472,0.611)	(0.567,0.623)	(0.456,0.558)
Scenario 1							
Δ Price		-6.73%	-6.64%	-6.65%	-5.55%	-7.01%	-4.33%

Table 17: Demand-side variables with different instruments: year 2016

	OLS	IV 1	IV 2	IV 3	IV 4	IV 4	IV 5
Mean utility							
Intercept	-4.605 (0.036)	-1.596 (0.250)	-1.559 (0.255)	-2.725 (0.259)	-2.690 (0.188)	-3.413 (0.126)	-2.692 (0.260)
Price	-0.421 (0.008)	-2.336 (0.154)	-2.360 (0.156)	-1.558 (0.156)	-1.582 (0.109)	-1.416 (0.092)	-1.580 (0.157)
Stops	-1.208 (0.024)	-0.505 (0.085)	-0.484 (0.086)	-1.858 (0.082)	-1.845 (0.063)	-2.072 (0.053)	-1.840 (0.082)
OriginConn	2.301 (0.058)	5.064 (0.286)	5.126 (0.288)	2.069 (0.273)	2.111 (0.200)	1.752 (0.176)	2.119 (0.273)
Distance	-0.197 (0.028)	-0.603 (0.073)	-0.613 (0.074)	-0.463 (0.058)	-0.463 (0.055)	0.244 (0.037)	-0.465 (0.059)
Distance2	0.060 (0.007)	0.348 (0.030)	0.354 (0.031)	0.199 (0.027)	0.201 (0.021)		0.202 (0.027)
Nesting Parameter (λ)	0.610 (0.003)	0.469 (0.028)	0.467 (0.028)	0.742 (0.037)	0.741 (0.028)	0.75995 (0.026)	0.740 (0.038)
Carrier FEs							
AA	0.001 (0.019)	0.243 (0.044)	0.241 (0.044)	0.257 (0.035)	0.261 (0.033)	0.253 (0.032)	0.259 (0.036)
DL	0.246 (0.020)	0.434 (0.045)	0.437 (0.045)	0.384 (0.035)	0.387 (0.035)	0.368 (0.033)	0.385 (0.036)
WN	-0.474 (0.021)	-1.200 (0.075)	-1.212 (0.076)	-0.724 (0.071)	-0.732 (0.055)	-0.698 (0.051)	-0.734 (0.071)
LCC	0.250 (0.026)	-1.168 (0.115)	-1.178 (0.117)	-1.051 (0.115)	-1.065 (0.086)	-0.973 (0.075)	-1.064 (0.115)
Statistics							
J-statistic	N/A	58.6	51.9	0.17	5.69	2.04	2.09
Degree of overidentification	N/A	5	3	1	4	4	2
First stage: Price	N/A	173.6	234.5	297.3	161.8	186.7	210.7
First stage: Nest share	N/A	1165.5	1565.7	884.3	469.7	459.5	692.6
Elasticities							
Own-price elasticity 95% CI	N/A	7.32 (5.87,9.06)	7.42 (5.96,9.13)	3.43 (2.58,4.40)	3.49 (2.88,4.20)	3.06 (2.53,3.61)	3.49 (2.56,4.47)
Con. semi-elasticity 95% CI	N/A	0.327 (0.240,0.397)	0.317 (0.220,0.388)	0.711 (0.685,0.731)	0.709 (0.688,0.729)	0.738 (0.722,0.753)	0.709 (0.680,0.729)

However, its market share, s_3, depends on the prices of all the products proposed in the market (here p_1, p_2 and p_3).

In Scenario 2, we simulate the case were both AA and US are merging. Imagine AA is firm 1 in the example above and US is firm 2, the new entity needs to take into account that, when it changes p_1, it impacts s_1, s_2, like before, but, also s_3. The new equilibrium prices satisfy the following system:

$$\begin{bmatrix} s_1(p_1,p_2,p_3) \\ s_2(p_1,p_2,p_3) \\ s_3(p_1,p_2,p_3) \end{bmatrix} + \begin{bmatrix} \frac{\partial s_1}{\partial p_1}(p_1,p_2,p_3) & \frac{\partial s_2}{\partial p_1}(p_1,p_2,p_3) & \frac{\partial s_3}{\partial p_1}(p_1,p_2,p_3) \\ \frac{\partial s_1}{\partial p_2}(p_1,p_2,p_3) & \frac{\partial s_2}{\partial p_2}(p_1,p_2,p_3) & \frac{\partial s_3}{\partial p_2}(p_1,p_2,p_3) \\ \frac{\partial s_1}{\partial p_3}(p_1,p_2,p_3) & \frac{\partial s_2}{\partial p_3}(p_1,p_2,p_3) & \frac{\partial s_3}{\partial p_3}(p_1,p_2,p_3) \end{bmatrix} \begin{bmatrix} p_1 - mc_1 \\ p_2 - mc_2 \\ p_3 - mc_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Therefore, for this counterfactual, we need to update the matrices Δ_t for all markets and solve for the new equilibrium prices using fixed-point algorithms.

C.2 Estimation procedure

We estimate the demand and supply-side jointly using the Generalized Method of Moments (see Hansen, 1982). Since we employ a nested logit model on the demand side, estimating both sides jointly comes at a negligible computational burden. We form moments that are interactions of the demand-and supply-side shocks with exogenous instruments introduced above.

On the demand side (see equation (6)), the linearity of the system allows us to express the unobservable ξ_{jt} as a function of the demand parameters θ_d and the explanatory variables. We then obtain the moment conditions by interacting the resulting demand-side unobservables with instruments:

$$E[z_{jt}^d \xi_{jt}] = 0,$$

where z_{jt}^d is a $k_1 \times 1$ vector of instruments. On the supply-side (see equation (9)), the supply-side unobservable ζ_{jt}, is the difference between the implied marginal costs from equation (8) and their deterministic part, $w_{jt}\theta_s$. We can then form moment conditions in the same way we did on the demand side:

$$E[z_{jt}^s \zeta_{jt}] = 0,$$

where z_{jt}^s is a $k_2 \times 1$ vector of instruments, in fact the product attributes.

We build sample analogs of the moment conditions by averaging first across products within a given market and then across markets:

$$\bar{g}(\theta) = \begin{pmatrix} \frac{1}{T} \sum_{t \in \mathcal{T}} \frac{1}{\mathcal{J}_t} \sum_{j=1}^{\mathcal{J}_t} z_{jt}^d \xi_{jt} \\ \frac{1}{T} \sum_{t \in \mathcal{T}} \frac{1}{\mathcal{J}_t} \sum_{j=1}^{\mathcal{J}_t} z_{jt}^s \zeta_{jt} \end{pmatrix},$$

with $\theta = (\theta_d, \theta_s)$. $\bar{g}(\theta)$ is a $(k_1 + k_2) \times 1$ vector of means. The GMM objective function to be minimized (with respect to θ) is a distance of $\bar{g}(\theta)$ to 0, i.e.:

$$f(\theta) = \bar{g}(\theta)^\top \Omega \bar{g}(\theta),$$

with Ω a positive definite weighting matrix.

We employ a two-step procedure in which we obtain a first set of estimates using an initial weighting matrix (the identity) before getting the final set of estimates using an estimate of the optimal GMM weighting matrix.