

Do Pre-Merger Concentration Screens Predict Realized Fare Increases? Evidence from Five U.S. Airline Mergers

Jose Aleman

University of Maryland, College Park

This version: September 15, 2026

Abstract

An antitrust agency must decide which markets a merger will harm before the merger closes, and the 2023 Merger Guidelines rank candidate markets by concentration. This paper asks whether that ranking identifies the markets where prices actually rose. Using the Department of Transportation's Origin and Destination Survey — 350 million ticket records over sixty quarters, 2005 to 2019 — I identify 8,009 overlap routes across five U.S. airline mergers, estimate realized fare effects with a stacked difference-in-differences design on matched control routes, and generate the predictions that an agency could have computed beforehand from a demand system and a Bertrand merger simulation built from pre-merger data alone. Fares on overlap routes rose 1.31 percent (95 percent interval 0.42 to 2.19), and 2.28 percent once the carriers flew under a single operating certificate; the effect does not appear until the fifth quarter after closing. The concentration presumption does not sort the routes: flagged routes saw smaller increases than unflagged ones, and the relationship between the change in the Herfindahl-Hirschman Index and the realized effect is indistinguishable from zero at every threshold tested. Diversion-based measures do sort them, with a one-standard-deviation increase in GUPPI worth 2.77 percentage points of realized fare increase. Both measures were computable before the mergers closed.

JEL classification: L41, L93, K21, L13, C23

Keywords: merger retrospective; horizontal merger guidelines; HHI; GUPPI; diversion ratio; stacked difference-in-differences; merger simulation; airline pricing

1 Introduction

Merger review is a prediction problem under a deadline. An agency must decide, before a transaction closes, which markets it will harm, and it must do so with data that predates the transaction. The 2023 Merger Guidelines resolve this with a structural presumption: compute market shares, form the Herfindahl-Hirschman Index before and after, and presume harm where the merger leaves the market concentrated and raises the index by more than one hundred points. The appeal of the screen is that it requires only quantities.

The alternative is a diversion-based measure. Estimate how customers substitute between the merging firms, compute the share of one firm's lost sales that the other recaptures, and value it. Gross upward pricing pressure (GUPPI) expresses that recapture as the implied increase in the firm's own opportunity cost of a sale. This is considerably more work and requires a demand system, which is why the structural screen remains the first cut in practice.

Both are predictions, and airline mergers offer an unusually clean setting in which to score them. The Department of Transportation publishes a ten percent sample of every domestic ticket sold, the same dataset

the agencies use for airline competition analysis. Five large mergers closed between 2008 and 2016, each leaving thousands of city pairs where both carriers had been selling and thousands more where only one had. Fifteen years of subsequent fares are on the record.

This paper scores the screens against the outcomes. It finds that fares did rise where the merging carriers overlapped, by 1.31 percent on average against matched controls, but that the effect takes five quarters to appear and that the concentration screen has no power to rank the routes: by decile of the change in the index the realized effect is flat, and flagged routes saw marginally smaller increases than unflagged ones. Measures built on diversion — GUPPI, the diversion ratio to the merger partner, and the simulated fare change — all rank the routes correctly, using only pre-merger data.

The paper also documents, at some length, three specifications that produced the wrong answer before the reported one. A naive two-way fixed effects regression on the pooled sample returns a fare decrease. An intuitive but incorrect control group reverses the sign. And inverse-odds weighting, the textbook reweighting estimator, collapsed one merger’s control arm to an effective sample size of 3.4 routes. These are reported in Section 7 because each is a plausible first choice, and the difference between them and the reported specification is larger than the effect being measured.

2 The screens

For an airline merger, the markets at risk are overlap routes: city pairs on which both merging carriers were already selling tickets, so that the transaction removes a competitor. I define an overlap route as a city pair on which each merging carrier held at least one percent of passengers, averaged over the four quarters preceding the merger’s announcement rather than its closing. The announcement baseline matters because carriers reallocate capacity once a deal is public; a baseline measured just before closing would already embed the merger’s anticipated effects. Five mergers yield 8,009 overlap routes.

The Herfindahl-Hirschman Index is the sum of squared market shares on a zero to ten thousand scale; a monopoly is 10,000 and four equal firms 2,500. A merger raises it by exactly twice the product of the merging firms’ shares. Applied to these routes, the Guidelines presumption flags 74 percent of them.

The diversion ratio is the share of customers who, on abandoning carrier A because its fare rose, buy from carrier B. It is not observed and must come from an estimated demand system. GUPPI is the value of the sales A diverts to B when A raises its fare, expressed as a fraction of A’s fare; a GUPPI of five percent means the merger gives A the pricing incentive of a five percent increase in marginal cost.

3 Data

Table 3.1. Sample construction

Item	Value
Source	US DOT Bureau of Transportation Statistics, Airline Origin and Destination Survey (DB1B), Market file
Window	60 quarters, 2005Q1–2019Q4
Volume	350 million ticket records, 5.3 GB
After filtering	19,189 routes; 696,963 route-quarters; 2.3 million carrier-route-quarters
Overlap routes	8,009 across five mergers

Item	Value
Deflator	CPI-U from FRED; fares in constant 2019 dollars
Cost instrument	US Gulf Coast jet fuel spot price, from FRED
Mergers	Delta/Northwest (2008Q4); United/Continental (2010Q4); Southwest/AirTran (2011Q2); American/US Airways (2013Q4); Alaska/Virgin America (2016Q4)

Notes: DB1B is a ten percent sample of domestic itineraries; each record carries origin, destination, ticketing carrier, fare and passenger count. The window ends at 2019Q4 deliberately: the pandemic makes 2020 fares uninterpretable as a control period, and the last merger requires twelve quarters of follow-up. Merger dates are not taken on trust — a validation stage requires each acquired carrier to stop appearing as a ticketing carrier when its brand is retired, and each does.

Because each quarterly file unpacks to roughly six million ticket records, the pull aggregates each quarter on arrival and deletes the raw file, recording the source URL, byte count and SHA-256 hash in a manifest. The full download takes about forty minutes; everything downstream of it runs in minutes.

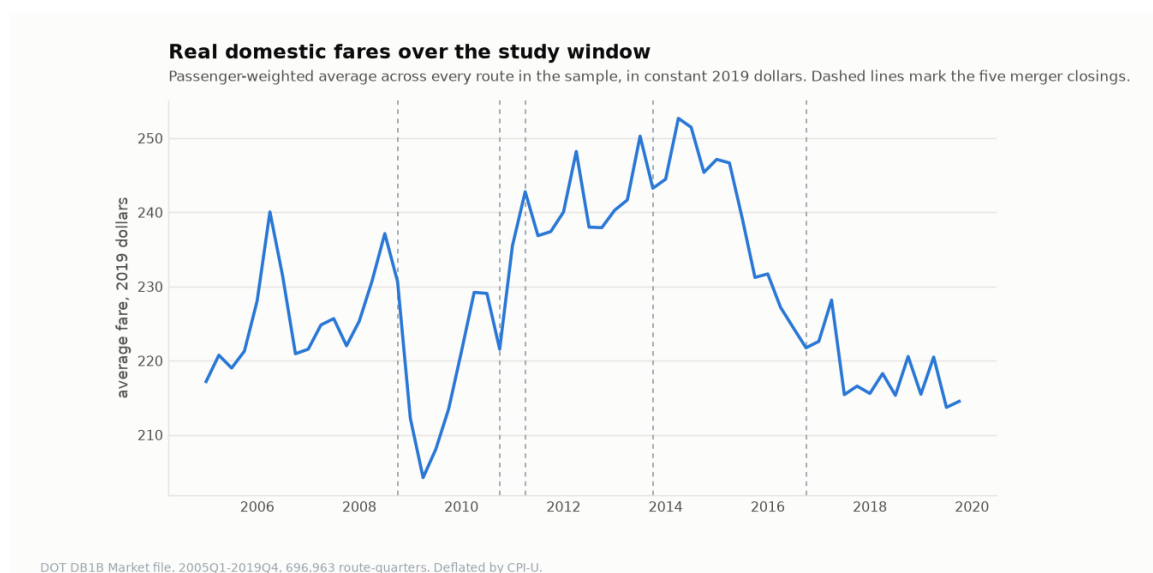


Figure 3.1. Real average fares on overlap and control routes, 2005–2019

Notes: Fares in constant 2019 dollars, deflated by CPI-U. The series are plotted on matched samples; the identifying variation is the change in the gap around each merger’s closing quarter, not the level.

4 Empirical strategy

4.1 Control group

The control group is routes on which exactly one of the two merging carriers was selling tickets. The more intuitive comparison — routes where neither carrier flew — is the wrong one, and Section 7 shows how wrong. A route served by both American and US Airways is long, busy and concentrated by construction; a route served by neither is short, thin and often flown by a single low-cost carrier. Conditioning on exactly one merging carrier holds constant whatever the merger did to the combined airline’s network and cost structure, and isolates the loss of a competitor, which occurs only on overlap routes.

Treated routes are matched to controls on log distance, log passengers, the pre-merger index, carrier count, log fare, nonstop share and the pre-merger fare trend, by nearest-neighbour matching on the propensity score. Treated routes with no comparable control are dropped and counted rather than matched to something unlike them; a unit test enforces this.

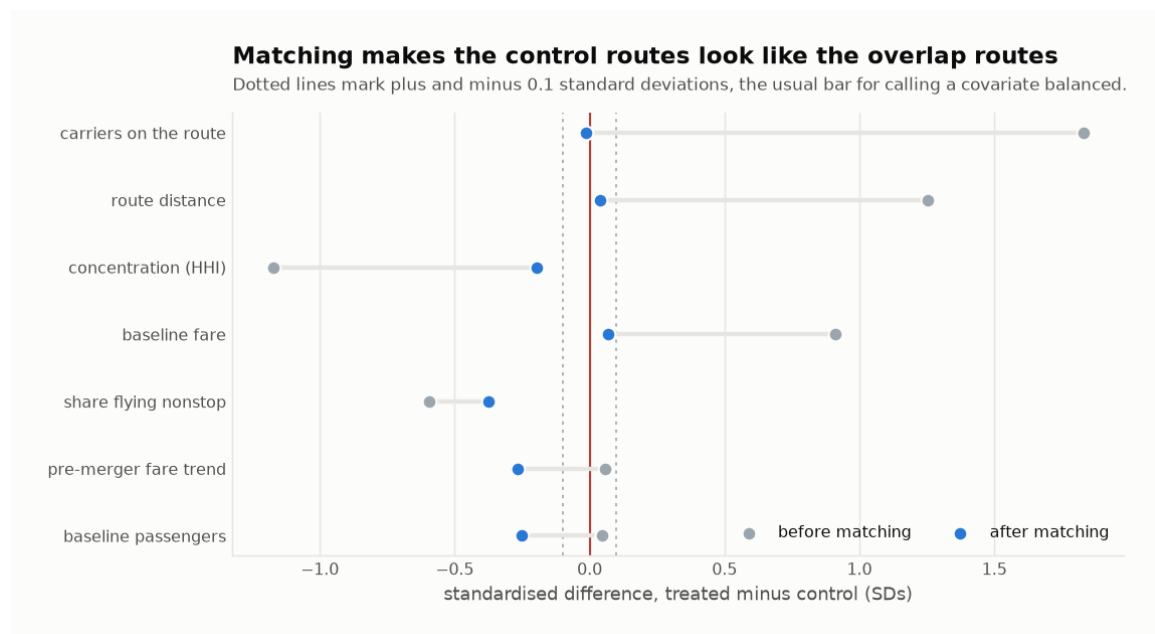


Figure 3.2. Covariate balance before and after matching

Notes: Standardized differences between treated and control routes. Carrier count and log distance move from standardized differences of 1.83 and 1.25 before matching to -0.01 and 0.04 after. Nonstop share and the pre-merger trend remain the least well balanced covariates.

4.2 Stacked difference-in-differences

Each merger is estimated in its own block: its overlap routes, its matched controls, and its own twenty-one quarter window, with route-by-block and quarter-by-block fixed effects. No comparison crosses mergers, which by construction eliminates the contamination that arises when already-treated units serve as controls for later-treated ones. The estimation sample is 247,050 route-quarters on 7,556 routes.

4.3 Inference

Standard errors are clustered on the route, since fares on a route are serially correlated. Because the number of clusters is large but the treatment is assigned at the route level, I also report a wild cluster bootstrap with 299 draws as a check on the asymptotics. Throughout, estimates whose interval covers zero are shown in gray rather than reported as effects.

5 Demand estimation and simulation

5.1 The instrument audit

GUPPI requires a demand system, and a demand system requires an instrument for the fare, because airlines charge more on routes where demand is strong and the raw price-quantity relationship therefore understates

price sensitivity. I constructed five instruments and tested each separately rather than pooling them and reading an overidentification statistic.

Three of the five return a price coefficient of the wrong sign, implying that passengers buy more as fares rise. The reason is common to all three: entry responds to demand. A carrier adds a route when it expects the route to sell, so a positive demand shock draws in rivals, rivals push the fare down, and any instrument constructed from rival counts carries the shock it was meant to exclude. Their first-stage F statistics run into the hundreds and thousands, so conventional strength diagnostics would not have caught the problem.

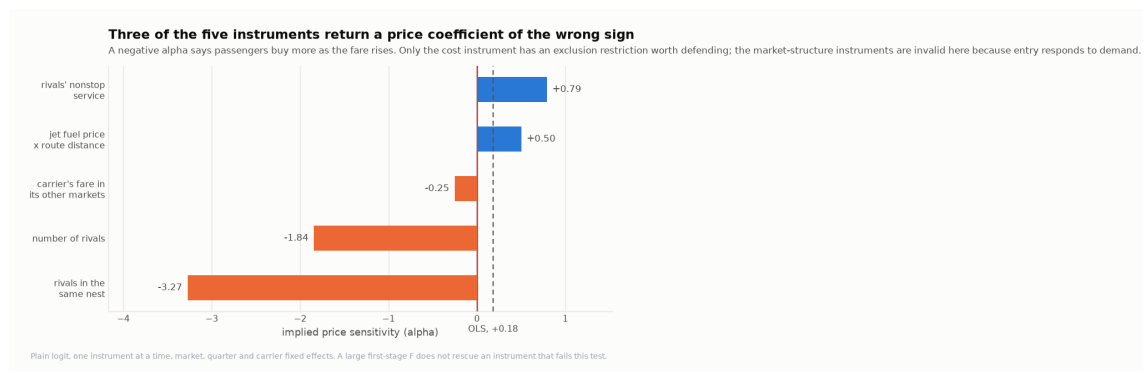


Figure 3.3. Instrument audit: first-stage strength against the sign of the price coefficient

Notes: Each instrument is estimated separately. Instruments built on rival presence are strong in the first stage and return a positive price coefficient. Only jet fuel price interacted with route distance — a cost shifter with no demand channel — returns the expected sign, and it is much the weakest first stage ($F = 23$, against 707 for the next weakest).

5.2 Calibration

The surviving instrument gives the right sign and a coefficient larger in magnitude than OLS, as the endogeneity argument predicts, but it implies a median own-price elasticity of -1.25 . A profit-maximizing firm does not price on the inelastic portion of its own demand: at that elasticity the Bertrand first-order condition returns a negative marginal cost. The estimate is therefore not usable for simulation.

The simulation consequently runs on a calibration rather than an estimate. I set the price coefficient to deliver a median own-price elasticity of -2.5 , within the range of published estimates for U.S. airline demand, and re-run everything at -1.5 and -4.0 . This is labelled a calibration wherever it appears, and the dollar figures the simulation produces should be read as scaled to that assumption.

The calibration matters less for the paper's claim than it might. Under logit demand, diversion ratios are functions of market shares and not of the price coefficient, so the level of every GUPPI scales with the calibration while the ordering of routes barely moves: the rank correlation between the reported run and every alternative exceeds 0.92, and exceeds 0.999 across the elasticity range. The paper's finding is about ordering.

5.3 Simulation

Marginal costs are recovered from the pre-merger first-order conditions, the post-merger Bertrand equilibrium is solved route by route, and GUPPI, upward pricing pressure, compensating marginal cost reduction and consumer surplus are computed from it. The solver converges on 7,932 of the 8,009 overlap routes.

6 Results

6.1 Realized fare effects

Table 3.2. Realized fare effects on overlap routes

Specification	Effect	95% interval	Observations
Stacked DiD, matched controls	+1.31%	[+0.42, +2.19]	247,050
Quarters +5 to +12 only	+2.28%	[+1.30, +3.26]	—
Weighted by passenger traffic	+1.75%	—	247,050
Delta / Northwest only	−1.74%	[−3.09, −0.39]	—

Notes: Effects are log points on real average fare, relative to matched control routes on which exactly one merging carrier was selling. Intervals are clustered on the route. The estimation sample covers 7,556 routes across five merger blocks. Quarters +5 to +12 restricts to the period after the carriers combined operating certificates.

The headline effect is +1.31 percent, and +2.28 percent once attention is restricted to the quarters in which the carriers actually operated as one firm. Weighting by traffic raises the estimate to +1.75 percent, so the average passenger experienced a larger increase than the average route.

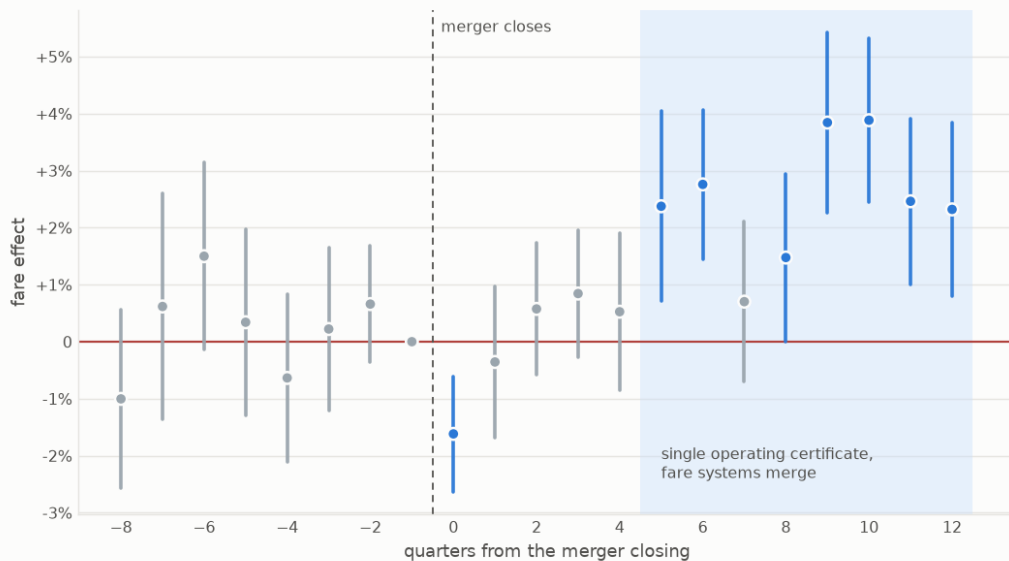
6.2 Timing

The effect does not appear immediately. Every event-study coefficient from the closing quarter through quarter +4 lies between −1.6 and +0.8 percent, and none is distinguishable from zero. The effect emerges at quarter +5 and persists.

This is not a statistical artifact. Two carriers that have legally merged continue to fly under separate operating certificates and to file separate fares for four to six quarters after closing. Until the certificates combine there is no single firm setting one price, and no mechanism by which the merger could raise fares. The practical implication is that a retrospective using a one-year post-merger window would find nothing here, and would report that finding as evidence of no effect.

Fares on overlap routes rise, but only after the airlines actually merge

Effect on the average fare of an overlap route relative to matched control routes, by quarter from closing. Grey where the 95% interval covers zero. The quarter before closing is the omitted baseline.



DOT DB1B, 2005-2019. Stacked difference-in-differences, 7,556 routes, standard errors clustered on the route. Average over quarters +5 to +12: +2.28%.

Figure 3.4. Event study: fare effect by quarter relative to merger closing

Notes: Coefficients on quarter-relative-to-closing indicators with 95 percent intervals clustered on the route. Pre-merger coefficients are individually indistinguishable from zero and the largest is 1.5 percent, but a joint test across them rejects at $p < 0.001$ ($F = 5.62$). With 6,400 clusters that test detects very small deviations; the coefficients scatter rather than trend, which is the desired pattern, but the design is not perfectly clean and the pre-trend test is reported as a failure rather than omitted.

6.3 The concentration screen does not sort the routes

Table 3.3. Realized fare effects by whether the Guidelines presumption flagged the route

	Routes	Average fare effect	95% interval
Flagged by the presumption	5,864	+2.01%	—
Not flagged	2,072	+2.72%	—
Difference		-0.71%	[-1.41, -0.01]

Notes: The presumption flags a route where the merger leaves it concentrated and raises the Herfindahl-Hirschman Index by more than 100 points, computed from pre-announcement passenger shares. Routes flagged by the screen recorded smaller fare increases than routes it did not flag. The interval on the difference barely excludes zero, so the sign should not be leaned on; the relevant point is the absence of positive sorting.

By decile of the increase in the index the profile is flat: +2.31 percent in the bottom decile against +1.79 percent in the top. Moving the threshold does not rescue the screen. At every cut-off tested, from 50 points to 2,000, the flagged group records the smaller average increase.

6.4 Diversion-based measures do sort them

Table 3.4. Realized fare effect per standard deviation of each pre-merger prediction

Prediction	Effect per SD	Standard error	Rank correlation
GUPPI	+2.77 pp	(0.22)	+0.189
Diversion to the merger partner	+2.26 pp	(0.19)	+0.118
Simulated fare change	+2.20 pp	(0.22)	+0.116
Increase in HHI	+0.24 pp	(0.20)	−0.017
Combined share of the two carriers	−0.59 pp	(0.19)	−0.060

Notes: Same 7,936 routes and the same realized outcomes in every row; only the ranking variable changes. The dependent variable is the realized fare effect and the regressor is the standardized prediction, so the coefficient is percentage points of realized fare increase per standard deviation of the prediction. Rank correlation is Spearman's rho against the realized effect. The coefficient on the increase in HHI is not distinguishable from zero ($p = 0.25$); combined share is significantly negative. Every prediction uses pre-merger data only.

A one-standard-deviation increase in GUPPI corresponds to 2.77 percentage points of realized fare increase. For the increase in the concentration index the same figure is 0.24 points with a standard error of 0.20, which does not reject zero. The combined share of the two carriers — the most immediately available structural measure of all — points the wrong way.

The information needed to rank these routes correctly existed before the transactions closed. The distinction is not between what was knowable ex ante and what was only visible ex post; it is between two ex ante measures, one of which carries information about substitution and one of which does not.

Do routes the tools rank as worse actually see bigger fare increases?

Overlap routes sorted into deciles by each prediction, with the average realized fare effect in each. Grey where the interval covers zero.

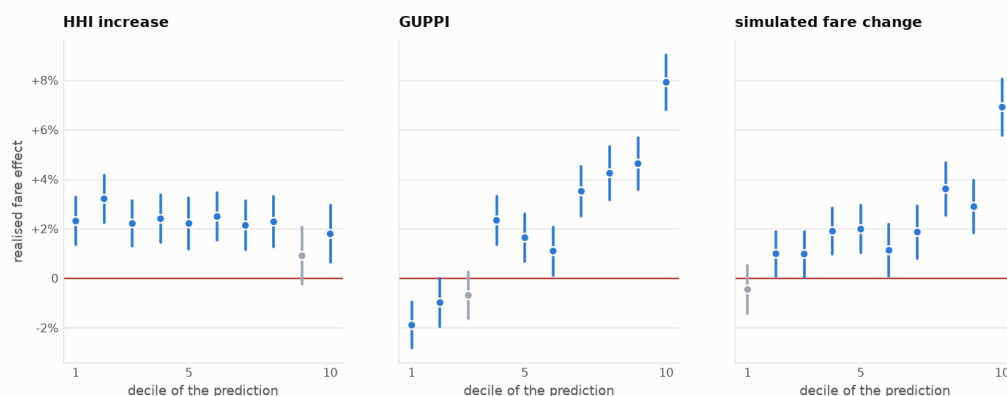


Figure 3.5. Realized fare effect by decile of each pre-merger prediction

Notes: Routes sorted into deciles by each prediction, with the mean realized fare effect plotted for each decile. Sorted by the increase in concentration, realized effects are flat across the whole range. Sorted by the value of sales diverted to the merger partner, they climb from −1.9 to +7.9 percent.

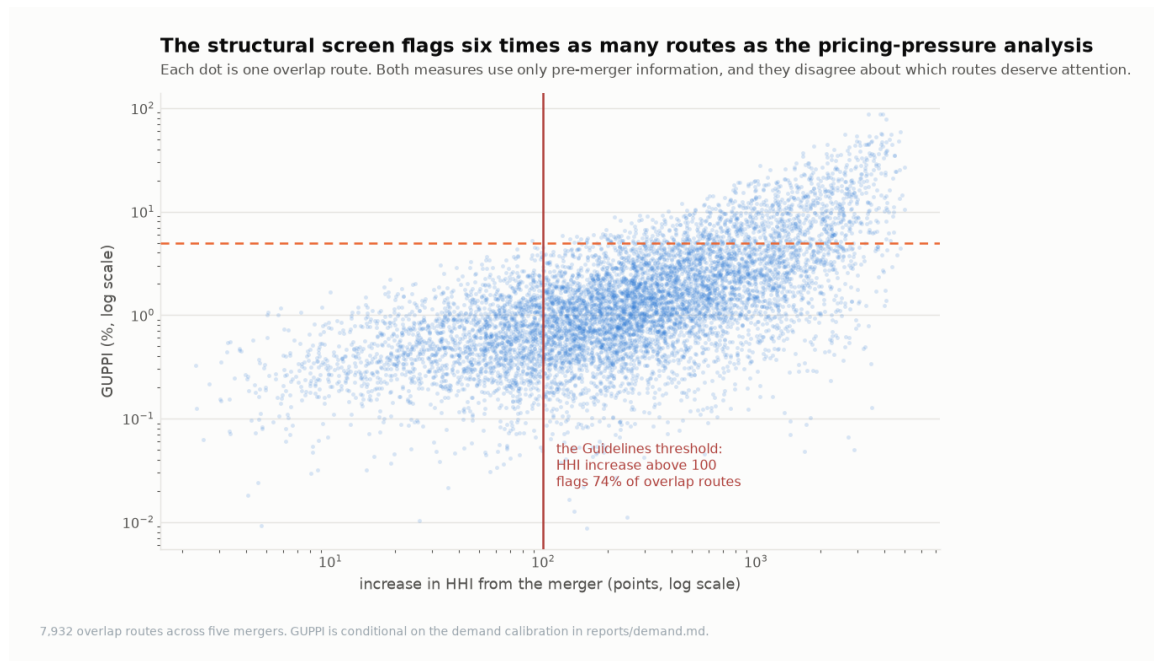


Figure 3.6. The structural screen against the diversion-based measure, route by route

Notes: Each point is an overlap route. The two measures are close to orthogonal, so the screen and GUPPI select substantially different sets of routes for investigation.

6.5 The simulation ranks well and scales badly

Regressing realized effects on simulated ones gives a slope of 0.57 with a 95 percent interval of 0.46 to 0.68. That interval rejects both relevant nulls: the simulation is not uninformative, and it is not correctly scaled. It overstates the magnitude of the effect by roughly a factor of 1.8 while ordering routes correctly. This is the pattern a calibrated price coefficient would be expected to produce, and it is the reason the paper's claims are stated in terms of ranking rather than in dollars.

6.6 Heterogeneity across mergers

Four of the five mergers raised fares on overlap routes. Delta/Northwest lowered them by 1.74 percent, with an interval from -3.09 to -0.39 . It is the earliest deal in the sample and the one with the most overlap routes, at 3,612. Alaska/Virgin America contributes only nineteen simulated overlap routes meeting the sample requirements, and its interval is correspondingly uninformative.

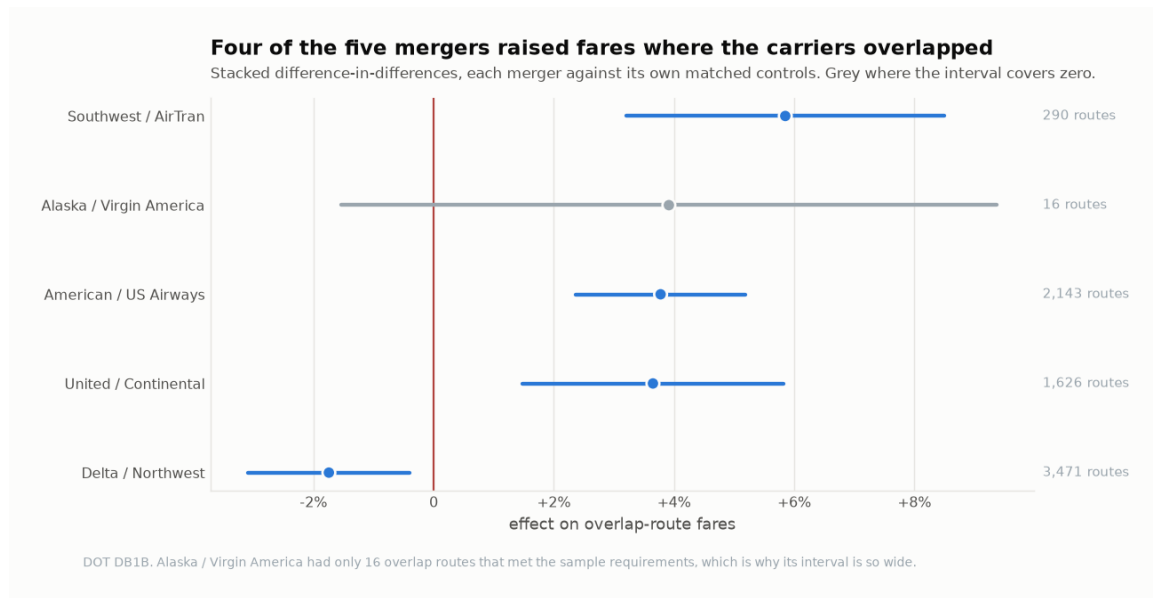


Figure 3.7. Realized fare effect on overlap routes, by merger

Notes: Estimates with 95 percent intervals clustered on the route. Delta/Northwest is the one negative estimate whose interval excludes zero. Alaska/Virgin America has too few qualifying overlap routes to be informative.

7 Three specifications that gave the wrong answer

7.1 Naive two-way fixed effects

Pooling all five mergers into a single regression with route and quarter fixed effects returns -2.41 percent — a fare decrease, significant at conventional levels. Two problems combine to produce it. The first is the control group, discussed below. The second is staggered treatment timing.

Goodman-Bacon (2021) established that a two-way fixed effects coefficient under staggered adoption is a weighted average of every two-by-two comparison available in the data, and that some of those comparisons use already-treated units as controls for later-treated ones. Applied here, 14.3 percent of the weight sits on comparisons that use routes already affected by an earlier merger as the counterfactual for a later one, and those comparisons carry a coefficient of -0.72 percent against $+2.41$ percent on the clean treated-versus-never comparisons. Because merger effects grow over the first two years, the forbidden comparisons subtract one merger's ongoing increase from another merger's estimate. The implementation is self-checking: the weighted components must reproduce the regression coefficient exactly or the routine raises an exception.

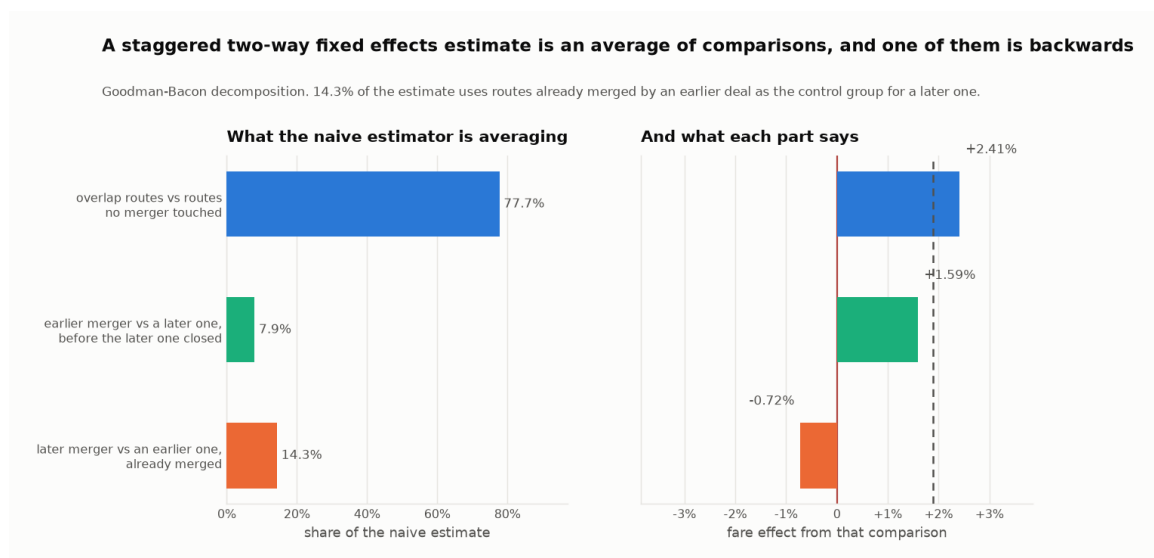


Figure 3.8. Goodman-Bacon decomposition of the pooled two-way fixed effects estimate

Notes: Each point is a two-by-two comparison, positioned by its weight and its coefficient. Comparisons using later-treated routes against earlier-treated ones carry 14.3 percent of the total weight and a negative coefficient. The stacked design used in the paper eliminates these comparisons by construction.

7.2 The wrong control group

My first control group was routes on which neither merging carrier flew. It is the cleaner-looking comparison and the incorrect one, for the reason given in Section 4.1. Table 3.5 shows that this single choice, interacted with matching, moves the answer more than anything else in the paper.

Table 3.5. Estimated fare effect by control group and matching

Control group	Without matching	With matching
Routes where neither carrier flew	-2.63% [-4.11, -1.14]	+1.11% [-1.33, +3.55]
Routes where exactly one carrier flew	-0.11% [-1.26, +1.05]	+1.31% [+0.42, +2.19]

Notes: The top-left cell is what an unreflective specification would have reported: a statistically significant fall in fares following merger. Matching brings the two control groups into agreement on sign, but matching onto routes that neither carrier served leaves few comparable controls and yields an interval more than twice as wide. The bottom-right cell is the specification reported throughout the paper.

7.3 Inverse-odds weighting collapsed the control group

Weighting control units by the odds of treatment is the textbook reweighting approach and it failed here in a way that the standard output did not reveal. It produced confidence intervals of roughly plus or minus eight percentage points on tens of thousands of observations. The cause was a small number of control routes with propensity scores near one, whose odds ran into the tens; after multiplying by passenger traffic, one merger's control arm had an effective sample size of 3.4 routes, with ten routes carrying 97 percent of the weight.

Nearest-neighbour matching bounds this by construction, since a control route can receive weight only from the treated routes it stands in for. Effective control counts moved from single digits into the hundreds, and intervals contracted to roughly plus or minus one percentage point. A unit test encodes the failure directly: it

constructs poor overlap between treated and control units and asserts that inverse-odds weighting degenerates while matching does not.

8 Robustness

Table 3.6. Robustness checks

Check	Question	Result
Placebo on one-carrier routes	Is the design capturing something other than lost competition?	−0.04% [−0.98, +0.90]
Wild cluster bootstrap, 299 draws	Does the interval depend on large-cluster asymptotics?	p = 0.003
Weighted by traffic	Average route, or average passenger?	+1.75%
Matching switched off	How much work is matching doing?	−0.11%
Quarters +5 to +12 only	What happens once the carriers are one firm?	+2.28% [+1.30, +3.26]

Notes: The placebo compares routes on which exactly one merging carrier flew against routes on which neither did. Neither group loses a competitor, so a nonzero effect there would indicate that the design was picking up a network or cost consequence of the merger rather than a competitive one. It does not.

The estimators themselves are validated against implementations written by others. Coefficients and cluster-robust standard errors match statsmodels to eight decimal places, weighted and unweighted, one-way and two-way; the instrumental-variables estimator matches linearmodels to the same tolerance. Both packages are test dependencies only and are never imported by the pipeline. The demand and simulation code is checked against closed forms: at a nesting parameter of zero the nested logit must reproduce plain logit diversion ratios, own-price elasticities and single-product Lerner margins exactly; re-solving the pre-merger equilibrium must return the observed prices; a merger must raise the merging carriers' prices, rivals must respond upward by less, and a merger to monopoly must do more than a partial one. Eighty-three tests run in continuous integration on every commit.

9 Limitations

This is one industry and five transactions. Airlines have features that may not transfer: hub economics, connecting itineraries, and frequent-flyer lock-in all affect substitution in ways that have no analogue in most merger review. The finding concerns airline routes with established overlap, not concentration screens in general.

The demand system is calibrated rather than estimated, for the reason given in Section 5.2. The ordering of routes survives that choice; the dollar magnitudes do not, and are reported as scaled to an assumption.

Most importantly, the paper compares a screen against an outcome, not against the decision an agency should have made. A screen that flags too many routes costs investigative resources; a screen that misses routes costs consumers. Which error is worse is a question about the agency's loss function, and this data does not answer it.

10 Conclusion

Fares rose on the routes where two merging airlines had both been selling: 1.31 percent against matched controls, and 2.28 percent once the carriers combined operating certificates. The effect takes five quarters to

appear, for the institutional reason that the merged firm does not set a single price until then, which is long enough that a short-window retrospective would miss it entirely.

The concentration presumption flagged 74 percent of these routes and did not distinguish among them. Routes it flagged recorded slightly smaller fare increases than routes it did not, the relationship between the change in concentration and the realized effect is flat across deciles, and no threshold between 50 and 2,000 points reverses that. Measures built on estimated substitution — GUPPI above all — rank the routes correctly, at 2.77 percentage points of realized increase per standard deviation. Both kinds of measure were computable from pre-merger data. The difference between them is not information availability but what the measure is about.

References

- Abadie, Alberto, and Guido W. Imbens. 2006. “Large Sample Properties of Matching Estimators for Average Treatment Effects.” *Econometrica* 74 (1): 235–267.
- Ashenfelter, Orley C., Daniel Hosken, and Matthew Weinberg. 2014. “Did Robert Bork Understate the Competitive Impact of Mergers? Evidence from Consummated Mergers.” *Journal of Law and Economics* 57 (S3): S67–S100.
- Berry, Steven T. 1994. “Estimating Discrete-Choice Models of Product Differentiation.” *RAND Journal of Economics* 25 (2): 242–262.
- Borenstein, Severin. 1989. “Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry.” *RAND Journal of Economics* 20 (3): 344–365.
- Callaway, Brantly, and Pedro H. C. Sant’Anna. 2021. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics* 225 (2): 200–230.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics* 90 (3): 414–427.
- Farrell, Joseph, and Carl Shapiro. 2010. “Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition.” *B.E. Journal of Theoretical Economics* 10 (1), Article 9.
- Goodman-Bacon, Andrew. 2021. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics* 225 (2): 254–277.
- Kim, E. Han, and Vijay Singal. 1993. “Mergers and Market Power: Evidence from the Airline Industry.” *American Economic Review* 83 (3): 549–569.
- U.S. Department of Justice and Federal Trade Commission. 2023. *Merger Guidelines*. Washington, DC.
- Werden, Gregory J., and Luke M. Froeb. 1994. “The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy.” *Journal of Law, Economics, and Organization* 10 (2): 407–426.

Data and code availability

All data are public and require no credentials. Fares are from the U.S. Department of Transportation, Bureau of Transportation Statistics, Airline Origin and Destination Survey (DB1B), Market file, 2005Q1–2019Q4; CPI-U and U.S. Gulf Coast jet fuel spot prices are from the Federal Reserve Bank of St. Louis (FRED). The twelve-stage pipeline, the econometric routines written from NumPy and SciPy rather than taken from a package, the eighty-three tests, every intermediate analysis output and a two-page memo are available at github.com/josealemanm/us-airline-merger-price-effects. An interactive dashboard runs in the browser at josealemanm.github.io/us-airline-merger-price-effects/dashboard.html. The two large panels and the raw quarterly files are not committed because stage four rebuilds them, but every analysis output behind the numbers reported here is, so the results can be checked without repeating the download.