An Empirical Analysis of Merger Efficiencies*

Alon Eizenberg† Omri Zvuluni‡

June 2024

First public draft: January 2024

Abstract

We develop an econometric method to study merger efficiencies. Classification techniques are employed first to determine the sign of the merger's effect on output levels in specific markets. These classifications are then combined with familiar oligopoly theory results to yield bounds on marginal cost savings. Applying this framework to the 2013 merger of US Airways and American Airlines we find that the merger led to output expansions in more than half of the markets where the sign of the output effect could be determined, and in at least 44% of the total number of markets that were directly affected by the merger and where the market structure was otherwise stable. Pro-competitive effects were more prevalent in larger markets and, to some extent, in markets that serve the merging carriers' hubs. Averaging across the markets experiencing output expansions, the lower bound on the marginal cost reduction was slightly above 2 USD, capturing 0.8% of the market price. The analysis provides insights regarding the nature and magnitude of merger efficiencies.

*We are grateful to workshop participants at Ben Gurion University, Boston College, Hebrew-U, Toulouse, University of Virginia, and the University of Vienna, and to conference participants at IIOC 2024 (Boston) for helpful comments. We are also grateful to Simon Anderson, Heski Bar Isaac, Effi Benmelech, Steve Berry, Tomer Blumkin, Federico Ciliberto, Pierre Dubois, Naomi Feldman, Andrey Fradkin, David Genesove, Klaus Gugler, John Mayo, Charlie Murry, Daniel Silva, Phil Haile, Ariel Pakes, Robert Parnham and Ro'i Zultan for helpful conversations and feedback. All errors are our own. Yotam Nir provided outstanding research assistance. This research is supported by the Israel Science Foundation (grant No. 2512/21).

†Department of Economics, Hebrew University of Jerusalem, and CEPR. email: alon.eizenberg@mail.huji.ac.il

‡Department of Economics, Hebrew University of Jerusalem. email: omri.zvuluni@mail.huji.ac.il

1 Introduction

Horizontal mergers reduce the number of competitors but may also result in efficiency gains (Williamson (1968)). While the magnitude of merger efficiencies is of considerable interest, their empirical assessment remains a difficult challenge, owing in part to the fact that economic costs are not directly observed.

This scarcity of evidence has implications for antitrust policies. Rose and Shapiro (2022) argue that agencies' skepticism with regard to efficiencies is "...warranted because there is no robust body of empirical evidence showing that most mergers realize cognizable efficiencies." Skepticism is also echoed from a theoretical perspective. While merger proposals typically promise various forms of costs savings, a merger could also result in dis-economies associated with the challenge of managing a larger operation or with business culture clashes (White (1987)). On the flip side, efficiencies may be under-estimated as they are realized via subtle and complicated channels (Bork (1978)).[1]

Typical merger studies in the industrial organization literature do not, typically, offer direct evidence on the nature and magnitude of efficiencies. Merger simulations are performed *ex-ante* given estimates of demand elasticities and an assumed, rather than estimated, degree of marginal cost reduction.[2] *Ex-post* analyses offer an in-sample estimation of the effect of the merger on outcomes such as prices. The estimated effect lumps the pro-competitive and the anti-competitive effects together, revealing which of the effects prevailed, but not their individual magnitudes.[3]

In light of this scarcity of evidence we propose an econometric method for estimating merger efficiencies. The method is appropriate for industries featuring a homogeneous good and requires the econometrician to observe a cross-section of markets affected by the merger, along with some post-merger outcomes. It delivers a classification of markets into those where the merger prompted output expansions and output contractions, respectively, as well as bounds on marginal cost savings. We therefore depart considerably from typical *ex-post* merger studies.[4]

In a first step, we classify the markets affected by the merger according to the sign of the merger's effect on output. Namely, we determine which of the affected markets experienced an increase, or a decrease, in output as a result of the merger. This classification is performed via a comparison of each affected market to various control markets. In a second step, we combine these classifications with structural assumptions on demand and supply to estimate simple bounds on the magnitude of efficiency gains in the affected markets. This step builds on familiar oligopoly theory results derived by Farrell and Shapiro (1990) for the case where the industry features a

[1] Bork notes that "(e)conomists, like other people, will measure what is susceptible to measurement and will tend to forget what is not, though what is forgotten may be far more important than what is measured." In Ashenfelter et al. (2014).

[2] See, for example, Björnerstedt and Verboven (2016).

[3] Some studies do, however, estimate merger efficiencies. Chen (2021) studies merger efficiency gains in U.S. freight railroads. Demirer and Karaduman (2022) study efficiency gains from mergers in the electricity generation industry using granular data on physical output and input levels. See also Bokhari (2023) and Khmelnitskaya et al. (2024). Yet another approach estimates the cost efficiencies that would compensate for the anti-competitive effect as in Affeldt et al. (2021).

[4] Retrospective merger evaluation has attracted considerable attention from both practitioners and scholars (see, for example, Ashenfelter et al. (2014)). Gathering evidence on realized efficiencies should be of similar value.

Cournot equilibrium before, and after the merger. Of note, some of our analysis does not depend on the Cournot structure, and, moreover, some of our bounds are robust to the possible presence of coordinated merger effects.

Intuitively, where the pro-competitive effect prevailed and output has increased, the efficiency could not have been "too low," revealing a lower bound on it. Where the anti-competitive effect prevailed, an upper bound on the efficiency presents itself instead. The econometric structure allows us to translate this intuition into a single bound on marginal cost savings in each market. Importantly, the bounds are simple functions of observed or easily-estimated pre-merger quantities: prices, market shares, and the market demand elasticity.

These two methodological steps — classifying markets according to the sign of the merger's output effect, and obtaining a bound on the efficiency gain — reveal a rich, albeit incomplete picture of the nature and magnitude of efficiencies. The first step uncovers the distribution of the sign of the merger's output effect across the affected markets. It allows us to get a sense of *where* merger efficiencies were realized, and to examine whether this distribution is aligned with economic reasoning. This analysis therefore speaks to the *nature* of efficiencies. The second step results in partial identification of the *magnitude* of efficiencies. We apply this methodology to the airline industry and, specifically, to the 2013 merger between American Airlines and US Airways.

Methodology. We observe a cross-section of local markets for a homogeneous good. Prior to the merger, the merging firms compete in some of these markets, but not in others. We refer to the former group of markets as the "affected markets" as those are the markets that are most directly affected by the merger.[5] In all markets we observe the pre-merger outputs of the merging parties, the pre-merger price, and the post-merger aggregate output. Also observed are cost shifters, used as instruments when estimating the market demand curve.

The first step of the analysis determines the sign of the merger's effect on output in each affected market. This classification is not trivial because post-merger quantities may be affected not only by the merger but also by changes in input prices or demand. This is the familiar hurdle in *ex-post* merger studies and it is routinely addressed via a difference-in-differences approach that yields an estimated Average Treatment Effect of the merger across markets.

Our approach builds on this logic but results in a market-specific classification of the merger's effect. For each affected market we identify a control market using two familiar methods: the Nearest Neighbor method (Fix and Hodges (1951)) and the Synthetic Control method (Abadie (2021)). Both methods construct a control market that mimics the pre-merger output behavior of the affected market as closely as possible. We then use the post-merger output of the control market to predict the affected market's post-merger output in a no-merger counterfactual. We thus classify the affected market as having experienced an output expansion (resp., output contraction) if its post-merger output exceeds (resp., falls short of) the post-merger output in the control market.[6]

[5]These are not, however, the only markets affected by the merger. For example, the merger may also have an impact on markets where only one of the parties was present via changes to that competitor's cost structure. Our bounds, nonetheless, can only be computed in markets where both merging firms were originally present.

[6]In one of our robustness checks we compare the *changes* in output across these two markets. This approach

The classification is, in essence, a prediction exercise. In the standard difference-in-differences approach the researcher uses the sample to test hypotheses regarding *the mean effect* of the merger across markets. We, in contrast, are interested in market-specific conclusions. Since a statistical test is not available at the market level, we employ a range of classification exercises, and refrain from classifying the output effect in a given market if the exercises do not point in a clear direction. Namely, we perform the classification exercise using different pools of candidate control markets, and different post-merger time periods. We use various majority rules such that some consensus across different exercises is necessary for classification. Where a majority of exercises do not point in the same direction, we classify the output effect in that market as "undetermined."

In a second step, a bound on the magnitude of the efficiency gain is computed in each affected market that was successfully classified in the first step. The bounding strategy rests on the notion that the sign of the merger's output effect, while confounding pro-competitive effects together with anti-competitive ones, does contain information regarding the former. This notion builds heavily on theoretical results from Farrell and Shapiro (1990) who analyze a homogeneous-goods industry where Cournot competition obtains before, and after a horizontal merger. They derive a critical level of the merged entity's post-merger marginal cost function, evaluated at the merging parties' pre-merger joint output level. The merger results in a decrease (resp., increase) in prices if and only if this cost lies below (resp., above) the critical level.

Observed price (or, equivalently, output) effects therefore imply bounds on this marginal cost, denoted by $C_x^M(\overline{x}_1 + \overline{x}_2)$, where $\overline{x}_i$ are the pre-merger outputs for the merging parties, $i = 1, 2$. If the merger resulted in an output expansion in a given market, we obtain an upper bound on $C_x^M(\overline{x}_1 + \overline{x}_2)$, whereas an output contraction analogously implies a lower bound. Motivated by this theoretical bound we define our measure of Merger Efficiencies (ME):

$$ME \equiv min\left\{c_x^1(\overline{x}_1), c_x^2(\overline{x}_2)\right\} - C_x^M(\overline{x}_1 + \overline{x}_2),$$

where $c_x^i(\overline{x}_i)$ denote the pre-merger marginal costs of the merging parties evaluated at their pre-merger outputs.

The ME measure is motivated as the reduction in the social cost associated with employing the resources of the merging parties to produce the marginal unit of output. Prior to the merger, this social cost is $min\left\{c_x^1(\overline{x}_1), c_x^2(\overline{x}_2)\right\}$ as society would wish to allocate the marginal unit to the firm that can produce it at the lowest cost.[7] Following the merger, and holding output fixed, this social cost becomes $C_x^M(\overline{x}_1 + \overline{x}_2)$. The ME captures the difference between these costs and provides a measure of the cost efficiency embedded in the merger.

The ME answers the following question: were the merged firm to continue to produce the same output as its constituent firms, by how much would its marginal cost be lower? It is a well-defined

yields classifications that are very highly correlated and often identical to the ones obtained when simply comparing the post-merger quantities.

[7] In principle, society may also wish to re-allocate output between the merging parties prior to the merger. Our thought experiment, however, holds output levels fixed for reasons that we further motivate below.

measure that does not confound cost changes with output changes. It allows us to focus on the impact of the merger on the marginal cost function, on which we place very minimal structure necessary to guarantee the existence of a Cournot equilibrium. In particular, we do not assume that marginal costs are constant in output, as is often done in empirical work.

At the same time, the ME also reflects some of Bork's concerns as it neglects certain aspects of efficiencies. It is only estimated in markets where both merging parties operated prior to the merger, and therefore does not account for efficiencies that allow the merged firm to enter new markets, or to enjoy lower costs in markets where just one of the merging parties operated originally. In addition, this measure focuses on marginal cost gains, and does not speak to elimination of duplicate fixed costs. Of note, it is the marginal cost savings that affect firms' pricing decisions, making them particularly relevant from an antitrust policy perspective.

Bounding the ME. Bounds on ME are available given the Cournot structure and the observed data. The market's pre-merger demand elasticity is identified from market-level price and output data along with instrumental variables addressing the familiar simultaneity bias.[8] Inverting the Cournot first-order conditions at the pre-merger output levels, and plugging in the identified demand elasticity, delivers unique solutions for the pre-merger marginal costs $c_x^i(\overline{x}_i)$, $i = 1, 2$. Their minimum, which is the first element appearing in the ME measure, is hence point-identified.

The second element of ME is only partially identified: we obtain a single bound on $C_x^M(\overline{x}_1+\overline{x}_2)$ in each affected market where the sign of the merger's output effect is determined via the classification exercise. This bound is the critical level of $C_x^M(\overline{x}_1+\overline{x}_2)$ derived by Farrell and Shapiro (1990) and is easily computed given observed data and the estimated demand elasticity. The ME is therefore set-identified in each affected market where the first-step classification was successful. This is achieved without relying on accounting cost data, avoiding typical pitfalls.[9]

The bound from Farrell and Shapiro (1990) requires that a Cournot oligopoly regime prevails before, and after the merger. Nonetheless, the ME lower bounds are robust to the possibility of coordinated merger effects: if output expansion obtains in the presence of coordinated merger effects, then it would have also obtained without them, validating the lower bound.

The 2013 American Airlines - US Airways merger. We apply the econometric framework to study the nature and magnitude of efficiencies associated with a high-profile merger in the airline industry. While the Cournot structure is obviously restrictive, the important role played by capacity choices in this industry speaks to its relevance.

The merging parties accounted for two of the remaining four "legacy" carriers and their merger created the largest US carrier at the time. Initially challenged by the Department of Justice and several States' General Attorneys, it was subsequently approved via a settlement involving the sale of slots and gates in congested airports to low cost carriers.

We apply our method to airline markets (directional airport pairs) where both merging parties

[8]Some structure on the market demand curve is also necessary. We use a constant-elasticity specification.

[9]Apart from familiar measurement issues, accounting data likely reflect average rather than marginal costs, and accounting rather than economic costs. They may also be strategically manipulated. Those issues provided an important motivation to the emergence of the New Empirical Industrial Organization literature (Bresnahan (1989)).

operated prior to the merger. There are close to four thousand such markets, but we focus on a smaller subset of those markets, 1,014 in number, that satisfy an additional requirement: the market underwent minimal changes to its structure other than the merger itself.[10] We refer to those 1,014 markets as the "affected markets" and classify them, when possible, according to the sign of the merger's output effect.

This exercise reveals that pro-competitive effects prevailed in a considerable number of markets.[11] Using the synthetic control approach we determine the sign of the output effect in 87% of the affected markets. Of those classified, 56% of the markets displayed a positive output effect. Pro-competitive effects therefore prevailed in at least 49% of all affected markets.[12] Using the nearest-neighbor approach we classified 68% of the affected markets, with 65% of these classifications pointing to a positive output effect. Across the two methods we conclude that at least 44%-49% of the 1,014 affected markets experienced output expansions.

With these classifications in hand we proceed to explore the nature of the pro-competitive effects. This is accomplished via linear probability regressions where an indicator function for output expansion serves as the dependent variable. Across such regressions we consistently find that larger markets are more likely to experience output expansions. The probability of output expansion rises by 3-4.5 percentage points when the market population (the geometric mean of the two endpoints' populations) grows by 1 million. This accords with economic reasoning as one expects larger networks to be of greater value in larger markets. Similar reasoning explains why we find a higher prevalence of pro-competitive effects in markets that serve certain hub airports of the merging parties, though this effect is less systematic than that of market size.

We next employ the classifications along with the econometric structure to compute bounds on marginal cost savings. Across the markets classified as experiencing an output expansion, the mean lower bound on the ME measure is slightly above 2 USD. Again averaging across these markets, the marginal cost savings amount to roughly 0.8% of the market price. Importantly, those means are computed *conditional* on output expansion taking place. They do not estimate a mean lower bound on efficiency across all markets. Nonetheless, given the prevalence of output expansion across the affected markets, these figures are informative regarding cost savings in many markets. Analogous upper bounds on the efficiency measure are available in markets experiencing an output contraction. Those average at 2.3-2.4 USD and about 0.9% of the market price.[13]

Literature. The literature on mergers and consolidation in the airline industry is vast and reflects many important changes in market structure since the industry's deregulation in the 1970s. Studies of the early consolidation phases include Carlton et al. (1980), Borenstein (1990), and Werden et al.

[10]Specifically, we require that both the identity of the legacy carriers and the number of the low-cost carriers remained stable between the pre-merger and the post-merger periods.

[11]While we focus on the 1,014 markets where the market structure remained largely stable, this result also holds within the larger group of about four thousand markets where both merging airlines operated prior to the merger.

[12]To clarify this calculation, $0.87 \times 0.56 = 0.49$.

[13]Due to selection, the upper bounds and the lower bounds are not computed over the same objects, and the reported averages do not capture upper and lower bounds on the mean efficiency. Dealing with such selection issues is possible, for example, by following Eizenberg (2014), but would require additional assumptions.

(1991). The past couple of decades witnessed a renewed consolidation wave, the reasons for which have been widely analyzed by economists (see Berry and Jia (2010)).

The American Airlines - US Airways merger (hereafter AA-US) itself has commanded substantial empirical work. Some studies have analyzed this merger *ex-post* using a difference-in-differences approach. Vaze et al. (2017) find that the merger led to a slight price decrease but also to a decline in service frequency, resulting in a small loss of consumer welfare. Le (2019) finds that the merger was pro-competitive in markets where the merging parties competed via nonstop service connecting major hubs, but anti-competitive in markets where they competed via connecting service and in markets where potential competition was eliminated. Das (2019) reports price decreases that were concentrated in larger markets, consistent with our own finding regarding the role of market size.

Other studies of the AA-US merger use pre-merger data to estimate a model of supply and demand and predict the merger's effect as a counterfactual exercise. Ciliberto et al. (2021) and Li et al. (2022) treat not only prices but also entry or service type choices as endogenous. Merger efficiencies are not estimated and merger analysis is performed given several efficiency scenarios. Li et al. (2022) consider a "best case" scenario where the merged entity assumes the minimum of the (constant in output) marginal costs of the merging parties. Ciliberto et al. (2021) similarly model a "best case" in which the merged entity assumes the "best" characteristics of the merging parties, and report that different assumptions regarding efficiencies result in markedly different conclusions regarding the effect of the merger.

Some studies of the AA-US merger focus more intensely on its effect on carriers' cost structures. Bontemps et al. (2023) perform a merger simulation treating both prices and the network structure as endogenous. They explicitly consider the effect of network size and find that a larger network reduces marginal costs but increases fixed costs. Park (2020b) focuses on the remedies associated with the AA-US merger, namely, the sale of gates and landing slots to LCCs in the congested DCA airport, considering the carrier's load factor as a marginal cost shifter.[14]

The current paper departs considerably from existing studies of the AA-US merger, and of mergers in general. First, we tackle the estimation of efficiencies as a primary goal. Second, we use classification methods to determine market-specific output effects, and then plug the inferred effect into a structural model that yields bounds on marginal cost efficiencies.[15]

The rest of the paper is organized as follows. Section 2 lays out the econometric model. Section 3 describes the application to airlines, and Section 4 concludes.

[14]For another recent study of the slot divestment aspect of the merger see Ali (2022).

[15]Some studies estimate airline efficiency from a system approach by focusing on the transformation of inputs into output (seat-miles) and of output into revenue. As reviewed in Khezrimotlagh et al. (2022), most of these studies do not focus on the effect of mergers. Their study reports efficiency gains stemming from the AA-US merger at the level of the entire merged airline.

2 Model

This section lays out the model in several steps. Section 2.1 follows Farrell and Shapiro (1990) closely to derive an important theoretical condition characterizing the sign of the merger price effect. Section 2.2 introduces a complete structural model of supply and demand that embeds the theory in section 2.1 to deliver a set identification result for efficiencies.

This empirical strategy involves, as explained above, a step where the sign of the output effect in specific markets is determined via classifications methods. It is important to note that these classification methods, while forming a component of the strategy implemented via the structural model, are methodologically independent from it. Some of our findings regarding merger efficiencies in the airline case (e.g., the overall frequency of pro-competitive effects and their relationship to market size) derive solely from the classification exercise and are therefore robust to assumptions made in the structural model, including the Cournot behavior.

2.1 Cournot oligopoly theory

Following Farrell and Shapiro (1990), consider a homogeneous-goods market featuring Cournot oligopolists indexed by $i = 1, ..., n$. An exogenous merger between firms 1 and 2 is analyzed. The exogeneity of the merger is appropriate to the airline context where a national-level merger affects many local markets. The set of merging parties, or the "insiders," is denoted by $I = \{1, 2\}$.

Following the merger, the set of active firms is $\{M, 3, ..., n\}$ where M is the merged entity. This implies the assumption that the merger does not prompt entry or exit. This assumption does not fit the data: entry and exit are observed in many markets affected by the merger. It is also inconsistent with the finding in Ciliberto et al. (2021) that the merger affected entry incentives. In the empirical analysis we tackle this issue by focusing on markets that exhibited minimal changes to market structure.

Inverse demand is given by $p(X)$ where X is the aggregate output. Demand is downward sloping: $p'(X) < 0$. The elasticity of demand is given by $\epsilon(X) = p(X)/Xp'(X)$. Firm i's output is denoted x_i, and the total output of its rivals is denoted $y_i = \sum_{i \neq j} x_j$. Firm i's cost, as a function of its output level, is $c^i(x_i)$ and its marginal cost is $c^i_x(x_i)$. Firms may have different marginal costs and hence different outputs in equilibrium. Firm i's profit depends on its own output choices as well a on rivals' aggregate output: $\pi^i(x_i, y_i) = p(x_i + y_i)x_i - c^i(x_i)$.

Two stability conditions are employed. First, the marginal revenue declines with rivals' output:

$$\partial MR^i / \partial y_i = p''(X)x_i + p'(X) < 0. \tag{1}$$

Second, the firm's residual demand curve intersects its marginal cost curve from above:

$$c^i_{xx}(x_i) > p'(X). \tag{2}$$

Firm i's profit maximization given rival's outputs is reflected in its first order condition:

$$p(X) + p'(X)x_i - c_x^i(x_i) = 0, \;\; i = 1, ..., n. \tag{3}$$

The Cournot equilibrium is an output vector $(x_1, ..., x_n)$ satisfying (3) for all firms. The slope of the firm's reaction curve, denoted R_i, is:

$$\frac{dx_i}{dy_i} \equiv R_i = -\frac{p' + x_i p''}{2p' + x_i p'' - c_{xx}^i}, \tag{4}$$

where the last equality is obtained by totally differentiating the first order condition (3). The numerator is negative by (1), and the denominator is negative by a second-order condition for profit maximization. It follows that $R_i < 0$, and, by condition (2), $R_i > -1$. Therefore, if rivals' total output increases, the firm will decrease its own output, but will not fully offset the rivals' output expansion. The relationship between a firm's output change and the aggregate output change can be expressed as follows:

$$dx_i = -\lambda_i \cdot dX, \;\text{ where } \lambda_i \equiv -\frac{R_i}{1 + R_i} = -\frac{p'(X) + x_i p''(X)}{c_{xx}^i(x_i) - p'(X)}. \tag{5}$$

The second equality is obtained by manipulating (4).[16] Importantly, the stability conditions (1) and (2) imply that $\lambda_i > 0$ for all $i = 1, ..., n$. Finally, it is assumed that following the merger, quantities adjust to re-establish a Cournot equilibrium.

These assumptions give rise to two results from Farrell and Shapiro (1990), hereafter FS90. The first is a Lemma that characterizes responses to an exogenous change in output by a single firm.

Lemma 1. *[Farrell and Shapiro (1990)] Consider an exogenous change in firm 1's output, following which the other firms adjust their output choices to re-establish a Cournot equilibrium among themselves. Aggregate output moves in the same direction as the output of firm 1, but by a smaller magnitude.*

Proof for this Lemma is provided in Appendix B. That appendix also shows the proof for the following Theorem, building on the Lemma.[17]

Theorem 1. *[Farrell and Shapiro (1990)] Consider a merger of firms $i = 1, 2$ and denote the merged entity by M. The merger raises prices if and only if*

$$\overline{p} - c_X^M(\overline{X}_M) > (\overline{p} - c_x^1(\overline{x}_1)) + (\overline{p} - c_x^2(\overline{x}_2)),$$

where $\overline{p}, \overline{x}_i$ are the pre-merger price and output for firm i, respectively, and $\overline{X}_M = \overline{x}_1 + \overline{x}_2$.

Theorem 1 provides a necessary and sufficient condition for the merger to have an anti-competitive effect (i.e., for a price increase): the merged entity's markup, evaluated at the joint pre-merger quan-

[16] The first equality is derived as follows: the definition of R_i implies that $dx_i = R_i \cdot dy_i$, from which we obtain that $dx_i(1 + R_i) = R_i(dx_i + dy_i) = R_i \cdot dX$.

[17] Farrell and Shapiro (1990) state this result for mergers that may involve more than two firms. As such mergers are less relevant for an empirical context we focus here on the two-firm merger case.

tities of the merging parties, must be lower than the sum of the pre-merger markups. Combining this with the first-order conditions in (3) leads to the following Corrolary.

Corollary 1. *Given that Cournot equilibrium prevails both before and after the merger, the merger results in a price increase if and only if*

$$c_X^M(\overline{X}_M) > \overline{p}\left(1 - \frac{\overline{s}_1 + \overline{s}_2}{\epsilon(\overline{X})}\right).$$

Notice that the right-hand side of the inequality in Corollary 1 is completely determined by pre-merger quantities: the market price, the market shares of the merging parties, and the elasticity of demand. In the context of the econometric model presented below, all these quantities shall be either observed or point-identified.

Given the *ex-post* nature of the econometric analysis, it would also be determined whether the merger prompted a price increase (equivalently, an output contraction), or a price decrease (an output expansion) in each market, subject to methodological challenges elaborated below. Corollary 1 then delivers a single bound on $c_X^M(\overline{X}_M)$ in each market: a lower bound if the market experienced a price increase, and an upper bound if the price was decreased. Along with additional structure provided in the next section, this Corollary shall provide the foundation for the set-identification of the ME measure of merger efficiencies.

2.2 A structural econometric model

The Cournot theory is next embedded in a structural model of demand and supply applied to the airline industry. The econometric framework involves two logical steps. In a first step we use classification methods to determine the sign of the merger's effect on output in specific airline markets. The second step uses these classifications, along with assumptions on the structure of demand and supply, to compute bounds on the ME measure of efficiency within these markets.

Denote the set of all airline markets $\mathcal{T} = \{1, 2, ..., T\}$ and let t index individual markets in this set. The active airlines in market t prior to the merger are indexed by $i = 1, ..., I_t$.[18]. Further denote by $\mathcal{T}^a \subset \mathcal{T}$ the set of markets where both AA and US were present prior to the merger. These are the markets directly affected by the merger. In those markets, let AA and US be denoted by $i = 1$ and $i = 2$, respectively.

We observe p_t^{pre}, the observed pre-merger price in each market $t \in \mathcal{T}$. In addition, the observed variables include airline-specific pre-merger quantities $\overline{x}_{it}$ for each airline $i = 1, ..., I_t$ in each market t. The aggregate pre-merger quantity is therefore observed and given by $\overline{X}_t = \sum_{i=1}^{I_t} \overline{x}_{it}$. We also observe the market-specific post-merger aggregate output, X_t^{post}. Also observed is a vector of market characteristics ℓ_t containing, among other variables, the flying distance and the geometric

[18]The model ignores the possibility of entry and exit around the merger event and treats this set as time-invariant. In practice we restrict the set of affected markets to consider only those where very minimal entry and exist occurred as discussed in Section 3.

mean of population at the market's endpoints, a typical measure of market size. Finally, the observed variables include a vector z_t of market-specific cost shifters.

Let the subsets $\mathcal{T}^+ \subset \mathcal{T}^a$ and $\mathcal{T}^- \subset \mathcal{T}^a$ contain affected markets where the effect of the merger on output was positive and negative, respectively. A naive comparison of the observed post-merger market output, X_t^{post}, to its observed pre-merger counterpart, $\overline{X}_t$, will not do to classify the sign of this effect. The reason is that changes in output could stem from factors unrelated to the merger, such as changes in the cost of jet fuel or in demand. We now very briefly explain our approach to dealing with this challenge, leaving additional details to Section 3.

The two methods employed to perform the classification (the synthetic control and the nearest-neighbor methods) associate a "control market" to each affected market. Both methods consider an original pool of potential control markets. The nearest-neighbor method chooses the single market within that pool that matches the pre-merger output behavior of the affected market as close as possible, in the sense of minimizing the euclidean distance between the output vectors of the two markets in a "training period" that takes place before the merger. We use the post-merger output in that control market as a prediction for the output of the affected market in a no-merger counterfactual. By comparing these counterfactuals to the observed output we can learn about the sign of the merger's effect on output in the affected market.

The synthetic control approach, instead, forms the counterfactual no-merger output prediction by applying weights to the output of potential control markets, as well as to other market characteristics. The weights are chosen to minimize the discrepancy between the synthetic output and the actual output of the affected market in the training period, and are then used to predict the post-merger counterfactual output.

In both methods, we compute these predictions for four different post-merger quarters in 2016 (and, in the nearest-neighbor case, we also do this using four different pools of potential control markets). Within each method, we require that a majority of the predictions points in the same conclusion. We relay additional details on this process to the empirical application section below.

Having obtained the classification of affected markets into those experiencing output expansion versus those experiencing output contraction as a result of the merger, we proceed to compute bounds on the ME measure of efficiencies in each such market. To this end, we place structure on the supply and demand conditions across markets.

On the supply side we assume that each market is in a Cournot equilibrium before, and after the merger, retaining the assumptions of section 2.1 precisely. Recall that these assumptions allow different carriers to have different cost functions and also allow these functions to be quite general. In particular, we do not impose that marginal costs are constant in output, a common assumption in empirical work.

On the demand side we assume that passengers consider the product to be homogeneous. In the airline context, this amounts to assuming that passengers do not consider the service offered by different carriers as differentiated. This is, of course, a simplifying assumption, but one that allows us to build directly on the theory in FS90.[19]

We model the market demand curve using a simple constant-elasticity specification that depends on market characteristics:

$$X_t(p_t) = a_t \cdot (p_t)^{-e_t}, \tag{6}$$

where $a_t = exp(\ell_{1t}\alpha + \xi_t)$ and $e_t = \ell_{2t}\eta$. The demand parameters α and η govern the level, and the elasticity of demand, respectively, whereas ℓ_{1t} and ℓ_{2t} are (partially overlapping) subsets of the ℓ_t vector of market characteristics. Finally, ξ_t is an unobserved demand shifter in market t. Applying a log transformation to both sides of (6) yields the following:

$$ln(X_t) = \ell_{1t}\alpha - \Big(\ell_{2t} ln(p_t)\Big)\eta + \xi_t. \tag{7}$$

Equation (7) delivers a linear regression defined over the cross-section of markets. The market price p_t is endogenous, and is interacted with each element of the market characteristic vector ℓ_{2t}. The demand parameters α, η are therefore point-identified given a vector of instruments z_t with dimension of at least $dim(\ell_{2t})$, to be discussed in section 3.3. So is the absolute value of the demand elasticity $\epsilon(X) = (-e_t) = -\ell_{2t}\eta$. Notice that we obtain a demand elasticity that can vary across markets but is constant-in-output within market. In the demand estimation results presented below, ℓ_{2t} is allowed to contain only a constant, so the demand elasticity is also constant across markets.

Given the estimated market demand elasticity, and the classification of each affected market t according to the sign of the merger's output effect, we are ready to employ the results of section 2.1 and obtain a bound on ME.

Define $b_t(\overline{p}_t, \overline{x}_t, \ell_{2t}; \eta) \equiv \overline{p}_t \Big(1 + \frac{\overline{s}_{1t} + \overline{s}_{2t}}{\ell_{2t}\eta}\Big)$ where $\overline{p}^t$ is the observed pre-merger market price, and the vector $\overline{x}_t = (\overline{x}_{1t}, ..., \overline{x}_{nt})$ collects the observed pre-merger quantities of all firms in market t. Observing $\overline{x}_t$ therefore implies knowledge of the pre-merger market shares of each firm including those of the merging parties, $\overline{s}_{1t}$ and $\overline{s}_{2t}$. Along with the observed demand shifters ℓ_{2t} and the identified demand elasticity η, all the components of $b_t(\cdot)$ are either observed or point-identified. Corollary 1 implies that $b_t(\cdot)$ delivers a single bound on $c^M_{X,t}(\overline{x}_{1,t} + \overline{x}_{2,t})$, the merged entity's marginal cost evaluated at the total pre-merger output of the merging parties, in each affected market where the sign of the output effect has been determined:

$$\begin{aligned} c^M_X(\overline{x}_{1,t} + \overline{x}_{2,t}) &> b_t(\overline{p}_t, \overline{x}_t, \ell_{2t}; \eta) \text{ in each } t \in \mathcal{T}^- \\ c^M_X(\overline{x}_{1,t} + \overline{x}_{2,t}) &\leq b_t(\overline{p}_t, \overline{x}_t, \ell_{2t}; \eta) \text{ in each } t \in \mathcal{T}^+ \end{aligned} \tag{8}$$

That is, we obtain a lower bound on this marginal cost in markets where the merger prompted

[19] In ongoing work we explore additional assumptions regarding the demand side that would still be consistent with the theoretical results obtained above, and hence still allow us to use the bounds from Corrollary 1.

an output contraction, and an upper bound in markets where output expansion was displayed.

The inequalities in (8) provide a bound on the second component of the ME measure. The first component is the minimum of the merging parties' pre-merger marginal costs. These marginal costs, and hence their minimum, are point identified from the Cournot first order conditions (3):

$$\begin{aligned} c^i_{x,t}(\overline{x}_{i,t}) &= \overline{p}_t + \overline{x}_{i,t} p'(\overline{X}_t) \\ &= \overline{p}_t - \overline{x}_{i,t} \ell_{2t} \eta \Big(exp(\ell_{1t}\alpha + \xi_t) \ell_{2t} \eta \Big) (\overline{X}_t)^{-\ell_{2t}\eta - 1}. \end{aligned} \tag{9}$$

Firm i's marginal cost in (9) is indeed identified as the right-hand side features observed variables (the pre-merger market price and aggregate output, the firm's own output, and observed market characteristics ℓ_{1t} and ℓ_{2t}) and objects that are identified from the regression in (7): the parameters η and α and the market-specific demand error ξ_t. We therefore point-identify the smaller of the merging parties' pre-merger marginal costs, denoted $h_t(\overline{p}_t, \overline{x}_t, \ell_{1t}, \ell_{2t}; \alpha, \eta, \xi_t) \equiv min\Big\{ c^1_x(\overline{x}_{1t}), c^2_x(\overline{x}_{2t}) \Big\}$.

Finally, combining (8) and (9) we obtain bounds on the ME measure of merger efficiencies. To re-iterate, the ME in a given market t captures the difference between the minimum of the merging parties' pre-merger marginal costs, and the marginal cost of the merged entity evaluated at the merging parties' total pre-merger output:

$$ME_t \equiv h_t(\overline{p}_t, \overline{x}_t, \ell_{1t}, \ell_{2t}; \alpha, \eta, \xi_t) - C^M_x(\overline{x}_{1t} + \overline{x}_{2t}). \tag{10}$$

By the above arguments, the first component of (10) is point-identified, whereas the second component is set-identified in the sense that we obtain a single bound on it in each affected market t where the sign of the output effect was determined. The bound on the ME in each such market t can be written as:

$$\begin{aligned} ME_t &\leq h_t(\overline{p}_t, \overline{x}_t, \ell_{1t}, \ell_{2t}; \alpha, \eta, \xi_t) - b_t(\overline{p}_t, \overline{x}_t, \ell_{2t}; \eta) \text{ in each } t \in \mathcal{T}^- \\ ME_t &\geq h_t(\overline{p}_t, \overline{x}_t, \ell_{1t}, \ell_{2t}; \alpha, \eta, \xi_t) - b_t(\overline{p}_t, \overline{x}_t, \ell_{2t}; \eta) \text{ in each } t \in \mathcal{T}^+. \end{aligned} \tag{11}$$

Coordinated merger effects. The above analysis assumes that firms compete according to the Cournot model both before and after the merger, giving rise to the theoretical bound on post-merger marginal costs from FS90. While theoretically appealing, this assumption rules out important behavioral possibilities. Among them is the possibility of coordinated merger effects — namely, that the merger causes a shift towards less competitive regimes. As discussed below in section 3.1, this concern has been raised with regard to the AA-US merger by the Department of Justice and also by some recent studies.

It is therefore important to consider how the analysis changes if one assumed that Cournot competition prevails prior to the merger, but the merger then prompts a change in conduct towards a less competitive regime. In this case, we still obtain the upper bound in (8) on marginal costs in markets $\mathcal{T}^+$ where the merger caused an output expansion. The reason is that if output expands despite the shift to a less competitive regime, then it would have expanded even more had the Cournot behavior been maintained after the merger. For this to be true, marginal costs cannot exceed the upper bound prescribed in (8). As a consequence, the lower bounds on ME in (11) hold under post-merger coordination.

Obtaining lower bounds on post-merger marginal costs, and, consequently, upper bounds on cost efficiencies is more difficult in the presence of coordinated merger effects. This difficulty is not only computational but conceptual: repeated games have infinitely-many equilibria associated with different degrees of departures from the competitive benchmark, ranging from mild coordination to the monopoly solution. Each such equilibrium would intuitively be associated with a different upper bound on the efficiency implied by an observed output contraction. We do not pursue such calculations in this paper.[20]

3 The American Airlines - US Airways merger

3.1 The merger

Having created the largest US carrier at the time, the 2013 merger of American Airlines and US Airways prompted much government scrutiny before being approved via a settlement. The case against the merger is described in detail in Olley and Town (2018). The government identified many markets where the merging airlines competed directly and was particularly concerned with the possibility that US Airways served the role of a "maverick" in such markets. US Airways' hubs were located in smaller markets and it relied more heavily on connecting flights relative to the other three legacy carriers. It competed aggressively by charging low prices on connecting service via its "Advantage Fares" program.

These differences between US Airways and the other legacy carriers suggested the possibility that its merger with American Airlines would result not only in unilateral merger effects, but also in coordinated effects. The government indeed argued that the merger would render the remaining legacy carriers more symmetric, making collusion easier. Consistent with such concerns, Turner (2021) and Kim and Park (2023) find evidence in support of coordinated effects associated with the AA-US merger.

On the other hand, the merger could possibly result in efficiency gains delivering welfare benefits. One source of potential efficiencies is the merged entity's expanded network. Our study does not focus on reductions in fixed costs, or on "network efficiencies" reflected in an expanded choice

[20] Farrell and Baker (2021) offer a critical discussion of the repeated game framework in analyzing coordinated merger effects, particularly on account of the many-equilibria issue. One of their suggestions is to consider a paradigm where the merger induces a shift from a simultaneous-move game to a game where one firm moves ahead of its rivals. But they motivate this approach in the context of a price-setting game where actions are strategic complements, as opposed to the quantity-setting framework pursued here.

set for passengers.[21] Rather, we focus on estimating marginal cost reductions, and it is therefore instructive to consider what operational changes could result in such reductions.

One possible case for operational efficiencies was recently made in an ongoing merger case involving JetBlue and Spirit. JetBlue's CEO has made the argument that the merger would result in expanded, rather than contracted capacity because it would allow the airlines to use larger planes on existing routes and to keep planes in the air for longer periods of time.[22] While this claim may pertain to efficiency gains throughout a merged entity's network, our analysis focuses on estimating such gains in markets where the merging parties were in direct competition prior to the merger.

3.2 Data

This study uses publicly-available data from Domestic airline markets in the United States, obtained from the Department of Transportation (DOT) website. The Origin and Destination Survey (DB1B) database provides a random sample of 10% of tickets issued by reporting carriers where an observation corresponds to an itinerary, defined by the origin, destination, and any connections, and provides information with respect to each leg comprising the itinerary. The T-100 database offers information on service at the segment level such as carriers' passenger capacity and actual numbers of passengers.

We use quarterly data for the years 2009 through 2016. At the itinerary level we observe the price paid, as well as the ticketing and operating airlines. In line with much of the literature we exclude itineraries involving multiple stops, those with very high or low prices, and those involving a very short travel distance. Complete details are available in Appendix C.

Following common practices, an airline market is defined as a directional airport pair (e.g., connecting Chicago O'hare and JFK) in a particular quarter. While it is also common to define markets as city pairs, Li et al. (2022) make the point that in analyzing mergers, considering airports has the advantage of speaking to issues relating to airline presence in specific airports, often viewed as a source of both efficiency and market power. These authors also motivate a directional market definition because of the impact of a carrier's presence at the origin airport on consumer demand.

As a consequence, and again consistent with prior studies, round-trip tickets were split into two directional observations, where each direction was assigned 50 percent of the round-trip price. As in Park (2020b) and Li et al. (2022) markets connecting two airports that belong in the top-100 airports by boarding counts were considered. The US census was used to obtain the population of each metropolitan area associated with such airports. The market's population was computed as the geometric mean of the populations at the two endpoints, again consistent with these studies.

In a typical market airlines may offer nonstop service alongside connecting service options. For example, a market defined by its endpoints, airports A and B, may have airline 1 offering nonstop service, airline 2 offering a service connecting through some hub airport C, and airline 3 offering two different connecting services via hub airports D and E.

[21] For a critical view of such claims for efficiency in the context of the AA-US merger see, for example, Moss (2013).

[22] "U.S. rejects JetBlue, Spirit exemption request, citing lawsuit." David Shepardson, Reuters, March 25 2023.

A first look at the pre-merger presence of the merging parties across markets is available in Table 1. These two legacy carriers are present in many of the 7,911 markets (again, directional airport pairs connecting top-100 airports). Both airlines competed with one another, prior to the merger, in about 50% of these markets, namely in about 4,000 markets. Some 15%-18% of all markets feature AA only prior to the merger, whereas about a quarter of the markets feature US only. In these markets the merger may eliminate potential competition whereby US would enter AA-only markets and vice-versa. This, however, largely overstates the number of markets with potential competition.[23] The single merging party operating in such markets may also experience changes to its efficiency following the merger. Our study, however, abstracts from analyzing the impact of the merger on such markets and focuses solely on markets where both parties were present prior to the merger.

Table 2 reports statistics computed over the 32,140 quarter-market pairs in 2012. The market price, computed as a weighted average over all tickets sold in the relevant market and quarter, has a mean of 257 USD and a standard deviation of 81. Total quarterly passengers average at 974, again with substantial variation. The table further reports statistics on the market's population and distance, the latter expressed in terms of non-stop miles.

3.3 Estimation and results

The empirical analysis is carried out in several steps. First, markets are classified into the $\mathcal{T}^+$ and $\mathcal{T}^-$ sets that contain markets that experienced output expansion and contraction effects, respectively. Market-specific aggregate demand curves are estimated next (equation (7)). Given these demand estimates, pre-merger marginal costs are calculated by inverting the Cournot first-order conditions in (9). Post-merger marginal cost of the merged entity (evaluated at the pre-merger joint output) are then bounded using the classifications and the estimated demand elasticity, completing the partial identification of the ME measure of merger efficiencies. These tasks are taken up, in turn, in the following subsections.

3.3.1 Classifying output effects and the sources of efficiencies

A key step in the analysis is to identify markets where the merger resulted in an output expansion versus those that experienced a contraction in output. While output (and prices) are observed both before and after the merger, classifying markets into these two sets poses a nontrivial challenge.

The theory reviewed in section 2.1 studies the effect of the merger holding all other market conditions constant. In that setup, output may change only as a consequence of the merger. If this were the case in reality, then observing that post-merger output exceeded its pre-merger counterpart would have been sufficient to determine that the pro-competitive effect overwhelmed the anti-competitive one in a given market. In practice, multiple variables change concomitantly with the merger. Fuel costs, for example, experience a decline in the time period surrounding the

[23] Le (2019), for example, defines potential competition in a market if one of the merging parties is present whereas the other has presence in one of the market's endpoints, facilitating its potential entry.

AA-US merger. If we naively compare output before and after the merger, an anti-competitive merger effect could be falsely classified as a pro-competitive one because its true causal effect was offset by the fuel cost reduction.

This challenge is similar to the one faced by typical studies that examine the *ex-post* effect of consummated mergers. Difference-in-differences econometric strategies are often employed to net out the effects of changes in the economic environment other than the merger. Those studies, however, aim at identifying an average treatment effect of the merger across markets. Our setup, however, requires a classification of *each market* as experiencing either a pro-competitive or an anti-competitive effect. This is a prediction problem.

We tackle this challenge using two different strategies. The first strategy follows the KNN (K nearest neighbors) method, introduced by Fix and Hodges (1951). A second strategy involves a synthetic control group (SCG) method following Abadie (2021). Both strategies use information regarding markets that were presumably unaffected by the merger to deliver a prediction of the post-merger output for each treated market (we use the terms "treated market" and "affected market" interchangeably).

The predicted values for post-merger counterfactual market outputs are then compared to the actual post-merger outputs to determine the classification. If the counterfactual output exceeds the observed output in the affected market, we classify this affected market as experiencing an anti-competitive merger effect. Analogously, if the counterfactual output falls short of the observed one, we determine that a pro-competitive effect obtained. Both strategies require definitions for the group of affected markets, as well for the group of control markets.

We define the affected markets as those satisfying the following criteria: both US and AA were active in the market prior to the merger; the set of active legacy carriers did not change over the 2012-2016 period; and, the number of LCC competitors remained constant over this time period. These choices aim at selecting markets where the merger was the only major change to the competitive landscape over the studied period. A total of 1,014 such treated markets are present in the data, and our study of efficiency gains is performed over this set.[24] Below we conduct sensitivity checks to explore the implications of this limitation of the set of affected markets, noting again that the total number of markets where both carriers competed prior to the merger was about four thousand.

Having defined the group of markets affected by the merger, we next turn to defining the set of control group markets unaffected by the merger. This definition involves a trade-off: the more conservative is the definition, the more we succeed in keeping markets that are affected by the merger out of the control group, at the cost of having a smaller control group. But as the control group becomes smaller, it contains fewer markets that may provide a good match for the affected market. Analogously, insufficiently conservative criteria will result in a better fit, but may be more

[24] LCC carriers are less differentiated than legacy carriers, motivating our requirement that only their number, rather than their identity, remains constant over time. Requiring instead that the exact same set of LCC carriers would be present in the data throughout the studied period would have reduced the number of treated markets to 140, reflecting a considerable loss in sample size.

difficult to justify from an economic standpoint.

To address this trade-off we consider four different control group specifications. The most restrictive criteria require both merging parties to be inactive in the market prior to the merger, and both the set of active legacy carriers and the number of active LCCs to be fixed across all periods. This results in a control group of 146 markets. In a second, less restrictive set we remove the limitations on the legacy carriers and LCC activity, but still require inactivity by the merging carriers. This version results in 777 control markets. A third specification allows for a maximum of one merging party to be active in the market, leaving 4,044 markets. Lastly, the forth and least conservative set contains all 6,854 non-treated markets, essentially considering any market where at least one of the merging parties was not present prior to the merger. Table 3 summarizes the criteria used to define the treatment and control groups.

The nearest neighbor approach relies on the idea of "matching" markets and provides simple and easy-to-interpret predictions. For each of our 1,014 affected markets we pick its K "nearest neighbors:" control group markets that are the most similar to the affected market according to some similarity criterion. Then, we interpret the post-merger average output of these K neighbors as a counterfactual output that we would have expected to observe in the affected market absent the merger.

The neighboring criterion we use in our application is the euclidean distance between the vectors of market outputs (passenger quantities) in the 20 quarters prior to the merger, that is, in 2009-2013.[25] For simplicity we set $K = 1$, practically defining a one-market control group comprising of the single market that is most similar in its 2009-2013 output levels to the affected market. The chosen neighbor's post-merger quantities become the predicted quantities.

In examining post-merger outputs we use the control market output levels as predictors of the affected market output levels in each of the four quarters of 2016. We determine the classification of the affected market via a "majority rule." Essentially, we require the classification (output expansion of contraction) to be identical in at least three of the four post-merger quarters.

To clarify the meaning of this decision rule, note that a certain affected market obtains its best match from the set of control group markets. The counterfactual passenger quantities in this affected market, absent the merger, in the four quarters of 2016, are then predicted using the chosen neighbor's quantities in these four quarters. For each quarter we compare the observed quantity in the affected market with the prediction. We then count the number of quarters in which the observed quantity exceeds the predicted one. If the number is 3 or 4, the market is classified as having experienced a pro-competitive effect. If the number is zero or one, an anti-competitive classification obtains. If the number is 2, no classification is given.

Since four control groups were defined, the KNN process is done for each one, providing four classifications for each of the 1,014 markets. To combine it all to a single classification, we run another majority rule vote among the four control group based classifications. For this decision we

[25] While determining the "pre-merger" period is never a trivial challenge, we note that Vaze et al. (2017) use the last two quarters of 2013 as the pre-merger period in their difference-in-differences analysis of the AA-US merger.

require at least three valid votes (clear classifications), and a majority classification among them, that is, a majority of at least 3 out of 4 or 2 out of 3. This "double majority rule" classification process balances between prediction of all control groups with equal weights.

The *synthetic control approach* forms the counterfactual quantities in a different fashion. Rather than picking a single control market, it creates a synthetic control market for each affected market via a weighted average over potential control markets and their characteristics.

The method defines a 'training' period over which an optimal set of weights is defined. The weights are defined over the set of control units (in our case, markets deemed to be unaffected by the merger) and additional variables (e.g., the market population or flying distance of such control units). The weights minimize the euclidean distance between the observed outcome variable (in our case, the output) in the treatment group, and its predicted value, computed as the sum of the product of each variable and its assigned weight. The weights create a synthetic version of the treated unit that consists of the different markets and variables.

We continue to use the 20 quarters ending at 2013Q4 as the pre-merger training period, as in the KNN method described above. A synthetic market is created using weights on each control group market, as well as each quarter's mean quantity, population, mean flown market distance, and the numbers of active legacy and LCC carriers. Once the weights are calculated, they are applied to the post-merger period to create predictions for each treated market in each of the four quarters in 2016. As before, a majority rule is applied over the four quarters to enhance the robustness of our predictions.

In contrast to the KNN method, the SCG was applied only using the first, most conservative, donor pool of control markets for computational reasons. As a consequence, it utilizes a majority rule, as opposed to the "double majority" rule in the KNN case that also required some consensus over different sets of potential control markets.

Classification results: the prevalence of output expansion. The results of the classification exercises are reported in Table 4. Classification is attempted over a total of 1,014 affected markets where both merging parties were present on the eve of the merger, and that displayed minimal changes in market structure over the studied period other than the merger (namely, the identity of the other legacy carriers remained intact and the number of LCC carriers did not change).

The left column reports the results of the synthetic control classification. Recall that this classification was performed using a majority rule over predictions made with regard to output levels in the four post-merger quarters of 2016. A total number of 136 markets (or, 13% of the 1,014 affected markets) were not classified as no majority was achieved among the four predictions. A total number of 494 markets were designated as having experienced an output expansion, and a total number of 384 were classified as having experienced an output contraction. These results suggest that a substantial number of markets experienced pro-competitive effects. These 494 markets constitute 49% of all affected markets, and 56% of all such markets where a classification was available.

Results from the nearest-neighbor classification strategy, reported in the middle column, echo a similar overall picture. While the fraction of markets classified decreases here from 87% to 68% (possibly because of the more stringent "double-majority" rule), a total of 65% of those markets that were classified were determined to have experienced an output expansion. The pro-competitive effect therefore prevailed in almost two-thirds of those where the effect was classified, and in at least 44% of all affected markets (this is calculated by $0.65 \times 0.68 = 0.44$).

Finally, on the right-hand column of Table 4 we apply the most severe classification rule, whereby we only classify the sign of the output effect in markets where a consensus is obtained between the nearest-neighbor and the synthetic control methods. That is, both have to arrive at a classification, and agree about it, for a classification to be achieved according to this "consensus" rule. Not surprisingly, the fraction of classified markets decreases substantially to 44%. However, about two-thirds of the markets that are classified are determined to have experienced an output expansion. The notion that a considerable fraction of markets experienced a pro-competitive effect that overwhelmed the anti-competitive one is therefore robust across the various classification rules.

Where do output expansions occur? The fact that we obtain market-level classifications of the sign of the merger's output effect allows us, further, to explore where output expansions occur. To this end, we employ Linear Probability regressions where an indicator variable for an output expansion classification serves as the dependent variable. These results are reported in Table 5. Panels A and B of this table report results based on the synthetic control and nearest-neighbor classification schemes, respectively. While we do not interpret the coefficients as causal, they nonetheless capture patterns of interest.[26]

The market size is positively and strongly correlated with the incidence of pro-competitive merger effects: as the market population (the geometric mean of the populations at both end-points) increases by 1 million, the probability of output expansion rises by 2.7 to 3.7 percentage points. These effects are statistically significant at the 1 percent or at the 5 percent levels in all specifications. To the extent that marginal cost reductions are available via economies of scale or higher network density, one may expect those effects to be more pronounced in larger markets. Consistent with this, Das (2019) finds that fares have decreased following the AA-US merger in larger markets using a difference-in-difference approach.

The presence of hubs of the merging carriers at either endpoint of the market is also positively associated with efficiency gains. One may expect the merged carrier to be enjoy more pronounced efficiency gains in markets that interact with its hub network. Nonetheless, these effects are not always statistically significant. In particular, one often has to omit the market size variable (as in columns 3 and 4) to gain significance for these variables.

Finally, we attempt to explain efficiency gains via a variable to which we assign the name "network advantage." This variable divides the total number of markets served by the merging parties out of the origin airport by the number of markets they serve out of the destination prior to the merger. The logic underlying this variable is that more efficiency gains may be available

[26]The standard errors reported in these regressions do not take into account the error stemming from the classification process itself, in which the dependent variable is determined. Since that process relies on classification methods, there is no straightforward way to account for this issue.

in markets that shift aircraft towards locations where they may be used more intensely. As the table shows, however, the effect of this variable on the probability of output expansion is small and statistically insignificant. As we shall see below, however, a similar variable is found to be an effective instrument for the market price in our demand analysis.

To summarize the classification exercise, it appears that efficiency gains overtake the anti-competitive effects in a substantial portion of the markets affected by the merger. These efficiency gains seem to be more prevalent in larger markets, and there is also some weak evidence that they may be stronger in markets served by hub airports of the merging parties.

The results presented above have not, as yet, used the structure of the model, relying instead strictly on the statistical prediction exercises. Combining these classifications with the structure of the model, as we do next, enables us to also bound the magnitude of the marginal cost reductions.

3.3.2 Demand

The demand equation (7) is estimated via two-stage least squares to address the familiar simultaneity issue (or, equivalently, the endogeneity of price). Table 6 presents estimation results. Markets in the four quarters of 2012 are included in this regression, resulting in 31,590 observations. The use of pre-merger observations is warranted since it is the pre-merger demand parameters that are used to compute the bounds on the ME measure of efficiencies.

Several variables are included in ℓ_{1t}, i.e., as shifters of the level of the market demand. As expected, the market population has a positive effect on demand, and so does the variable "nonstop miles." Recall that this variable provides an approximation to the actual travel distance since the market demand estimated here aggregates over nonstop and connecting service options. As travel distance becomes longer, the value of the outside option (travel by modes of transportation other than a commercial flight) decreases, consistent with this positive distance effect.

Another variable shown to have a positive effect on the level of demand is the number of markets that are served out of the origin airport in the relevant quarter ("origin markets"). A larger number of markets served out of the destination could stimulate demand via loyalty programs.[27] Finally, dummy variables for the first three quarters of 2012 are included. The omitted quarter is the fourth quarter containing the busy travel season. The first and third quarters have a negative demand effect relative to the omitted quarter, while the second quarter has a positive effect.

The above patterns for the variables affecting the level of demand hold for both the OLS (left column) and the 2SLS (right column) specifications. The 2SLS specification addresses the familiar simultaneity of supply and demand by way of instrumenting for the endogenous price variable. Two such instruments are used. The first is the number of markets served out the destination airport divided by the number of markets served out of the origin airport. Large values of this variable suggest that serving the market helps carriers shift aircraft towards airports where they can be used more effectively. This implies cost savings that are unrelated to demand, noting that the number

[27] A similar argument is made in Li et al. (2022) with respect to the presence of specific carriers at the origin airport. The variable used here mirrors that interpretation for our aggregate demand specification.

of markets served out of the origin airport is controlled for as a demand shifter. Consistent with this logic, this variable has a negative effect on demand in the first-stage regression.[28]

The second instrumental variable is a cost measure capturing fuel costs. The DOT data allow us to compute a cost per gallon that is carrier-quarter specific. On a given market, we weight these measures (for one or two legs as appropriate) by the distance flown on the leg, as well as by total passengers, and then average over all itineraries within a market. This variable can only be computed for a subset of the dataset due to missing data. As a consequence, the 2SLS results are obtained over a sample of 17,614 observations. In the first-stage regression it has the expected positive effect on ticket prices.

The results presented here pertain to a model where ℓ_{2t} includes a constant only, so that the demand elasticity in all markets, in absolute value, is the inverse of the estimated parameter η. Demand is downward sloping in both the OLS (left column) and the IV specifications. As expected, demand is more elastic when the endogeneity issue is addressed: the elasticity is estimated to be (-2.60) under OLS and (-5.58) under 2SLS.

3.3.3 Pre-merger costs

Next, marginal costs are computed using the carrier first-order condition (9) and the estimated demand parameters. Table 7 reports statistics computed over 15,884 market-quarter cells in 2012 where both US and AA were present.

The distribution of pre-merger marginal costs appears similar across the two merging parties, with mean values of about 250 USD, minimum values of 90 USD and maximum values above 600 USD. On average across markets, US seems to have slightly higher marginal costs than AA, where the mean difference is 1.3 USD. This similarity in mean marginal costs masks considerable heterogeneity across markets, whereby in some markets the marginal cost difference comes close to 80 USD. The distribution of this cost difference across markets is displayed in Figure 1.

This figure shows that the mass of markets is quite evenly distributed between markets where AA enjoys a cost advantage versus markets where US has the advantage. Recall that in the Cournot framework a firm that produces a larger output than its rival has a lower marginal cost. So the markets with negative values for the cost difference variable are simply those where US carries more passengers than AA, and the markets with positive values correspond to cases where AA carries more passengers than US.

Finally, Table 7 reports that the mean pre-merger markup across markets is 2.4 percent for US Airways and 2.9 percent for American Airlines. The markup is computed by taking the difference between the market price $\overline{p}_t$ and the carrier's marginal cost divided by the market price.

[28]This approach bears some analogy to the approach in Brancaccio et al. (2020) who instrument for exporters' shipping costs using tariffs levied on goods exported from the destination.

3.3.4 Bounds on merger-induced efficiencies

Having classified markets into those that experienced pro-competitive effects versus anti-competitive effects, we may determine, for each affected market, whether it generates a lower bound, or an upper bound on the ME measure of merger efficiencies via (11). Given the estimated pre-merger demand and marginal costs, we are also able to compute these bounds. Table 8 presents the statistics regarding the distribution of estimated bounds across the treated markets.

Panels A and B display bounds statistics that are implied by the synthetic control and by the nearest neighbor classification schemes, respectively. The classification does not affect the computed bounds. It does, however, tell us whether the bound computed in a given market is a lower bound, or an upper bound on the ME measure. Panel C refers to a classification scheme that requires a strict consensus among the synthetic control and the nearest neighbor analyses.

Across the three panels, the distribution of the bounds does not vary significantly, which is not surprising given that the different classification schemes deliver results that are quite correlated. The mean lower bound across affected markets that were classified as experiencing an output expansion is 2.12-2.18 USD. The marginal cost reduction, on average across such markets, is therefore slightly above 2 USD, or, about 0.8% of the market price. Beyond these averages we observe quite a bit of variation in lower bounds on ME. In some markets, they are as high as 20 USD or 6.5% of the market price, whereas in some other markets they are very close to zero.

The distribution of the upper bounds on ME, likewise, is very similar across the three classification schemes. On average across markets experiencing output contractions, the upper bounds are about 2.20-2.40 USD, or about 0.9% of the market price.

Importantly, the upper and lower bounds are computed in different markets and hence bound different objects. There is no mechanical relationship between them, and in particular, the upper bounds do not have to be systematically greater than the lower bounds. Nevertheless, such a relationship is apparent in the estimated means of these bounds as reported in the table.

Interpreting the mean bounds across markets. Continuing to examine Table 8, we now elaborate on the interpretation of the mean upper and lower bounds. Let us focus on panel A, featuring the bounds implied by the synthetic control classification scheme and the structural model. It may be tempting to consider 2.18 USD, the mean lower bound, as a lower bound on the mean efficiency gain, and 2.31 USD, the mean upper bound, as an upper bound on the same mean efficiency gain. Such an interpretation, however, would be incorrect.

To understand why, note that averaging the lower bounds (which yields an estimate of 2.18) delivers a consistent estimator of the mean lower bound on ME *conditional on an output expansion.* It does not estimate an unconditional mean lower bound across all affected markets. In fact, we may expect our 2.18 estimate to be higher than such an unconditional mean lower bound since it was computed while conditioning on the efficiency gain being large enough to offset the anti-competitive effect. Likewise, the 2.31 estimate does not reveal an unconditional mean upper bound. Eizenberg (2014) offers a bounded-support strategy to address this type of selection problems in the context of bounds that may only be conditionally computed. That strategy would require additional assumptions and is not pursued here.

Instead, we prefer to recognize the selection issue when interpreting our results. Taking stock, we have seen that output expansion obtained in at least 44%-49% of all treated markets, and in a majority of those where the sign of the output effect could be determined. And, on average *across these markets where output expansion obtained as a causal effect of the merger*, the lower bound on the efficiency gain was about 2 USD or 0.8% of the market price. We have also learned that such output expansions were more prevalent in larger markets. Taken together, these findings provide a rich, though partial, picture of the incidence of efficiency gains following the merger.

3.3.5 Robustness checks

We conclude our analysis by conducting several sensitivity checks. We consider several departures from our baseline classification procedure and examine whether our conclusions are robust to such variations.

The first variation on the classification process concerns the definition of the pre-merger period over which control markets are matched to affected markets. While the baseline analysis considered a 20-quarter period covering the years 2009-2013, this variation only uses the 16 quarters in 2009-2012. This exercise is meant to gauge the extent to which the choice of the pre-merger period affects the results.

A second variation concerns the comparison of affected and control markets: in the baseline approach we determined that the output effect was positive (resp., negative) if the post-merger output in the affected market exceeded that of the control market. Under the variation, we determine that the output effect was positive (resp., negative) if the *difference in output* between the pre-merger and the post-merger period in the affected market exceeded the corresponding difference in the control market. This variation is meant to mimic the familiar concept of a difference-in-differences.

Appendix Table A1 presents the results of both robustness checks. The first two rows simply reproduce the baseline results under NN (nearest-neighbor) and SC (synthetic control) from Table 4. As the table clearly shows, neither one of the two variations described above (the length of the pre-merger period, or the exact manner with which we compute the output effect across affected and control markets) affects the classification results in any substantial fashion.[29]

A third robustness check addresses a posential selection issues affecting our analyzed set of markets. While a total of 3,824 markets featured both carriers prior to the merger, we analyze the merger's effect in a subset (hereafter the "restricted set") of 1,014 such markets.[30] Recall that to obtain the restricted set of markets we excluded markets where either the identity of the non-merging legacy carriers, or the number of the low-cost carriers, had changed from the pre-merger

[29] It is not surprising that the first and second differences yield similar classifications since the matching process chooses the control market, actual or synthetic, to minimize the pre-merger output difference relative to the treated market. This match is generally very close, so that comparing the post-merger outputs or the pre-post output differences amounts to an almost identical classification.

[30] The total number of affected markets, 3,824, is slightly smaller than those presented in the most-right column of Table 1. This stems from a requirement for complete data over all 32 quarters in 2009-2016 which leads to the elimination of a small number of markets from the analysis.

to the post-merger period. In essence, we require the merger to be the only material change to market structure over the studied period.

To examine the importance of this restriction, we first compare the characteristics of the 1,014 markets in the restricted set to those in the complete set of 3,824 affected markets. Using all 3,824 markets as our sample, we regress five market characteristics on a dummy variable taking the value 1 for markets included in the restricted group of 1,014 markets, and zero otherwise. The regressions, presented in appendix Table A2, indicate that the restricted markets feature cheaper fares, shorter travel distances, and lower output relative to the complete set of affected markets, consistent with selection concerns.

To address these concerns more directly we next perform the market classification exercises for all 3,824 markets. Reassuringly, appendix Table A3 shows that these classifications are both qualitatively and quantitatively in line with our original ones. The table reports the results from a nearest-neighbor classification (double-majority rule) and a synthetic control classification (majority rule) performed in exactly the same fashion as in the baseline analysis where only the restricted sample of 1,014 markets were used (the results of which were reported in Table 4).

In the nearest-neighbor exercise, 49% of all markets are classified as experiencing an output expansion following the merger — exactly the same fraction as in the baseline analysis. The fraction of unclassified markets is 30%, which is larger than the unclassified fraction within the restricted set (13% per Table 4). This is not surprising since the complete set of markets includes cases where the market structure underwent considerable changes concomitantly with the merger, possibly driving more volatile variation in outcomes relative to control markets. The fraction of output-expansion markets out of those that were classified is 70% compared to 56% within the restricted set of 1,014 markets. The synthetic control analysis using the full sample of 3,824 affected markets again produces a similar overall fraction of output expansion markets (52%) to that obtain in the baseline case using the restricted markets (44%). As a fraction of classified markets, the output expansion share is 65%, exactly the same fraction obtained using the restricted sample.

To sum, while the restricted sample of 1,014 markets does appear to have somewhat different characteristics than the full sample of 3,824 affected markets, our results are not driven by such selection: the prevalence of markets experiencing output expansion is similar across the two sets of markets. Our focus on the restricted set of markets in the baseline analysis stems from theoretical considerations and from the desire to stay as closely as possible to the comparative static analyses of the model.

3.3.6 Additional issues

We finally turn to discuss additional interpretations for the results reported above. One interesting possibility is that the merged airline was able to enjoy efficiencies in some markets by virtue of shifting managerial attention and other resources from other markets. In this case, some markets may have cross-subsidized other markets as part of the merged entity's re-organization. Despite the high prevalence of output-expansion markets, this possibility cannot be directly ruled out by our results.

Identifying such cross-market effects may be quite difficult in the presence of heterogeneous oligopolistic markets. Among other issues, such an analysis would have to take into account exit and entry driven by unobserved equilibrium selection mechanisms.[31] Nonetheless, we are able to examine this issue in a somewhat cursory fashion by examining the effect of the merger on the total output across the affected markets. Once again we compare observed post-merger output levels in affected and control markets, but this time we focus on the magnitude of those differences and aggregate them across affected markets. By adding together the output expansions and contractions over our set of affected markets we can gauge the possibility that those cancel each other out in the aggregate.

Appendix Table A4 provides the results of this exercise, performed under different classification schemes (namely, the Nearest-neighbor method performed using four different sets of potential control groups described in earlier sections, and the synthetic control method) and different definitions of the pre-merger period similarly as in the robustness section above).

As shown in the most-right (fourth) column, the post-merger ratio of the total output in the affected markets to that in the control markets ranges between 0.97-1.4, depending on the classification specification. In fact, only one instance showed a result below 1. This indicates that, by and large, the merger led to an overall output expansion in the affected markets, rather than merely shifting output across such markets. We again caution that this is a somewhat basic check and ruling out cross-subsidies across markets would require additional analysis that is beyond our current scope.[32]

Finally, we note that the merger deal was preceded by a period where American Airlines' parent company filed for bankruptcy. There is evidence that airlines tend to shed capacity and lay off employees both before and during bankruptcies.[33] The question is whether efficiency gains afforded by the bankruptcy proceedings could be erroneously picked up in our analysis as merger-induced efficiencies.

Importantly, however, such capacity reductions would most likely manifest themselves in reductions of fixed costs and should, in fact, result in upward pressure on marginal costs. If anything, this should work against our finding that output expansion and marginal cost reduction obtained in many markets.

[31] See Ciliberto et al. (2021).

[32] For completeness, the third column of the table reports the same ratio computed for the pre-merger period. If the matching of affected to control markets were perfect, this ratio should be 1. This, however, happens only in two of the ten specifications, and otherwise the *total output* in the affected markets is somewhat larger than that in the control markets. Nonetheless, at the individual market level, the quality of the match is typically much higher, and the aggregate difference stems from a particular subset of markets.

[33] See Benmelech and Bergman (2008) and Benmelech et al. (2012).

4 Concluding remarks

This paper proposes and implements a methodology for estimating merger efficiencies from an *ex-post* perspective. In a first step, data on pre-merger and post-merger output (as well as some market characteristics) are used in a classification exercise that determines the identity of markets where output expansion, or output contraction, obtained following the merger. This step builds on statistical methods rather than on an econometric model. Linear probability models are then estimated to investigate which market characteristics are associated with a higher probability of a market expansion, enabling one to connect economic reasoning regarding the possible sources of efficiencies with the statistical exercise.

In a second step, a structural model of supply and demand is estimated. Along with the classification of the sign of the output effect of the merger, this model allows us to place a single bound on magnitude the efficiency gain in each market. This gain is modeled via a new measure of Merger Efficiencies to which we refer as ME. This measure captures the difference between the pre-merger minimum of the merging parties' marginal costs, and the merged entity's post-merger marginal cost, evaluated at the pre-merger joint output.

This approach places very minimal structure on cost functions, and, in particular, does not impose that they are constant in output as is typically done in many relevant studies. Nonetheless, the structure is otherwise restrictive. We assume that consumers view the product as homogeneous, model the market demand via a simple constant-elasticity specification, and prescribe that firms engage in Cournot competition both before, and after the merger (though our lower bounds are valid under coordinated merger effects as well). This structure enables us to compute simple bounds following classic theoretical results from Farrell and Shapiro (1990).

This trade-off should be carefully considered in applications of this framework. It is especially well-suited to industries where the homogeneous good and quantity competition assumptions provide a reasonable approximation of actual behavior. Of note, the New Empirical Industrial Organization (NEIO, Bresnahan (1989)) started with analyzing such cases, and then moved to consider product differentiation and other departures. We hope that our work forms the first step in an analogous process that achieves a more modest and specific task: learning about the nature and magnitude of efficiency gains, objects of interest that appear particularly elusive and yet draw substantial interest in antitrust studies as well as in practical discussions of merger policy. Indeed, future work should ideally focus on identifying additional sets of assumptions that yield analogous bounds on efficiency gains.

We implement our approach to the high-profile airline merger of American Airlines and US Airways. The role played by capacity choices in this industry suggests quantity competition as a relevant model for analyzing the supply side. In some agreement with other ex-post studies of this particular merger, the statistical analysis finds evidence of output expansion in a considerable share of the markets where both carriers competed head-to-head (and that satisfy some additional criteria, such as stability of the market structure except for the effect of the merger), and that output expansion was more prevalent in larger markets.

The structural model contributes the additional insight that, *conditional on an output expansion taking place*, the mean lower bound on the ME measure of marginal cost savings was slightly above 2USD, or, about 0.8% of the market price. The emphasis on the selection issue is important: this figure should not be viewed as revealing a lower bound on the unconditional mean efficiency across markets. Nonetheless, it is informative, in this case, with regard to the efficiency gains in a large set of markets. Future work may also focus on translating the conditional mean bounds into unconditional ones. This may be possible via a moment inequality approach subject to additional assumptions.

Tables and Figures

Table 1: Merging parties' presence across markets, 2012

A. Total market counts

Quarter	# markets	AA markets	US markets	Only AA	Only US	Both present
1	7,911	5,203	5,715	1,379	1,891	3,824
2	7,911	5,147	6,095	1,077	2,025	4,070
3	7,911	5,107	6,051	1,106	2,050	4,001
4	7,911	5,121	5,990	1,132	2,001	3,989

B. Percentage of total number of markets in quarter

Quarter	AA markets	US markets	Only AA	Only US	Both present
1	66%	72%	14%	24%	48%
2	65%	77%	14%	26%	51%
3	65%	76%	14%	26%	51%
4	65%	76%	14%	25%	50%

Notes: The total number of markets (directional airport pairs, top 100 airports) where neither one, one or both of the merging parties were present across the four quarters of 2012, i.e., pre-merger. Panel A reports total market counts while Panel B shows the same information in percentage terms out of the total number of markets. Source: authors' calculations using USDOT data.

Table 2: Market-level descriptive statistics

Variable	Mean	Standard deviation	Min	Max
Market price (USD)	257	81	35	731
Passengers (X)	974	1,846	10	26,918
Market population measure	2,585,612	2,172,205	104,271	16,100,000
Nonstop miles	1,291	894	152	5,136

Notes: Statistics were computed over the 32,140 market-quarter pairs in 2012. The population measure is the geometric mean of the populations of the MSAs at the market's two endpoints. Source: authors' calculations using USDOT data.

Table 3: Criteria defining the affected and control group markets

	Treatment	Control 1	Control 2	Control 3	Control 4
# merging parties present	2	0	0	< 2	≤ 2
Constant legacy carriers	✓	✓	x	x	x
Constant # of low-cost carriers	✓	✓	x	x	x
Number of markets	1,014	146	777	4,044	6,854

Notes: See text.

Table 4: Market-level classification of the sign of the output effect

	Synthetic control	Nearest-Neighbor	Consensus
Markets with output expansion	494	451	299
Markets with output contraction	384	238	148
Markets where the effect was not determined	136	325	567
Total number of affected markets	1,014	1,014	1,014
Output expansion fraction (out of those classified)	56%	65%	67%
Output contraction fraction (out of those classified)	44%	35%	33%
Output expansion fraction (out of total affected)	49%	44%	29%
Output contraction fraction (out of total affected)	38%	23%	15%
Unclassified fraction (out of total affected)	13%	32%	56%

Notes: The table presents statistics regarding the classification of markets according to the sign of the merger's effect on output. Classification is attempted over 1,014 affected markets where both merging parties were present prior to the merger. Synthetic control classifications are performed based on four quarters of post-merger activity. A majority of classifications among these four predictions is necessary for a classification to be determined (see text). The nearest-neighbor approach implements a double-majority rule over both these four quarters and four donor pools of control markets. Consensus classifications are made only when the Synthetic control and the nearest-neighbor rules result in an identical classification decision. Source: authors' calculations using USDOT data.

Table 5: Linear probability models for a pro-competitive classification

A. Synthetic control (majority)	(1)	(2)	(3)	(4)
network advantage	0.018	0.018	0.014	0.014
	(0.014)	(0.014)	(0.014)	(0.014)
Population (M)	.037***	0.036***		
	(0.011)	(0.010)		
US hub	0.067		0.088	
	(0.058)		(0.058)	
AA hub	0.052		0.132***	
	(0.050)		(0.044)	
Merge hub		0.058		0.117***
		(0.040)		(0.037)
Constant	0.437***	0.438***	0.514***	0.514***
	(0.034)	(0.033)	(0.026)	(0.026)
Observations	878	878	878	878
R2	0.03	0.03	0.01	0.01

B. Nearest-neighbor (double majority)	(1)	(2)	(3)	(4)
network advantage	-0.007	-0.006	-0.008	-0.008
	(0.017)	(0.017)	(0.014)	(0.017)
Population (M)	0.031**	0.027**		
	(0.012)	(0.012)		
US hub	0.141**		0.160**	
	(0.067)		(0.067)	
AA hub	0.041		0.105**	
	(0.054)		(0.047)	
Merge hub		0.078*		0.122***
		(0.045)		(0.040)
Constant	0.573***	0.579***	0.635***	0.634***
	(0.038)	(0.038)	(0.030)	(0.030)
Observations	689	689	689	689
R2	0.02	0.02	0.02	0.01

Notes: Linear probability models where an indicator for a pro-competitive classification serves as the dependent variable. Panels A and B refer to classifications performed via the synthetic control and the nearest-neighbor classification schemes, respectively. Standard errors appear in parentheses below the coefficients. The variable "network advantage" is the ratio of markets served out of the origin to the number of markets served out of the destination airport (see text). Population is the geometric mean of the endpoints' populations in millions. US hub, AA hub and Merge hub are dummy variables indicating that either one of the market endpoints is a hub of US Airways, a hub of American Airlines, or a hub of either one of them, respectively. ***, ** and * indicate significance at the 1%, 5% and 10% levels. Source: author's calculations using USDOT data.

Table 6: Passenger demand

	OLS		2SLS	
	Coefficient	SE	Coefficient	SE
Demand elasticity (η)	-2.6044	0.0253	-5.5847	0.1438
Demand level shifters (α)				
origin markets	0.0006	0.0000	0.0006	0.0000
population	0.1876	0.0028	0.0820	0.0055
nonstop miles (000)	0.3334	0.0085	1.0186	0.0352
Q1 dummy	-0.1363	0.0163	-0.1160	0.0273
Q2 dummy	0.0783	0.0164	0.1924	0.0277
Q3 dummy	-0.0462	0.0163	-0.0809	0.0269
Constant	20.4406	0.1363	36.2076	0.7577
Observations (2012 market-quarter pairs)	31,590		17,614	

Notes: Estimates of the demand equation (7). Population (geometric mean of populations of the MSAs at the market's two endpoints) was divided by 1M. Nonstop miles (in thousands) represent the distance flown in a direct flight between the two endpoints. Origin markets: the number of markets served out of the origin airport. The two instrumental variables used are (1) a cost measure (a passenger-mile weighted average of fuel cost per gallon) and (2) the ratio of the number of markets served out of the destination airport to the number of markets served out of the origin airport. See text. Source: authors' calculations using USDOT data.

Table 7: Carrier pre-merger marginal costs

		Mean	Min	Max
A. Carrier marginal costs across markets (USD)				
		Mean	Min	Max
US Airways		252.13	91.86	629.23
American Airlines		250.80	90.17	618.17
B. Carrier markup (% of market price)				
		Mean	Min	Max
US Airways		0.024	0.000	0.174
American Airlines		0.029	0.000	0.179
C. Cost difference (USD)				
		Mean	Min	Max
US-AA		1.334	-67.635	65.242
Number of markets:	15,884			

Notes: Pre-merger marginal costs and markup statistics computed for each of the two merging carriers over 15,884 markets (2012 market-quarter pairs where both carriers were present). Panel C presents statistics regarding the difference between the carriers pre-merger marginal costs. Source: authors' calculations using USDOT data.

Table 8: Distribution of ME bounds

A. Synthetic control (majority)

	Mean	Min	Max	# markets
Lower bound (USD)	2.18	0.01	20.07	494
Lower bound as a fraction of market price	0.80%	0.01%	6.56%	494
Upper bound (USD)	2.31	0.01	27.00	384
Upper bound as fraction of market price	0.88%	0.00%	6.87%	384

B. Nearest neighbor (double majority)

	Mean	Min	Max	# markets
Lower bound (USD)	2.16	0.02	18.78	451
Lower bound as a fraction of market price	0.83%	0.01%	6.56%	451
Upper bound (USD)	2.43	0.01	17.82	238
Upper bound as fraction of market price	0.94%	1.13%	6.87%	238

C. Consensus

	Mean	Min	Max	# markets
Lower bound (USD)	2.12	0.02	18.78	299
Lower bound as a fraction of market price	0.82%	0.01%	6.56%	299
Upper bound (USD)	2.26	0.01	17.82	238
Upper bound as fraction of market price	0.91%	1.16%	6.87%	238

Notes: The distribution of the bounds on the Merger Efficiencies (ME) measure is presented given each of the three classification schemes (synthetic control, nearest neighbor, and their consensus — see notes to Table 5). Lower bounds are available in markets classified as experiencing an output expansion, and upper bounds are available in markets classified as experiencing an output contraction following the merger. Source: authors' calculations using USDOT data.

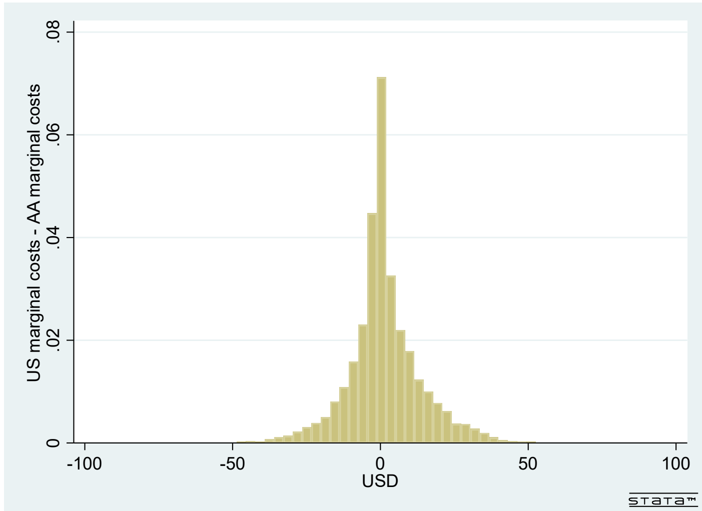

Figure 1: A histogram of the pre-merger difference between the marginal costs of US and AA across markets

A Additional Tables

Table A1: Classification robustness across methods

Method	Differences	Training period	Affected markets				Share of markets		
			Total	E	C	NA	E	C	NA
NN	1st	20	1,014	451	238	325	0.44	0.23	0.32
SC	1st	20	1,014	494	384	136	0.49	0.38	0.13
NN	2nd	20	1,014	436	252	326	0.43	0.25	0.32
SC	2nd	20	1,014	496	378	140	0.49	0.37	0.14
NN	1st	16	1,014	479	246	289	0.47	0.24	0.29
SC	1st	16	1,014	474	431	109	0.47	0.43	0.11
NN	2nd	16	1,014	470	250	294	0.46	0.25	0.29
SC	2nd	16	1,014	474	431	109	0.47	0.43	0.11

Notes: The table presents the classification of markets into E (output expansion), C (output contraction) and NA (output effect not classified) using different variations on the original classification scheme. NN and SC stand for nearest-neighbor (double-majority) and synthetic control (majority) methods (see text). The first two rows reproduce the baseline results from the first two columns of Table 4. First and second differences refer to the baseline approach (comparing post-merger output in the control and affected markets) versus an approach that compares the difference in output from the pre-merger to the post-merger periods across these markets. The training period refers to the number of pre-merger quarters over which control markets were matched to affected ones. Source: authors' calculations using USDOT data.

Table A2: Characteristics of the restricted set of affected markets

	(1)	(2)	(3)	(4)	(5)
Dependent variable	Population	Distance	Output	Revenue	Price
Restricted set indicator	-840,283	-346	-7,213	-1,540,322	-6.973
	(81,566)	(32.11)	(804.6)	(153,535)	(2.518)
Constant	3,163,412	1,627	16,194	3,321,987	240.8
	(42,002)	(16.53)	(414.3)	(79,062)	(1.297)
Observations	3,824	3,824	3,824	3,824	3,824
R2	0.027	0.029	0.021	0.026	0.002

Notes: Regressions where the dependent variable is a market-level characteristic, and the explanatory variable is an indicator for inclusion in the restricted set of 1,014 affected markets. The sample includes all 3,824 markets where both merging parties competed prior to the merger. The market characteristics are population (the geometric mean of the population at both endpoints), distance (nonstop miles), output (number of quarterly passengers), market revenue (USD) and price (USD). Source: authors' calculations using USDOT data.

Table A3: Output effect classification among all affected markets

Classification method	NN	SC
Markets with output expansion	1,864	1,995
Markets with output contraction	807	1,094
Markets with an unclassified output effect	1,153	735
Total number of markets	3,824	3,824
Output expansion fraction (out of those classified)	0.70	0.65
Output contraction fraction (out of those classified)	0.30	0.35
Output expansion fraction (out of total affected)	0.49	0.52
Output contraction fraction (out of total affected)	0.21	0.29
Unclassified fraction (out of total affected)	0.30	0.19

Notes: The sample includes all 3,824 markets where both merging parties competed prior to the merger. Source: authors' calculations using USDOT data.

Table A4: Aggregate output effects

Classification	Training quarters	Pre-merger Q(affected)/Q(control)	Post-merger Q(affected)/Q(control)
KNN Control group 1	20	1.07	1.42
KNN Control group 2	20	1.07	1.09
KNN Control group 3	20	1.07	1.13
KNN Control group 4	20	1.00	0.97
Synthetic control group	20	1.06	1.21
KNN Control group 1	16	1.08	1.42
KNN Control group 2	16	1.08	1.14
KNN Control group 3	16	1.07	1.13
KNN Control group 4	16	1.00	0.96
Synthetic control group	16	1.06	1.22

Notes: The total output in the affected markets divided by the total output in the control markets, before the merger (third column) and after the merger (fourth column), given different classification schemes detailed in the first column and different definitions for the pre-merger period reported in the second column. Source: authors' calculations using USDOT data.

B Proof of analytical results

Both the results and the proofs follow Farrell and Shapiro (1990) exactly.

Proof of Lemma 1. Let Δx_1 and ΔX represent the change to firm 1's output, and in total output, respectively. Any firm $i \neq 1$ satisfies its Cournot first-order conditions implying, by (5), $dx_i = -\lambda_i dX$. Recall that by definition, $y_1 \equiv \sum_{i\neq 1} x_i$, and hence:

$$dy_1 \equiv \sum_{i\neq 1} dx_i = -dX \sum_{i\neq 1} \lambda_i.$$

Adding dx_1 to both sides of the above equation yields:

$$dy_1 + dx_1 = -dX \sum_{i\neq 1} \lambda_i + dx_1,$$

but since $dy_1 + dx_1 = dX$,

$$dX(1 + \sum_{i\neq 1} \lambda_i) = dx_1.$$

And as $\lambda_i > 0$ for all i, we obtain that dX and dx_1 are of the same sign, but dX is smaller in magnitude. Since this result holds for infinitesimal changes, it also holds for Δx_1 and ΔX. $\square$

Proof of Theorem 1. Starting at the pre-merger Cournot equilibrium, we allow firms 1 and 2 to merge into a new entity M. Denote pre-merger quantities by $\overline{x}_i$ for each firm i. Holding quantities for non-merging parties fixed at their pre-merger levels, we derive a condition for firm M's post-merger quantity to fall short of $\overline{x}_1 + \overline{x}_2 \equiv \overline{X}_M$. By Lemma 1, total quantity would then also fall.

For M to produce less than $\overline{X}_M$, its marginal cost should be higher than its marginal revenue at that quantity: $c_x^M(\overline{X}_M) > p(\overline{X}) + \overline{X}_M \cdot p'(\overline{X})$, suggesting an upper bound on its markup there:

$$p(\overline{X}) - c_x^M(\overline{X}_M) < -\overline{X}_M \cdot p'(\overline{X}).$$

This upper bound also turns out to be the sum of the merging parties' pre-merger markups:

$$\sum_{i=1,2} \left(p(\overline{X}) - c_x^i(\overline{x}_i) \right) = \sum_{i=1,2} \left(-\overline{x}_i p'(\overline{X}) \right) = -p'(\overline{X}) \sum_{i=1,2} \overline{x}_i = -p'(\overline{X})\overline{X}_M.$$

The merger would therefore reduce the aggregate quantity if and only if

$$p(\overline{X}) - c_X^M(\overline{X}_M) > (p(\overline{X}) - c_x^1(\overline{x}_1)) + (p(\overline{X}) - c_x^2(\overline{x}_2)).$$

$\square$

C Data appendix

This appendix section provides details regarding data cleaning and other relevant details. The data cleaning choices described below are in line with recent Industrial Organization studies of the Airline industry.

T-100 database processing. Observations that do not belong to the civilian domestic market are removed. In particular, observations not pertaining to passenger aircraft configurations were removed. About 87% of observations belong in the F or L classes pertaining to civilian passenger classes and only those classes were retained. In a negligible number of observations a zero number of seats is reported and those are also dropped.

Following Park (2020b) and Li et al. (2022) only markets connecting the largest 100 airports are considered. To this end, airports are first ranked by their total passenger boarding counts in 2012. Then, only segments connecting those 100 airports are retained.

Observations from the T-100 database were then aggregated up to combinations of year-quarter-segment (origin-destination)-carrier and merged onto the DB1B data which processing is described next.

DB1B database processing. Following Dana and Orlov (2014) itineraries with bulk fares are dropped as those are typically sold via travel agents. Still following that study, tickets for which the price data is flagged by the DOT as lacking credibility, as well as tickets costing less than 12.5 USD or more than 1,250 USD are also dropped from the sample (Park (2020a)).

Tickets that have more than one connection or more than two trip breaks are dropped. Only directional tickets departing from and arriving the top-100 airports discussed above are kept. Markets (directional origin-destination airport pairs) with less than 200 passengers in any quarter are also dropped (following Ciliberto and Williams (2014), Park (2020b)), as are markets for which the non-stop distance is less than 150 miles (Ciliberto et al. (2021)).

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Affeldt, P., Duso, T., Gugler, K. P., and Piechucka, J. (2021). Assessing eu merger control through compensating efficiencies.

Ali, R. M. (2022). Effect of airport slots in competition and antitrust policy: Evidence from a recent merger. *Essays in Competition Economics*.

Ashenfelter, O., Hosken, D., and Weinberg, M. (2014). Did robert bork understate the competitive impact of mergers? evidence from consummated mergers. *The Journal of Law and Economics*, 57(S3):S67–S100.

Benmelech, E. and Bergman, N. K. (2008). Liquidation values and the credibility of financial contract renegotiation: Evidence from us airlines. *The Quarterly Journal of Economics*, 123(4):1635–1677.

Benmelech, E., Bergman, N. K., and Enriquez, R. J. (2012). Negotiating with labor under financial distress. *The Review of Corporate Finance Studies*, 1(1):28–67.

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

Björnerstedt, J. and Verboven, F. (2016). Does merger simulation work? evidence from the swedish analgesics market. *American Economic Journal: Applied Economics*, 8(3):125–64.

Bokhari, F. (2023). Retrospective merger evaluation: Gsk-pfizer consumer health. Technical report, Working paper.

Bontemps, C., Gualdani, C., and Remmy, K. (2023). Price competition and endogenous product choice in networks: Evidence from the us airline industry'. Technical report, Working Paper.

Borenstein, S. (1990). Airline mergers, airport dominance, and market power. *The American Economic Review*, 80(2):400–404.

Brancaccio, G., Kalouptsidi, M., and Papageorgiou, T. (2020). Geography, transportation, and endogenous trade costs. *Econometrica*, 88(2):657–691.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization*, 2:1011–1057.

Carlton, D. W., Landes, W. M., and Posner, R. A. (1980). Benefits and costs of airline mergers: A case study. *The Bell Journal of Economics*, pages 65–83.

Chen, Y. (2021). Network structure and efficiency gains from mergers: Evidence from us freight railroads. Technical report, Working paper.

Ciliberto, F., Murry, C., and Tamer, E. (2021). Market structure and competition in airline markets. *Journal of Political Economy*, 129(11):2995–3038.

Ciliberto, F. and Williams, J. W. (2014). Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry. *The RAND Journal of Economics*, 45(4):764–791.

Dana, J. D. and Orlov, E. (2014). Internet penetration and capacity utilization in the us airline industry. *American Economic Journal: Microeconomics*, 6(4):106–137.

Das, S. (2019). Effect of merger on market price and product quality: American and us airways. *Review of Industrial Organization*, 55(3):339–374.

Demirer, M. and Karaduman, O. (2022). Do mergers and acquisitions improve efficiency: Evidence from power plants. Technical report, Working Paper.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *Review of Economic Studies*, 81(3):1003–1045.

Farrell, J. and Baker, J. B. (2021). Natural oligopoly responses, repeated games, and coordinated effects in merger analysis: A perspective and research agenda. *Review of Industrial Organization*, 58(1):103–141.

Farrell, J. and Shapiro, C. (1990). Horizontal mergers: an equilibrium analysis. *The American Economic Review*, pages 107–126.

Fix, E. and Hodges, J. L. (1951). Nonparametric discrimination: consistency properties. *Randolph Field, Texas, Project*, pages 21–49.

Khezrimotlagh, D., Kaffash, S., and Zhu, J. (2022). Us airline mergers' performance and productivity change. *Journal of Air Transport Management*, 102:102226.

Khmelnitskaya, E., Marshall, G., and Orr, S. (2024). Identifying scale and scope economies using demand-side data. *Available at SSRN 4757085*.

Kim, S. J. and Park, Y. (2023). Examining the coordinated effects of the aa/usair merger.

Le, H. B. (2019). An ex post analysis of the us airways/american airlines merger. *Review of Economic Analysis*, 11(3):383–398.

Li, S., Mazur, J., Park, Y., Roberts, J., Sweeting, A., and Zhang, J. (2022). Repositioning and market power after airline mergers. *The RAND Journal of Economics*, 53(1):166–199.

Moss, D. L. (2013). Delivering the benefits? efficiencies and airline mergers. *Efficiencies and Airline Mergers (November 21, 2013).*

Olley, G. and Town, R. (2018). End of an era: the american airlines-us airways merger. *The Antitrust Revolution: Economics, Competition, and Policy*, pages 448–470.

Park, Y. (2020a). *Essays on Industrial Organization*. PhD thesis, University of Maryland, College Park.

Park, Y. (2020b). Structural remedies in network industries: An assessment of slot divestitures in the american airlines/us airways merger. *US Airways Merger (October 12, 2020).*

Rose, N. L. and Shapiro, C. (2022). What next for the horizontal merger guidelines. *Antitrust Magazine, Spring*.

Turner, D. (2021). Coordinated effects in the american airlines-us airways merger. *Available at SSRN 3917112*.

Vaze, V., Luo, T., and Harder, R. (2017). Impacts of airline mergers on passenger welfare. *Transportation Research Part E: Logistics and Transportation Review*, 101:130–154.

Werden, G. J., Joskow, A. S., and Johnson, R. L. (1991). The effects of mergers on price and output: Two case studies from the airline industry. *managerial and Decision economics*, 12(5):341–352.

White, L. J. (1987). Antitrust and merger policy: a review and critique. *Journal of Economic Perspectives*, 1(2):13–22.

Williamson, O. E. (1968). Economies as an antitrust defense: The welfare tradeoffs. *The American Economic Review*, 58(1):18–36.