What's the Difference?
Measuring the Effect of Mergers in the Airline Industry

[Preliminary and Incomplete: Please do not Circulate]

Reed Orchinik *
Swarthmore College

Marc Remer†
Swarthmore College

January 24, 2020

Abstract

We analyze the effect of the four most recent US airline mergers using three retrospective techniques: standard difference-in-differences regression, synthetic control, and nearest neighbor matching. In doing so, we compare the performance of the most recently developed merger-retrospective methodologies. Each method compares an outcome of interest, such as price, in a set of "treated" markets where a merger occurred to a set of "control" markets that are unaffected by the merger. The three methodologies, however, differ in how they select and weight control markets. We find that the three methods do not always align in the direction or statistical significance of the effect of the merger. Furthermore, synthetic control and nearest neighbor matching are sensitive to the set of characteristics used to select control markets. Overall, we find that the American Airlines/USAir merger lowered prices, whereas the United/Continental merger increased prices. Even so, these findings are not robust across methodologies or specifications. We are unable to draw any consistent conclusions about the Southwest/AirTran or Delta/Northwest mergers.

*Swarthmore College, 500 College Ave., Swarthmore, PA, 19081. Email: reed.orchinik@gmail.com.
†Swarthmore College, 500 College Ave., Swarthmore, PA, 19081. Email: mremer1@swarthmore.edu.

1 Introduction

It is well understood that a merger between competitors can lead to the exercise of market power and a loss of total welfare. However, a horizontal merger may also generate marginal cost efficiencies, which can result in lower prices and increased welfare. Preventing the former mergers and allowing the latter is a central goal of antitrust enforcement. From an empirical standpoint, economists have developed two general sets of tools to help distinguish between pro and anti-competitive mergers. The first set of tools are prospective, such as demand estimation and merger simulation, which are used to forecast the potential competitive impact of a merger. The second set are retrospective, which are used to study the actual impact of a previously consummated merger.

In this article, we analyze three merger retrospective techniques: standard difference-in-differences regression (DID), synthetic control, and nearest neighbor matching. These estimation methodologies are informed by the same underlying logic: compare an outcome of interest, such as price, in a set of "treated" markets where a merger occurred to a set of "control" markets that are unaffected by the merger. The primary distinction between the techniques is how they select control markets. DID estimation places equal weight on each observation in a pre-specified control group. The synthetic control method weights each control observation such that the weighted-average approximates pre-merger characteristics of the treatment group. Nearest neighbor matching uses a selection criteria to match each treated unit to a subset of control units and then equally weights each control unit in the subset. Using each of these three methods, we analyze the competitive effects of the four most recent mergers in the United States airline industry. We demonstrate that techniques recently developed to create more suitable control groups are sensitive to choices made by the researcher, and that in practice DID, synthetic control, and nearest neighbor matching should not be expected to produce similar results. The analysis makes a general contribution to the applied microeconomics literature, which frequently uses one of these three estimators to assess the impact of policy interventions and natural experiments.

In April, 2019, the Federal Trade Commission held public hearings discussing the importance of merger retrospectives.[1] In the hearings, the extent to which merger retrospectives can be used to understand the outcomes of previous mergers and generalized to prospective mergers was debated. While the exact value of merger retrospectives is an open question, there is no doubt that accurately measuring the impact of horizontal mergers will greatly benefit antitrust policy. We contribute to this goal by implementing and assessing three

[1] https://www.ftc.gov/news-events/events-calendar/ftc-hearing-14-merger-retrospectives

leading merger retrospective techniques on real-world consummated mergers. We find that in one instance prices likely increased due to the merger (United/Continental) and in another prices decreased (American/USAir). In the other two, we find no robust impact across methods from the merger on prices charged to consumers. One common finding across all mergers is that the estimated price effects are sensitive to either the applied methodology or reasonable choices made within a given methodology. In some instances, the sign of the effect changes across econometric techniques.

The accuracy of merger prediction tools is often gauged by how closely they align with the results of the DID estimates of a consummated merger. A number of articles first estimate a structural demand model using pre-merger data and then simulate a merger. The accuracy of the simulations are then measured relative to the DID estimates.[2] This analysis assumes that the merger retrospective accurately measures the true effect of the merger. We demonstrate that this may currently place too high of a burden on commonly used retrospective techniques to accurately estimate merger effects. The three retrospective methodologies analyzed in this article may produce estimates of greatly different magnitude, and even sign. We discuss the relative merits of each methodology and settings in which one may be favored over another. As best practice, we encourage researchers to implement either nearest neighbor or synthetic control alongside of standard DID when conducting a merger retrospective analysis.

The United States airline industry presents an ideal setting within which to perform merger retrospectives. Since the industry was deregulated in the 1970's there has been a steady stream of consolidation. In the past 15 years, horizontal mergers have left the industry with only three legacy airlines (American, Delta, and United) and one predominant low-cost carrier (Southwest). In its lawsuit to block the American Airlines/USAir merger, the US Department of Justice noted that this merger wave has resulted in higher fares and reduced service between city pairs.[3] More recently, the industry has been under investigation by the DOJ for potentially coordinating capacity discipline in an effort to increase prices.[4] Given the recent consolidation and airlines' importance to the overall economy, it is imperative to understand precisely if mergers in the industry have resulted in price increases and lower consumer welfare.

From a methodological perspective, the airline industry is well suited to apply DID-style

[2] See, for example, Peters (2006), Weinberg and Hosken (2013). Garmon (2017) uses the synthetic control method to form a baseline estimate of merger price effects.

[3] See paragraph 35 of DOJ's complaint against American Airlines and USAir, which is available at `https://www.justice.gov/atr/case-document/file/514531/download`.

[4] See Ciliberto, Watkins, and Williams (2019) and cites therein.

estimators to measure the impact of a horizontal merger. To analyze the industry, antitrust practitioners and researchers typically specify each route to be a unique market. Given the hubbing system utilized by most airlines, a merger results in a relatively small number of "treated" overlap markets and a large number of potential "control" markets that are not directly affected by the merger. To measure the impact of the merger on an outcome of interest, such as price, we can compare the difference in prices between the treated and control routes, both before and after the merger. A central challenge to this analysis is identifying a set of control routes that serve as a proxy for what prices would have been on the treated routes, but-for the merger. In the mergers we analyze, there are between 5 and 13 routes affected by the merger and more than 25,000 potential control routes. It is therefore of critical importance to select an appropriate subset of control routes in order to accurately assess the impact of the merger.

Additionally, mergers provide a unique context for policy evaluation where the policy in question has a direct effect on a relatively well-understood demand system. In many contexts where DID-style estimation is well-suited, the relationship between the outcome variable of interest and unobservables is weakly understood. Here, price is codetermined with passengers in a system of simultaneous equations. However, passengers cannot be used as a matching or control variable in estimation and must remain unobservable to the matching algorithm. Passenger estimation, thus, provides a compelling means to test the fit of the control group in DID-style estimation on unobservables, one of the supposed benefits of matching algorithms.

We analyze the four most recent mergers in the US airline industry: Delta/Northwest, United/Continental, Southwest/AirTran, and American Airlines/USAir. The primary focus of the analysis is the extent to which the price of a round-trip ticket changed due to the merger. As a secondary component, we analyze the impact on the number of passengers flown on overlap routes. If price and passenger counts move in the opposite direction it suggests that the merger either led to an exercise of market power (prices increase/passengers decrease) or marginal cost efficiencies (prices decrease/passengers increase). Conversely, if they move in the same direction then it could be due to merger-induced quality changes (e.g. improved service) or unrelated demand shocks.

We find that airline mergers have a heterogeneous effect on prices. The American/USAir merger is estimated to generate lower prices and more passengers on routes in which the airlines competed head-to-head prior to the merger. Conversely, the United/Continental merger generated higher prices on overlapping routes. The impacts of the Delta/Northwest and Southwest/AirTran mergers are much less clear. Indeed, we estimate that the Southwest/AirTran merger significantly increased prices (7%-10%) using DID regression, but significantly decreased prices (-9% to -19%) using nearest neighbor matching. These findings highlight that mergers within an industry should be evaluated on a case-by-case basis, and there is no a priori reason to expect observationally similar mergers to generate the same outcome.

Even for the two mergers where there is general agreement across estimation methods, reasonable and minor changes to the routines can alter statistical significance, the magnitude of the effect, or the sign of the estimated coefficients. For example, adding proxies for pre-merger competition to the matching criteria for nearest neighbor or synthetic control changes the estimated effect of American/USAir merger from negative to positive. Adding the same competition variables to the United/Continental synthetic control and nearest neighbor routines kills the statistical significance of the estimates. This occurs in the synthetic control despite the fact that the competition variables receive almost no weight in constructing the synthetic control routes. These findings demonstrate that researcher degrees of freedom may influence the findings of ex-post policy analysis, even using methods developed to minimize the need for researchers to make such choices.

The remaining paper is organized as follows. Section 2 reviews the related literature, Section 3 details the three estimation methodologies. Section 4 summarizes the data, details pre-merger pricing trends, and how well different methodologies fit the pre-merger data. Section 5 presents the estimation results, and Section 6 discusses the relative value of each estimation methodology.

2 Literature Review

There is a rich literature exploring the extent to which non-experimental estimators can accurately estimate treatment effects when the effects are known from experimental data. LaLonde (1986) evaluated the accuracy of commonly used parametric regression techniques on experimental data and found that the results could not be replicated without bias. Imbens and Angrist (1994) derive conditions under which local average treatment effects can be identified from non-experimental data. Dehejia and Wahba (2002) use the same experimental data as LaLonde (1986) and find that propensity score matching estimators are able to replicate the experimental findings with low bias. On the other hand, again using LaLonde's experimental data, Smith and Todd (2005) find that propensity score matching estimators are sensitive to the variables used to generate propensity scores and which sub-samples of the data are used for estimation. Smith and Todd (2005) find that DID matching estimators were best able to produce estimates consistent with the experimental findings. We contribute to this literature by evaluating the performance of the leading DID estimators. However, our analysis is distinct in one important dimension - we evaluate the methods using non-experimentally derived data. We then ask if the most common implementations of these estimators yield similar results. Specifically, we implement DID regression (Imbens and Angrist (1994)), synthetic control (Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010)), and nearest-neighbor mathching (Abadie and Imbens, 2006, 2011).

There is a growing literature that uses DID regressions to estimate the competitive impact of consummated mergers. For example, the price effect of horizontal mergers between health insurance providers (Dafny, Duggan, and Ramanarayanan, 2012), beer companies (Ashenfelter, Hosken, and Weinberg, 2015), and academic journals (McCabe, 2002) have all been estimated using DID. In a closely related literature, the DID estimated effects of a consummated merger are compared to the predictions of a structural analysis estimated on pre-merger data. Using this approach, the airlines industry (Peters, 2006), hospital markets (Garmon, 2017), motor oil and maple syrup products (Weinberg and Hosken, 2013), and retail gasoline markets (Houde, 2012), among others, have been analyzed. Ashenfelter, Hosken, and Weinberg (2014) perform a comprehensive review of the merger retrospective literature, and there are two notable findings. First, almost all retrospectives use standard DID regression, and mergers in oligopoly markets tend to generate higher prices. We contribute to this literature by estimating the effects of four separate mergers in the airline industry, and comparing the estimates across several different methodologies.

There is a long history in economics of analyzing the US airline industry, in general, and horizontal mergers within the industry. Following the Airline Deregulation Act of 1978, there was a large wave of industry consolidation. Borenstein (1990) evaluated the 1986 mergers of TWA/Ozark and Northwest/Republic by comparing prices on overlap routes to those on unaffected and similar length routes, and finds that prices were unchanged in the former merger and increased by almost 10% in the latter. Using a set of control routes distinct from Borenstein (1990), Morrison (1996) analyzes the same two mergers (in addition to USAir/Piedmont) and finds that over a similar time frame used in Borenstein (1990) TWA/Ozark prices increased and Northwest/Republic decreased. Kim and Singal (1993) use standard DID regressions to estimate the price effect of 14 airline mergers that occurred from 1985 to 1988, and find that market power effects outweighed efficiency gains. More recently, Huschelrath and Muller (2013) analyzes six US airline mergers from the late 1990's up through the Delta/Northwest merger in 2008 and finds that prices increased by 3%-6%, depending upon the post-merger time frame considered.

In terms of mergers analyzed, Carlton, Israel, MacSwain, and Orlov (2019) is most closely related to this article. It uses DID to estimate the effects of the three most recent US legacy airline mergers: Delta/Northwest, United/Continental, and American/USAir. Following the advice of Mehta and Miller (2012), Carlton et al. (2019) take seriously the potential sensitivity of estimates to the assumed control group. The analysis includes a large number of robustness checks wherein the merger effects are measured relative to different sets of control routes pre-selected by the authors. The study finds that Delta/Northwest and American/USAir led to significant price decreases on overlap routes, and there was no signficant change in prices due to the United/Continental merger.

The empirical analysis in this article differs from Carlton et al. (2019) along a number of important dimensions. First, we implement two additional empirical methodologies that allow the data to guide the selection of control routes, rather than relying exclusively on our own judgment. While synthetic control or matching estimators are frequently utilized in applied microeconomics, they are seldom employed in industrial organization.[5] Second, a primary aim of this article is to compare the DID estimates to the newer techniques that weight control units to more closely match treated units. Our results demonstrate that synthetic control and nearest neighbor estimators may lead to very different findings than standard DID regressions. Finally, we also analyze a recent merger between low cost carriers (Soutwest/AirTran), as these may have a market impact unique from legacy mergers.

3 Methodology Overview

To analyze the effect of airlines mergers on prices and quantity, we implement three separate methodologies: DID, synthetic control, and nearest neighbor matching. The aim of each technique is to infer the effect of the merger by comparing an outcome on overlap routes relative to routes unaffected by the merger. The primary distinction between methods is how they select and weight unaffected routes. In this section, we detail each estimation routine.

3.1 Difference-in-differences estimation

DID regression is the most common technique used in merger retrospectives that involve multiple geographic markets differentially affected by the merger. In the airline industry,

[5] A notable exception is Deryugina, MacKay, and Reif (2019), which uses nearest neighbor matching to estimate the effect of electricity regulations on market demand.

a market is often defined as a route between specific origin and destination airports.[6] We restrict our attention to non-stop routes, as this is the typical market definition employed in the US Department of Justice's antitrust analysis and makes our results comparable to other US airline merger retrospectives (e.g. Werden, Joskow, and Johnson (1991) and Carlton, Israel, MacSwain, and Orlov (2019)). Routes affected by the merger are those where the merging airlines both operate regular non-stop service prior to the merger.

To identify the causal effect of the merger, OLS regression is employed along with important assumptions about the evolution of outcomes in "treated" and "untreated" markets. Specifically, an equation of the following form is estimated:

$$y_{mt} = \alpha + \beta D_{mt} + \gamma_m + \lambda_t + \epsilon_{mt}. \tag{1}$$

Equation (1) is specified at the market-level, m, and is estimated using panel data that varies over time, t.[7] Here, y_{mt} is an outcome of interest, γ_m is a market fixed effect, λ_t is a time fixed effect, and ϵ_{mt} is an iid normally distributed error term. Assuming the post-merger period begins at time $t+k$ and M is the set of affected markets, then D_{mt} equals one for all $t \geq t+k$ and $m \in M$ and zero otherwise. The coefficient β therefore measures how the outcome, y_{mt}, changes post-merger on affected routes *relative* to unaffected routes.

In addition to the standard OLS assumptions, there are additional conditions that must hold for β to be an unbiased estimate of the average treatment effect of the merger on routes M. For the purposes of this article, we focus on two important assumptions and refer readers to Wooldridge (2002) for a full treatment of DID estimation. First, within each market, it must be that the merger is an exogenous event. This assumption is violated if, for example, two airlines merge because they expect future demand to increase on their overlap routes relative to non-overlap routes.

Second, it is critical that unaffected routes provide a baseline for what would have occurred "but-for" the merger. Specifically, there must no systematic or time-varying differences between the treatment and control groups that are not accounted for. In equation (1), route fixed effects control for time-invariant differences across routes, and time fixed-effects control for time varying factors common across all routes. While this fixed-effect structure helps account for many unobservables, they are not a panacea. If, for example, population or income is growing faster on treatment routes then the estimate of β may be biased upwards. More specifically, DID assumes that treatment and control groups follow a parallel trend

[6]See, for example, Ciliberto and Tamer (2009) and Berry and Jia (2010).

[7]Equation (1) can also be estimated using firm/market-level data. However, due to institutional details described below and ease of comparing across methods, we aggregate the data to the market level.

prior to the treatment, and would have continued on parallel trends but-for the treatment. In a setting with multiple pre-treatment periods, such as in our analysis, this assumption may be tested directly.

Even if the parallel trend assumption holds, further refinements are often made to the control group. In the context of horizontal mergers, the control group may be limited to those with a similar competitive environment as the pre-merger overlap routes. In airline mergers, if the treatment group only includes routes with large hubs at either end, then it may be appropriate to exclude routes between small regional airports. These decisions, however, are typically guided by economic intuition rather than a precise economic model. Finally, estimating a DID regression using OLS places equal weight on all control observations, giving each equal importance when measuring the treatment effect of the merger.

The following two methodologies offer a data-driven approach to selecting a control group that is potentially a better reflection of the but-for world of the treatment group. These techniques, at least in theory, decrease the amount of discretion or economic intuition needed to select an appropriate control group.

3.2 Synthetic control estimation

Synthetic control was first developed to assess policy effects in a setting with relatively few treatment and control units (Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010)). In such an environment, it may be that each control unit serves as a poor proxy for the treated unit(s). However, a weighting of the control units may produce a "synthetic" control that provides a close match to the treatment unit prior to the policy enactment. The treated unit is then compared to the synthetic control to measure the treatment effect.

In merger retrospectives, there are typically multiple treated units and a large number of potential control units. The synthetic control method offers the potential to, a priori, remain agnostic about the relevance of each control observation. By selecting a set of characteristics, such as lagged prices or the number of competitors, a synthetic control can be constructed by solving for a set of weights that minimize the distance between the pre-treatment characteristics of the treated unit and a weighted average of the control units.

In this article, we apply the extension of Abadie, Diamond, and Hainmueller (2010) developed in Cavallo, Galiani, Noy, and Pantano (2013) for multiple treated units.[8] Again,

[8]We give an overview of the methodology here, and refer readers to those articles for a more complete exposition.

define the pre-merger period at time $t = 1, .., t + k - 1$ and the post-merger period to be $t = t + k, ..., T$. Let Y_{mt}^N be the price that would be observed on route m at time t if no merger occurred, and let Y_{mt}^I be the price if the merger did occur. The effect in market m of the merger at time $t \geq t + k$ is then defined as $\alpha_{mt} = Y_{mt}^I$ - Y_{mt}^N. In markets affected by the merger, however, Y_{mt}^N is unobserved for $t \geq t + k$, and therefore must be estimated.

Let C be the number for routes in the control group, which are unaffected by the merger. Also, let Z_m be a $(1 \times z)$ vector of observed predictors for price on route m. Abadie et al. (2010) show that if a set of weights, $w_c^* \geq 0$, exist such that the following conditions hold then an unbiased predictor of α_{mt} can be estimated,

$$\sum_{c=1}^{C} w_c^* Y_{ct} = Y_{mt}, \forall t \in [1, .., t + k - 1] \tag{2}$$

$$\sum_{c=1}^{C} w_c^* Z_c = Z_m \tag{3}$$

$$\sum_{c=1}^{C} w_c^* = 1 \quad . \tag{4}$$

Given these conditions hold, then the following is an estimator for the treatment effect for each post-merger time period in each market:

$$\widehat{\alpha}_{mt} = Y_{mt}^I - \sum_{1=c}^{C} w_c^* Y_{ct}. \tag{5}$$

There is no assurance that a set of weights, w^*, will exist such that equations (2)-(4) hold exactly, nor is a unique solution guaranteed. Finding the set of weights such that each equation approximately holds is the primary estimation challenge.

We implement the following procedure to compute the control weights.[9] Let T_0 denote the number of pre-merger time periods, and $K = (k_1, ..., k_{T_0})$ be a linear combination of pre-merger prices, such that $\bar{Y_m^K} = \sum_{s=1}^{T_0} k_s Y_{ms}$. Let G be a finite set of linear combinations, $K_1, ..., K_G$, and $X_1 = (Z_1'; \bar{Y_1^{K_1}}, ..., \bar{Y_1^{K_G}})$ be a $(n \times 1)$ vector of pre-merger linear combinations and price predictors on a route affected by the merger, where $n = z + G$. Then, define X_0 as a $(n \times C)$ matrix that contains the same pre-merger variables, and each row of X_0 corresponds to a control route. The optimal weights, W^*, are chosen to minimize the distance between the treated unit and a weighted sum of the control units.

[9]We implement the method using the "synth" package in R.

More specifically, the following equation is minimized:

$$\|X_1 - X_0W\| = \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)}. \tag{6}$$

Here, V is a $n \times n$ symmetric, positive definite matrix that weights the relevance of the pre-merger prices and characteristics. The synthetic control method is valid for any symmetric, positive definite V matrix; we solve for the optimal V matrix that minimizes the root mean-squared error defined by equation (6).

We solve for a synthetic control unit for each route where the merging firms both offer non-stop service prior to the merger. We then use equation (5) to calculate the treatment effect on each route. To test for significance, we follow the permutation tests developed in Cavallo, et al. (2013). First, we calculate a synthetic control route and "placebo" effect for each control route. Then, if there are M routes affected by the merger, we randomly select M placebo effects and take the average. We take 1,000 averages and then test where the average treatment effect falls in the distribution of placebo averages.[10]

Thus, the p-value for the test of whether the average merger price effect is positive is,

$$\text{p-value} = \frac{\sum_{i=1}^{1000} I(\bar{\alpha}^{tr} < \bar{\alpha}^{pl})}{1000}, \tag{7}$$

where $\bar{\alpha}^{tr}$ is the treatment effect on the affected routes and $\bar{\alpha}^{pl}$ is an average of M randomly chosen placebo routes.

While the synthetic control method allows the researcher to remain agnostic about the relevance of each control unit, there is still some discretion afforded. First, to the best of our knowledge, there is no definitive metric to determine whether the synthetic control unit is sufficiently close to the treatment unit. Abadie et al. (2010) exclude placebo routes where the mean square error is more than 20 times the treated unit; Cavello et al. (2013) informally gauges goodness of fit of the synthetic control relative to the average value of the dependent variable. Average treatment effect estimates may therefore be sensitive to the "convergence" tolerance, especially when there are a small number of treated units. Second, the variables upon which to base the synthetic control, Z_m and Y_{ct}, need to be specified. As we demonstrate below, both the magnitude and statistical significance of the treatment effect can be sensitive to this choice.

This choice is analogous to choosing dependent variables in a DID regression. One clear

[10] Cavello suggest taking all possible M unit averages. Given the number of control routes, this is computationally infeasible in our samples.

advantage of synthetic control, however, is that the parallel trend assumption will be satisfied at least as well as in DID.

3.3 Nearest neighbor matching

The objective of nearest neighbor matching is to select a finite subset of control units similar to each treated unit in order to measure the effect of treatment. We follow the methodology developed in Abadie and Imbens (2006, 2011). Again, let Y_{mt}^N be the price on route m at time t if no merger had occured, and let Y_{mt}^I be the observed price on a route where the merger did occur. Then, let $\mathcal{J}_N(m)$ be the N "nearest neighbors" of route m. A route's nearest neighbors are determined by first selecting a set of relevant covariates, $x = \{x_1, ..., x_f\}$. Then, the distance between routes m and a are:

$$\|x_m - x_a\|_S = \sqrt{(x_m - x_a)' S^{-1} (x_m - x_a)}. \tag{8}$$

Here, S is a symmetric, positive-definite matrix that weights the relative importance of the covariates, and x_m and x_a are $(1 \times f)$ vectors of observable characteristics for routes m and a, respectively. In this article, we specify S such that we calculate the Mahalanobis[11] distance between routes, which allows us to standardize covariates to be on comparable scales. As is in the synthetic control method, we include lagged values of the outcome variable in the characteristic vector, and in some specifications we also include additional measures of route-level competition.

The estimated treatment effect on route m is $Y_{mt}^I - \widehat{Y}_{mt}^N$, where:

$$\widehat{Y}_{mt}^N = \frac{1}{N} \sum_{j \in \mathcal{J}_N(m)} \{Y_{jt} + \widehat{\mu}_t(x_m) - \widehat{\mu}_t(x_j)\}. \tag{9}$$

In equation (9), $\widehat{\mu}_t$ is a bias correction term developed in Abadie and Imbens (2011) to account for nearest neighbor estimates being inconsistent when matching occurs with more than one continuous variable. The bias correction term, $\widehat{\mu}_t$, is estimated by regressing the outcome variable on the set of matching covariates, and then using the predicted value for each route. The average treatment effect at time t is then:

$$\widehat{\tau}_t = \frac{1}{M} \sum_{m=1}^{M} Y_{mt}^I - \widehat{Y}_{mt}^N, \tag{10}$$

[11] See Mahalanobis (1936) for an exact formula.

where M is the number of routes that are affected by the merger. Standard errors are derived as in Abadie and Imbens (2006, 2011).[12]

Nearest neighbor matching is similar to the synthetic control method in that it allows the data to determine the relevance of each control unit while still leaving important choices to the researcher. As with synthetic control, the set of matching characteristics must be selected. Additionally, the number of neighbors much be chosen. A potential disadvantage to nearest neighbor matching is that a lot of data from the control units are discarded. Where synthetic control may potentially place positive weight on each control unit, nearest neighbor constrains all but N units to have zero weight for each treated unit. As matching is typically done with replacement across treated units, this may result in a majority of data being ignored in estimation. Of course, this may be viewed as an advantage as all but the most relevant control units are considered.

4 Data and Summary Statistics

4.1 Data overview

The data used in this article comes from the US Department of Transportation's Origin and Destination Survey database (DB1B), and is commonly used in the economic analysis of the airline industry.[13] The data is a 10% sample of all domestic airline tickets, and include the quarter of travel, origin and destination airport, airline, number of stopovers, and the price paid for each ticket in the sample.

To conduct the analysis, we restrict the data to only include routes with regular traffic[14] and, similar to previous studies[15] we remove tickets with very low or high prices. We specify routes to be non-directional for a number of reasons. First, recent research demonstrates that airlines almost always set the same price for the same seat, whether or not the round trip is A-B-A or B-A-B (Lewis, 2019). Second, a given airline almost always faces the same set of competitors, regardless of whether the flight is A-B or B-A. Third, the US government often defines non-directional routes when performing its analysis. [16] Thus,

[12] See those respective articles for more detail.

[13] See, for example, Goolsbee and Syverson (2008) and Berry and Jia (2010).

[14] Following Carlton, et al. (2019), we only consider routes with an average of 20 passengers per day. Because the data is a 10% sample, this is a cutoff of approximately 200 passengers per day.

[15] Round trip equivalents below below $50 and above $2,000 are removed from the sample. Ciliberto, et al. (2019) similarly restrict their data.

[16] See, for example, United States Government Accountability Office's analysis of airline competition. https://www.gao.gov/assets/670/664060.pdf.

specifying directional routes avoids creating near-duplicate observations and is consistent with how antitrust practitioners define airline markets.

In accordance with antitrust practice and much of the academic literature, we restrict the data to only include non-stop routes. As in Carlton et al. (2019), we aggregate the data up to the route-quarter level and take the quantity-weighted average price as the primary outcome variable. The analysis is performed at the route-level rather than the airline-route level for two reasons. First, it is more natural to construct synthetic controls and perform matching at the route level, rather than matching on airline-route observations. Second, in response to an exercise of market power, non-merging firms will typically change their prices. Measuring price at the route level accounts for these reactions and gives a more complete accounting of the change in prices paid by consumers.

The data allow us to identify the number of passengers buying tickets at a given price, on each airline on each route. This information is employed to construct variables on which to base the synthetic control units and nearest neighbor matches. Specifically, we create quantity-based HHI's, which capture the level of competition on a route. We also use the total number of legacy carriers[17] as an additional competitive metric upon which to match routes. Previous research, such as Brueckner, Lee, and Singer (2013), demonstrates that legacy carriers create less competitive pressure than low-cost carriers (LCC's), and therefore it may be important to separate out legacy carrier impact when matching routes.

4.2 Selecting Treatment and Control Routes

We focus on the four recent US airline mergers: Delta/Northwest, United/Continental, Southwest/AirTran, and American/USAir. The Southwest/AirTrain merger combined low-cost carriers and the other three mergers were between legacy airlines. For each merger, the analysis includes the eight quarters before and after the merger, and we drop from the sample the quarter in which the merger was consummated. We focus on routes that are most likely to be impacted by the merger. As such, we define a route as an "overlap" if both merging airlines had at least a 10% market share in all eight quarters prior to the merger being consummated. Relative to the total number of routes each airline flies, the merging airlines overlap with non-stop service on relatively few routes.

Selecting the control routes against which to measure the impact of an airline merger is critical to the analysis. Before implementing any of the methodologies, similar to Carlton et

[17] The legacy carriers in the sample are, American Airlines, Continental, Delta, Northwest, United, and USAir

al. (2019), we discard routes with passenger counts that are too dissimilar from the overlap routes. For each merger, we pool the set of overlap routes and calculate the minimum and maximum number of observed passengers in any quarter. We then drop any route such that, in any quarter, the total number of passengers is 10% less than the minimum or 10% greater than the maximum observed on the overlap routes. In theory, synthetic control and nearest neighbor matching do not place weight these dropped routes if they are too dissimilar from the treated routes. Still, we apply these filters prior to implementing all three methods so that results are comparable across methodologies and to previous DID studies.

Finally, we split the control routes into two categories: (i) neither merging firm offered non-stop service in the two years prior to the merger and (ii) exactly one merging firm had at least a 10% share in each of the 8 quarters prior to the merger. Any non-overlap route that does not fall into these two categories is dropped from the analysis. We create two control groups in order to account for horizontal mergers having two potential countervailing effects on price: increased market power and marginal cost efficiencies. Measuring the merger effect relative to routes where neither firm has a pre-merger presence facilitates an estimate of the net impact of the merger.[18] That is, the effect on the overlap routes potentially include both market power and efficiency effects. If routes where one merging carrier operated pre-merger were included in the first control group, then potential cost savings would be included in both the treatment and control, which would lead to a biased estimate of the net merger effect.

Estimating the effect relative to routes where one carrier was present pre-merger allows us to separate out the market power effect from the cost efficiency effect. If cost savings are realized equally across routes and pass-through is, on average, the same across control and treated routes then this isolates the market power effect from the cost saving effect. For example, suppose we find a negative impact on overlap routes relative to control group 1, but a positive impact relative to control group 2. This would suggest that cost efficiencies outweighed the exercise of market power (negative relative to group 1), but that there was still a market power effect (prices increased relative to group 2). In other words, the pass-through of marginal cost efficiencies was lower on overlap routes than on routes with only one merging firm due to greater post-merger market power on overlaps.

This reasoning relies on two important assumptions. First, it assumes that, but-for the change in market power, the merging firm has, on average, the same cost pass-through rate on overlap and control group 2 routes. This may not be the case if, for example, the curvature of demand is systematically different on the two sets of routes. Second, overlap

[18] This assumes that routes where neither firm had a pre-merger presence were unaffected by the merger. This assumption would be violated if, for example, the merger facilitated network-wide price coordination.

routes may experience greater opportunities for cost savings and therefore may experience, on average, greater marginal cost efficiencies. Despite these assumptions, the second control group has the potential to isolate the market power effect from the cost savings effect. And, it is prudent to remove the merging parties from the first control group, so that a clean test of the net effect of the merger can be obtained.

4.3 Summary Statistics and Pricing Trends

For each merger, Table 1 summarizes the treatment and control routes and the competitive environment before and after the merger. The number of overlap routes range from 5 (Delta/Northwest) to 13 (Southwest/AirTran). While this is a small number of routes compared to both control groups, the change in market structure on overlap routes is dramatic. The average change in HHI on overlap routes is more than 3,000 for all four mergers;[19] the 2010 US Horizontal Merger Guidelines state that mergers resulting in a change in HHI greater than 200 are "presumed to be likely to enhance market power." Moreover, the affected routes are "highly concentrated markets" according to the Guidelines, with an average pre-merger HHI of more than 3,000 for each merger.[20] The small number of overlaps and high changes in market concentration are a result of the hubbing system, whereby airlines cluster their flights at particular airports.

Table 1 also highlights that overlap routes are, on average, structurally different from the control routes. Prior to the merger, overlap routes tend to be less concentrated than the typical route. This is because, by definition, overlap routes have at least two airlines with at least 10% market share, and most other routes tend to be dominated by a single airline. Furthermore, overlap routes in legacy (low-cost carrier) mergers tend to have a much higher (lower) overall legacy share compared to control routes. In total, the pre-merger market structure of overlap and control routes are systematically different.

Table 2 summarizes prices and passenger counts before and after each merger. Despite dropping control routes with passenger counts more than 10% outside the range of overlap routes, there still exist substantial differences between overlap and control routes. On average, passenger counts in the first control group are 30% to 40% less than on overlaps. There is not a clear pattern, however, in how price levels differ between treatment and control groups in the pre-merger period. In the Delta/Northwest merger, average pre-merger prices are similar betweeen overlap and the first control group; in United/Continental prices are higher in the

[19]We use the average quarterly passenger share on each route in the year prior to the merger to calculate the change in HHI.

[20]For access to the Guidelines see, `https://www.justice.gov/atr/file/810276/download`.

overlap routes; and in Southwest/AirTran prices are much lower in the overlap group.

Some of the four mergers occurred during unique periods where overall economic trends are first-order determinants of price and passenger counts. Delta received regulatory approval to acquire Northwest in the fourth quarter of 2008, just as the Financial Crisis began. This can be seen in the first panel of Figure 1, where prices for overlap and both control groups begin to decrease significantly just after the merger is consummated. The United/Continental merger was consummated in the third quarter of 2010, while the US economy was recovering from the Financial Crisis. As a result, the pre-merger period includes a substantial decrease in prices, followed by an upward trend that passes through the pre and post-treatment period. The Southwest/AirTran merger also occurred during the economic recovery from the crisis, whereas American/USAir coincided with a more stable economic climate. Given the relatively small number of treated units in each merger, there is some concern that unobserved heterogeneous effects of the crisis may bias the results. To the extent that these unobservables are captured in the pre-merger prices, and uncorrelated with treatment, synthetic control and nearest neighbor matching may be well-suited to address the problem. On the other hand, DID using route and time fixed-effects may not be able to address these heterogeneous effects.

Another potential concern is that some routes are more prone to seasonal fluctuations that others. In Figure 4, passenger counts on the United/Continental overlap routes (and to a lesser extent, Delta/Northwest and Southwest/AirTran) appear to be highly seasonal, whereas the control routes appear less prone to seasonal demand fluctuations. Time fixed effects will not completely account for this difference between groups, but synthetic control and nearest neighbor could potentially place more weight on routes that exhibit similar seasonal patterns. In the following section, we show that these latter two methods tend to produce a better pre-merger fit between the treatment and control group in terms of pre-merger prices. Furthermore, synthetic control and nearest neighbor routes exhibit seasonal demand patterns that are closer to the treatment routes, but are far from a perfect fit. We provide more details in the following subsection.

4.4 Summary of pre-merger matching

In Figure 1 it is not clear, to the eye, whether or not the parallel trend assumption holds in any of the mergers. As such, we follow Autor (2013) and use regression analysis to perform a formal test. Specifically, we estimate equation (1) but include an interaction between the treatment and time dummy variables. If these interactions are statistically significant in the pre-merger period then it demonstrates that overlap routes followed a different time trend prior to the merger, and therefore post-merger effects should be interpreted with caution. We find that in 3 of the 4 mergers, the parallel trend assumption is rejected with 95% confidence. In Delta/Northwest, United/Continental, and Southwest/AirTran, at least 4 of 8 pre-merger interactions are statistically significant with at least 95% confidence when the first control group is used. Nearly identical results hold when the second control group is used, although United/Continental is only statistically significant in 3 of 8 periods with 95% confidence. For American/USAir, only one time period interaction is significant pre-merger with the first control group, and none are significant using the second control group.[21] At best, these results suggest that in all but the American/USAir merger, overlap routes follow different seasonal patterns than do control routes, which is borne out in the passenger counts depicted in Figure 4. At worst, the results demonstrate a violation of the parallell trend assumption, and the DID regressions should be interpreted with caution.

Conversely, in the synthetic control routine, pre-merger prices on overlap routes look nearly identical to the synthetic control routes. Although synthetic control places positive weight on almost every control route, a small handful of routes tend to get a large majority of the weight. For each treated route in each merger, the five routes receiving the most weight account for, on average, 85.3% of the total weight.[22] When competition measures are used in addition to lagged prices, they tend to get minimal weight in the synthetic control; on average across mergers, HHI and the number of legacy carriers receive .006 and .013 weight, respectively. Moreover, HHI and the number of legacy carriers receive no weight in about two thirds of the synthetic control routes. In turn, including these additional matching variables has almost no impact on the amount of weight placed on a given route in constructing the synthetic control. The average absolute difference in the weight placed on a control route is 0.001 across all mergers and synthetic control routes. Figure 2 depicts the mean price for synthetic control routes and actual overlap routes. For each of the mergers, the pre-merger fit is very good; the synthetic control routes are able to match the pre-merger seasonality and price trends in all four mergers. Additionally, the synthetic control routes appear to match the post-merger seasonal patterns, although the gap in price levels appear to grow in all but the Southwest/AirTran merger.

Finally, the parallel trends assumption clearly holds for at least three of the four mergers.

[21] These results are available upon request.

[22] This statistics is for the first control group and matching on only pre-merger prices. This statistic ranges from 83.0% to 91.4% across control groups 1 and 2 and whether or not competition variables are used in matching.

As shown in Figure 3, the Delta/Northwest and United/Continental mergers have virtually identical treated and control pre-merger price trends. Lastly, the difference in levels for American/USAir is minimal, especially between the treated group and the one merged firm control group. The no merged firm group, still, remains relatively close to the treated route throughout the entirety of the pre-merger period and follows a similar, though not identical, seasonal trend.

5 Merger Retrospective Results

We now present the merger retrospective results. The focus is on the price impact of the merger, but we make reference to the total passenger results presented in the appendix. We begin by analyzing each merger separately. Then, we discuss overall implications of horizontal mergers in the US airline industry and compare results across the three retrospective techniques. We include two years of quarterly, route-level data before and after each merger, and drop the quarter in which the merger occurred. This results in 16 observations for each route included in the regression.

The DID estimates include two specifications,

$$y_{mt} = \alpha + \beta V_m P_t + \gamma_m + \lambda_t + \epsilon_{mt} \tag{11}$$

$$y_{mt} = \alpha + \beta_1 V_m P_{1t} + \beta_2 V_m P_{2t} + \gamma_m + \lambda_t + \epsilon_{mt}. \tag{12}$$

For the primary results, y_{mt} is the log of the average quarterly price on each route. In equation (11), $V_m P_t$ is the interaction of an overlap indicator and an indicator for the eight post-merger quarters. In equation (12), we separately estimate treatment effects for the first and second post-merger years.

For the synthetic control method and nearest-neighbor matching, we include two specifications. The first matches only on the eight pre-merger price observations. The second specification includes two additional match variables: the quarterly average HHI, the number of legacy carriers, and the market share of legacy carriers on a route in the year prior to the merger.

5.1 Delta/Northwest results

The Delta/Northwest merger was the first in a wave of airline mergers that began after years of eroding industry profits following the terrorist attacks of 9/11. The Delta/Northwest merger coincided with the onset of the 2008 financial crisis, which impacted many sectors of the economy including the airline industry. Given the small number of overlap routes (5), there is some concern that any estimated effect could be due to unobservable, heterogeneous impacts of the financial crisis. Therefore, synthetic control and nearest-neighbor may be better equipped to produce reliable estimates, as they more closely match pre-merger price trends leading up to the economic downturn.

Table 5 reports the DID estimates, and there is some support for prices decreasing in the first post-merger year, relative to routes where no merging firm was present pre-merger. There is some evidence that overlap route prices increased relative to routes with one merging firm pre-merger. This suggests that the merger resulted in efficiencies but, due to increased market power, pass-through was lower on overlap routes than routes with only one merging firm. The effect on total passengers is inconclusive relative to the first control group, and significantly negative relative to the second. Thus, DID offers some support for merger-specific efficiencies lowering overall prices.

On the other hand, synthetic control yields generally insignificant results. Whether or not competition variables are used in addition to pre-merger prices, there is no statistically significant difference between overlap and the synthetic control routes constructed from either of the two control groups.[23] Nearest neighbor matching also produces insignificant results relative to the first control group, whether or not competition controls are included. On the other hand, relative to routes with one merging firm, prices increase significantly by 7%-14%. These results largely lose significance, however, when extra competition matching variables are included.

In total the three methodologies produce disparate results. Standard DID suggests that there may have been cost efficiencies that led to lower price in the first post-merger year. However, merger-specific cost efficiencies would typically persist into the second post-merger year, which does not seem to be the case. Furthermore, neither synthetic control nor nearest neighbor appear to support these findings. As this merger occurred during a volatile time in the US economy, we place more weight on synthetic control and nearest neighbor matching, which may better account for unobservables across routes. However, as the pre-merger period largely occurred prior to the financial crisis, it's not clear how well pre-merger price trends predict the post-merger, unobserved impact of the crisis. Given the disparity across methodologies, we believe the impact of merger on prices is unclear from the analysis.

[23] The one exception is for the second post-merger quarter, where prices are 4.9% lower with 90% confidence when only pre-merger prices are used for matching.

5.2 United/Continental results

The United/Continental merger was the second legacy airline combination during the merger wave. It was completed in late 2010, two years after Delta/Northwest received antitrust approval. The two airlines had seven overlap routes prior to the merger, all of which contained endpoints at Newark International, Chicago O'Hare, and Denver International. Unlike the Delta/Northwest merger, all three methodologies tell the same general story; the United/Continental merger resulted in higher prices and no cost efficiencies.

The DID estimates presented in Table 8 show a 7% price increase in the first post-merger year relative to both control groups. The nearly identical results across control groups is consistent with news reports[24] of the two airlines having difficulty combining operations, and therefore being unable to realize marginal cost savings. The sign and magnitude of the estimated effects in the synthetic control and nearest neighbor methods (Tables 9 and 10, respectively) are consistent with the DID findings. The estimated effects across control groups are in the 1%-14% range, depending upon the post-merger quarter. Statistical significance, however, is not always robust to whether or not we match on competition variables in additional to pre-merger prices. Furthermore, while passengers are estimated to fall on overlap routes in each of the three methods, the decrease is not statistically different from zero in any specification other than on version of nearest neighbor (matching only on pre-merger prices using control group 1).

In total, all methods tell a story of increased prices and no marginal cost efficiencies from the merger. However, statistical significance is sensitive to including competition matching variables in addition to pre-merger prices.

5.3 Southwest/AirTran results

The Southwest/AirTran merger is the only combination of two low-cost carriers in the analysis. Prior to the merger, Southwest and AirTran overlapped on 13 non-stop routes, which is the most overlaps of the four mergers. 12 of the 13 routes include an endpoint at either Baltimore-Washington or Orlando International airports. Previous research has demonstrated that low-cost carriers have a unique competitive constraint on prices (Brueckner, Lee, and Singer (2013)). Furthermore, both airlines are thought to have a lower cost structure, which may limit the scope for merger-specific marginal cost efficiencies. In total, these factors may increase the potential for anti-competitive harm on overlap routes. The results, however, are

[24]See, for example, Carey and Nicas (2015).

decidedly mixed.

In Table 11, the DID results strongly indicate a post-merger exercise of market power on overlap roots. Prices increased on overlap routes by about 9%, and this holds even when the effect is separately estimated for the two post-merger years. The estimates are one to three percentage points lower when the control group includes only routes where one merging firm was present pre-merger. This suggests that there were not significant marginal cost efficiencies and that prices increased relative to both control group due to the exercise of market power. Across the 4 specifications, price effects are estimated precisely. In Table 23, significant passenger decreases are also estimated across the four specifications. The standard DID methodology, therefore, paints a clear picture of post-merger exercise of market power.

The other methodologies, however, offer a different story of how the merger affected overlap markets. The synthetic control method estimates no significant impact of the merger when constructing the control from routes where neither merging airline was present pre-merger. Using routes where one merging firm had a pre-merger presence, the merger is estimated to significantly increase price by 2%-3%, depending on the quarter. The significance, however, largely disappears when competition variables are used in addition to pre-merger prices to construct the synthetic control routes. In total, the synthetic control yields supportive but inconclusive evidence of the merger increasing prices.

In stark contrast, nearest-neighbor matching yields large and significant negative effects when matching only on pre-merger prices and using control routes where Southwest and AirTran had no pre-merger presence. In the first post-merger year, the price reductions are between 9% and 19% and significant with 99% confidence. However, these results flip by adding two additional competition matching variables (pre-merger HHI and the number of legacy carriers); positive and significant price effects between 9%-29% in the first year are estimated. While adding these additional matching criteria to the synthetic control had a minimal impact, they swing the results greatly in nearest neighbor matching. Using the second control group, the results are more consistent; post-merger prices increased. Still, statistical significance is dependent upon adding competition matching variables.

In summary, the estimated impact of the Southwest/AirTran merger is highly dependent upon both the methodology and the variables used for matching treatment to control routes. This highlights that reasonable, yet different choices made by the researcher can yield contrasting policy implications.

5.4 American/USAir results

American Airlines and USAir is the most recent legacy airline merger in the United States, and was completed in November of 2013. The merger left the United States with only three legacy airlines. The US Department of Justice initially sued to block the merger, but eventually reached a settlement. As part of the settlement, the merged company had to sell slots and/or gates at 6 airports. We exclude routes that experienced entry due to the divestiture from the analysis.The merger resulted in 10 overlap routes that did not experience post-merger entry due to divestiture. The overlaps predominantly had endpoints at the airlines' major hubs in Charlotte, Philadelphia, and Dallas.

Similar to the United/Continental merger, all retrospective methodologies point in the same direction. However, in this case, the merger appears to have resulted in a price decrease. Table 14 presents the DID results and, relative to both control groups, overlap prices fell by 7%-10%. Between the two control groups, there is no statistical difference in the magnitude of the price decreases. This suggests that marginal cost efficiencies were realized exclusively on overlap routes. Alternatively, overlap routes received an unobserved negative demand shock subsequent to the merger. However, this is unlikely, as total passengers are also estimated to increase post-merger on overlap routes (Table 26).

Synthetic control results are reported in Table 15 and tell a story similar to DID, although magnitudes are smaller and statistical significance is not as robust. When control routes do not include either merging firm, post-merger price effects are in the -2% to -6% range for all but one quarter, and they are only statistically different from zero in the first and eighth post-merger quarters. Adding additional competition matching variables flips the sign for half of the post-merger quarters, and the estimates are only significant in the first post-merger period. Relative to the second control group, overlap prices are found to decrease significantly using both sets of matching variables.

The nearest neighbor estimates are reported in Table 16. Using the control group with no merging firms present, the effect of the merger is estimated to be negative in every quarter and significant with at least 90% confidence in 7 of 8 post-merger quarters. However, including the two competition measures in the matching criteria flips the sign of the estimates in 6 of 8 quarters and all statistical significance is lost. Estimates are more robust when matching on the control group with one merging party present pre-merger. The merger is estimated to lower prices in every quarter, whether or not additional competition matching variables are used. The effects are also statistically different from zero in most post-merger quarters.

The passenger results reported in Appendix Tables 26 through 28 largely agree on a statically significant increase in post-merger passengers on overlap routes. Taken as a whole, the results are consistent with the merger leading to lower prices on overlap routes. However, the findings are not entirely robust to adding competition matching criteria in the synthetic control or nearest neighbor methods. As this is the only merger that passed the parallel trend test, the standard DID results may be more reliable than in the other three mergers.

6 Comparing Methodologies and Discussion of Results

The previous subsections highlight two important findings. First, the impact of airline mergers is heterogeneous and may generate either higher or lower prices. Furthermore, estimated effects may be sensitive to both the employed methodology and choices within a given methodology. In this section, we discuss some sources of these sensitivities and offer guidance as to when a particular methodology is best employed.

There are three important sources of estimation sensitivity that practitioners interested in using DID-style estimators to evaluate mergers, specifically, and policy changes, more generally, should be aware of. The first is that estimated treatment effects are subject to researcher degrees of freedom in each of three methods. Different specifications and models, all reasonable, may yield divergent results for the same event.

Second, nearest neighbor matching and synthetic control have very little overlap in the routes designated most similar to the treated route. Across mergers, only 6%-16% of the five "nearest neighbors" were also the top five most heavily weighted routes by the synthetic control method. Both methodologies are intended to remove research intuition from the process of selecting control groups, and allow the data to speak for itself. Yet, these two methods result in almost completely dissimilar control groups, both in weighting and composition.

Finally, recent literature on cross-validation methods in synthetic control, caution that given a large number of control units, idiosyncratic price variations in the donor pool may comprise large portions of the synthetic control group. With a large enough sample size, we can expect that random realizations of the error term can potentially lead to a well-fit linear combination of control units. Horizontal mergers across geographic markets are a unique "natural experiment" with which to compare methods, as they affect price and quantity which are simultaneously determined by a supply and demand system. Often, when analyzing policy interventions, important unobservables are unknown and their relationship with the outcome variable is difficult to determine. One of the benefits to synthetic control and nearest neighbor matching algorithms is that they, in theory, account for important unobservable determinants of the outcome variable. A good match should not only be a well-fitting linear combination of lagged outcomes, but there also be only small differences in unobservable outcome determinants. In theory, a good control group should not only match pre-merger price trends, but also follow similar quantity trends.

Figures 4, 5, and 6 show the passenger trends for the OLS, synthetic control, and nearest neighbor control groups, respectively. There are clear discrepancies in each method between the passenger trend of the control and treated units, with respect to both the level and seasonal components of the trends. It is not immediately clear, to the eye, which method produces the best-fitting control group on passengers.

We test how well the pre-merger passenger trends match that of the treated group. To do so, we use actual passengers counts on the control routes, and weight them according to the route weights used by DID, synthetic control, and nearest neighbor when price is the outcome variable of interest. We then calculate the root mean-squred error (RMSE) between the two groups over the 4 mergers, and report the results in Table 29. Synthetic control produces the best fitting pre-treatment fit control group with one control group present. However, the nearest neighbor control unit fits best for the no-merged firm control group. The OLS control unit, an unweighted mean of the donor pool, performs better than synthetic control for the control group with no merged firms.

It is possible, however, to have a well behaved control group with a comparatively higher RMSE when that control group follows a similar passenger trend and seasonal patterns as to the treated routes. To account for this, we estimate the correlation between the pre-treatment passengers of the treatment and control groups. Table 30 shows the correlation between a control group and the treated group for each merger. The average correlation over control groups and mergers within a method provides an estimate of, generally, how well control group passengers track the treated group. The nearest neighbor control group has an average correlation of 0.72, which is the highest of any method. Synthetic control has the highest correlation for the no merged firms control, at 0.72, but performs the worst among the methods for other control group, with a correlation of 0.54. The average correlation for synthetic control is 0.63. Yet, it performs particularly poorly in the DL/NW merger: the one-firm control group has a correlation of -0.05. Finally, OLS has an average correlation of 0.61. Surprisingly, on average across donor routes, the OLS control group performs almost as well as synthetic control, and substantially better in some instances. Taken together, these results show the importance of carefully constructing the control group in all DID-style methods and avoiding over-reliance on "out of the box" methods. The use of finely tuned cross-validation methods, while time consuming and econometrically rigorous for the user, could provide assurance that synthetic control groups are not composed of random realizations of the dependent variable.

7 Conclusion

We use three of the leading difference-in-differences-style methods to estimate the price-effects of four recent US airline mergers. The results differ across mergers and methods, and are subject to researcher decisions. The American/USAir merger appears to generate lower prices and more passengers on routes in which the airlines overlapped before the merger. Meanwhile, the United/Continental merger led to higher prices. There are meaningful differences in estimated effects between methods for the Southwest/Airtrain and Delta/Northwest mergers. We estimate that the Southwest/AirTran merger significantly increased prices (7%-10%) using DID regression, but significantly decreased prices (-9% to -19%) using nearest neighbor matching.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” *Journal of the American Statistical Association*, 105(490), 493–505.

Abadie, A., and J. Gardeazabal (2003): “The Economic Cost of Conflict: A Case Study of the Basque Country,” *American Economic Review*, 93(1), 112–132.

Abadie, A., and G. W. Imbens (2006): “Large Sample Properties of Matching Estimators for Average Treatment Effects,” *Econometrica*, 74, 235–267.

——— (2011): “Bias-Corrected Matching Estimators for Average Treatment Effects,” *Journal of Business and Economic Statistics*, 29, 1–11.

Ashenfelter, O. C., D. Hosken, and M. C. Weinberg (2014): “Did Robert Bork Understate the Competitive Impact of Mergers? Evidence from Consumated Mergers,” *Journal of Law and Economics*, 57(S3), S67–S100.

——— (2015): “Efficiencies Brewed: Pricing and Consolidation in the US Beer Industry,” *RAND Journal of Economics*, 46(2), 328–361.

Autor, D. H. (2013): “Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing,” *Journal of Labor Economics*, (1), 1–42.

Berry, S., and P. Jia (2010): “Tracing the Woes: An Empirical Analysis of the Airline Industry,” *American Economic Journal: Microeconomics*, 2(3), 1–43.

Borenstein, S. (1990): “Airline Mergers, Airport Dominance, and Market Power,” *American Economic Review*, 80(2), 400–405.

Brueckner, J. K., D. Lee, and E. S. Singer (2013): “Airline Competition and Domestic US Airfares: A Comprehensive Reappraisal,” *Economics of Transportation*, 2, 1–17.

Carey, S., and J. Nicas (2015): “United Continental Is Still Shaky Five Years After Merger,” *The Wall Street Journal*, July, https://www.wsj.com/articles/united–flights–grounded–due–to–computer–issue–1436361911.

Carlton, D., M. Israel, I. MacSwain, and E. Orlov (2019): “Are Legacy Airline Mergers Pro- or Anti-Competitive: Evidence From Recent U.S. Airline Mergers,” *International Journal of Industrial Organization*, 62, 58–95.

Cavallo, E., S. Galiani, I. Noy, and J. Pantano (2013): "Catastrophic Natural Disasters and Economic Growth," *Review of Economics and Statistics*, 95(5), 1549–1561.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828.

Ciliberto, F., E. Watkins, and J. W. Williams (2019): "Collusive Pricing Patterns in the US Airline Industry," *International Journal of Industrial Organization*, 62, 136–157.

Dafny, L., M. Duggan, and S. Ramanarayanan (2012): "Does Matching Overcome LaLonde's critique of nonexperimental estimators?," *American Economic Review*, 102(2), 1161–1185.

Dehejia, R. H., and S. Wahba (2002): "Propensity Score-Matching Methods for Nonexperimental Causal Studies," *Review of Economics and Statistics*, 84(1), 151–161.

Deryugina, T., A. MacKay, and J. Reif (2019): "The Long-Run Dynamics of Electricity Demand: Evidence from Municipal Aggregation," *American Economic Journal: Applied Microeconomics*, Forthcoming.

Garmon, C. (2017): "The Accuracy of Hospital Merger Screening Methods," *The RAND Journal of Economics*, 48(4), 1068–1102.

Goolsbee, A., and C. Syverson (2008): "How Do Incumbent Airlines Respond to the Threat of Entry? Evidence From the Major Airlines," *Quarterly Journal of Economics*, 123(4), 1611–1633.

Houde, J.-F. (2012): "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, 102(5), 2147–2182.

Huschelrath, K., and K. Muller (2013): "The Competitive Effects of Firm Exit: Evidence from the U.S. Airline Industry," *Economics of Transportation*, 2, 72–85.

Imbens, G. W., and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 476–475.

Kim, E. H., and V. Singal (1993): "Mergers and Market Power: Evidence from the Airline Industry," *American Economic Review*, 83(3), 549–569.

LaLonde, R. J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4), 604–620.

LEWIS, M. S. (2019): "Avoiding Competition-Enhancing Price Discrimination: Evidence from the U.S. Airline Industry," Discussion paper.

MAHALANOBIS, P. C. (1936): "On the Generalised Distance in Statistics," *Procedings of the National Institute of Sciences of India*, 2(1), 49–55.

MCCABE, M. J. (2002): "Journal Pricing and Mergers: A Portfolio Approach," *American Economic Review*, 92(1), 259–269.

MEHTA, A., AND N. H. MILLER (2012): "Choosing the Appropriate Control Group in Merger Evaluation," *More Pros and Cons of Merger Control*, pp. 189–210.

MORRISON, S. A. (1996): "Airline Mergers: A Longer View," *Journal of Transport Economics and Policy*, 30(3), 237–250.

PETERS, C. (2006): "Evaluating the Performance of Merger Simulation: Evidence From the U.S. Airline Industry," *The Journal of Law and Economics*, 49(2), 627–649.

SMITH, J. A., AND P. E. TODD (2005): "Does Matching Overcome LaLonde's critique of nonexperimental estimators?," *Journal of Econometrics*, 125, 305–353.

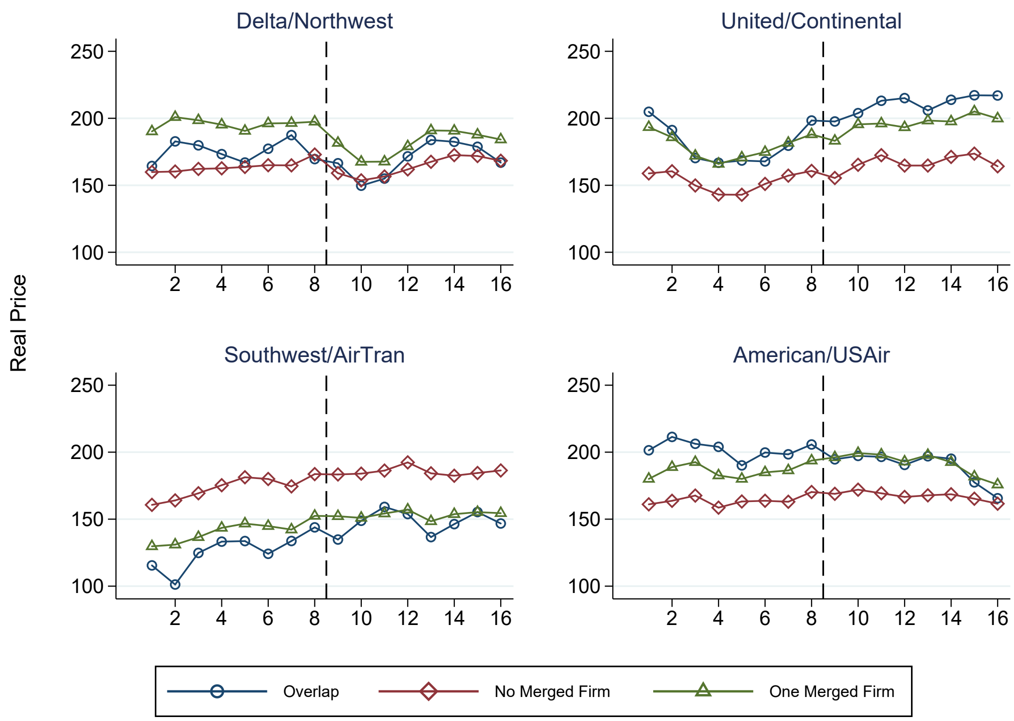
WEINBERG, M. C., AND D. HOSKEN (2013): "Evidence on the Accuracy of Merger Simulations," *Review of Economics and Statistics*, 95(5), 1584–1600.

WERDEN, G. J., A. S. JOSKOW, AND R. L. JOHNSON (1991): "The Effects of Mergers on Price and Output," *Managerial and Decision Economics*, 12, 341–352.

WOOLDRIDGE, J. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

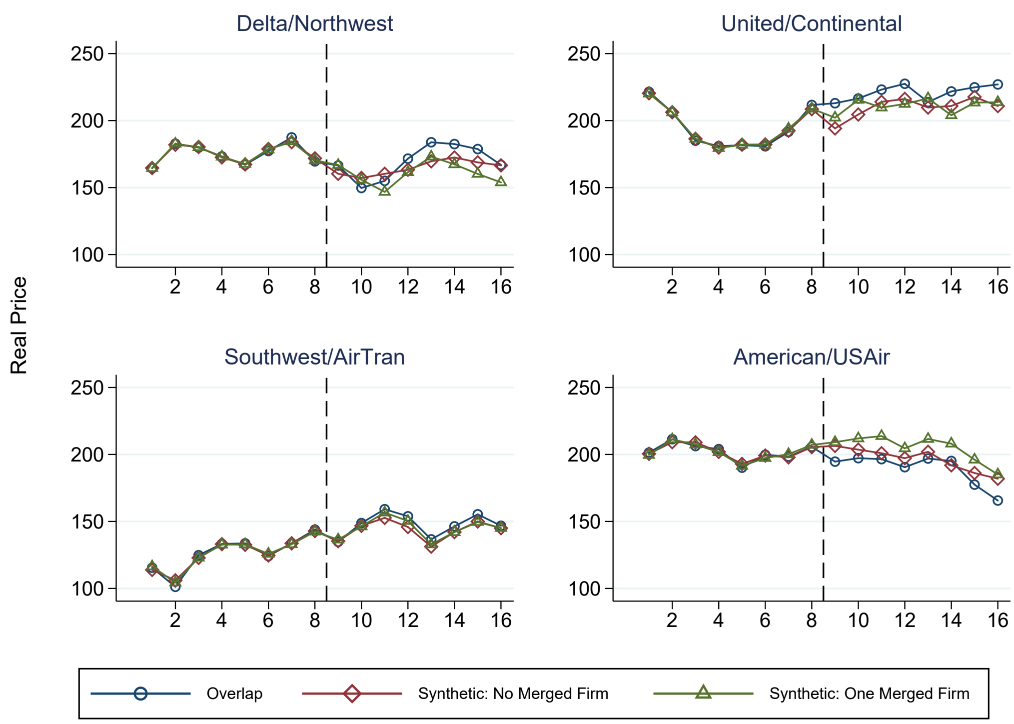
8 Figures

Figure 1: Prices Before and After Mergers



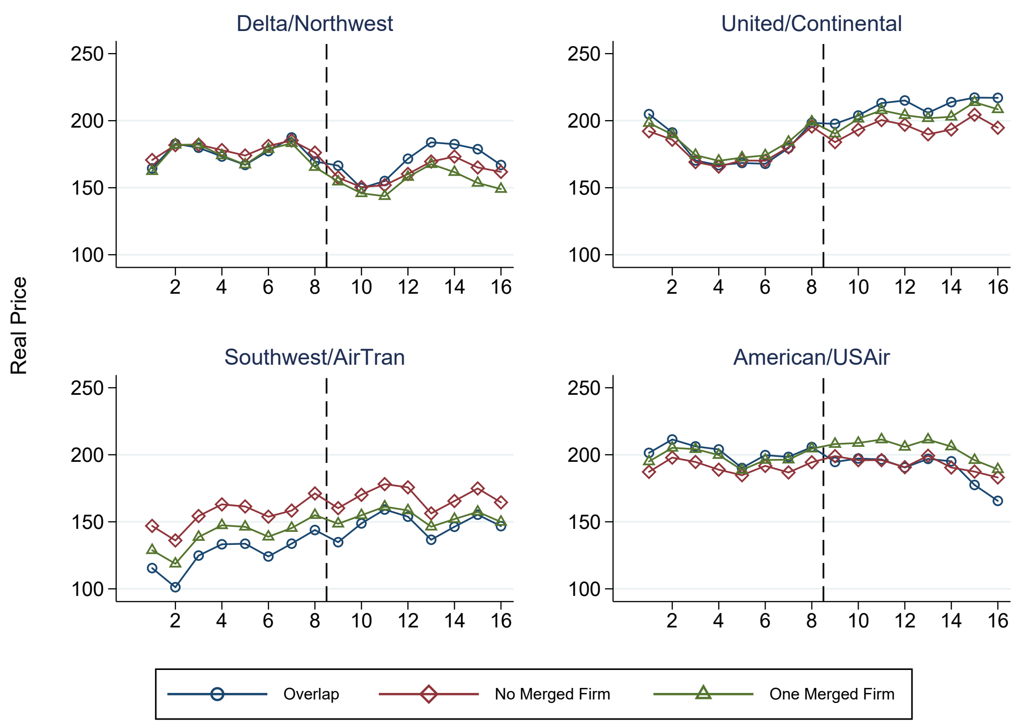
Notes:

Figure 2: Synthetic Control Price Trends



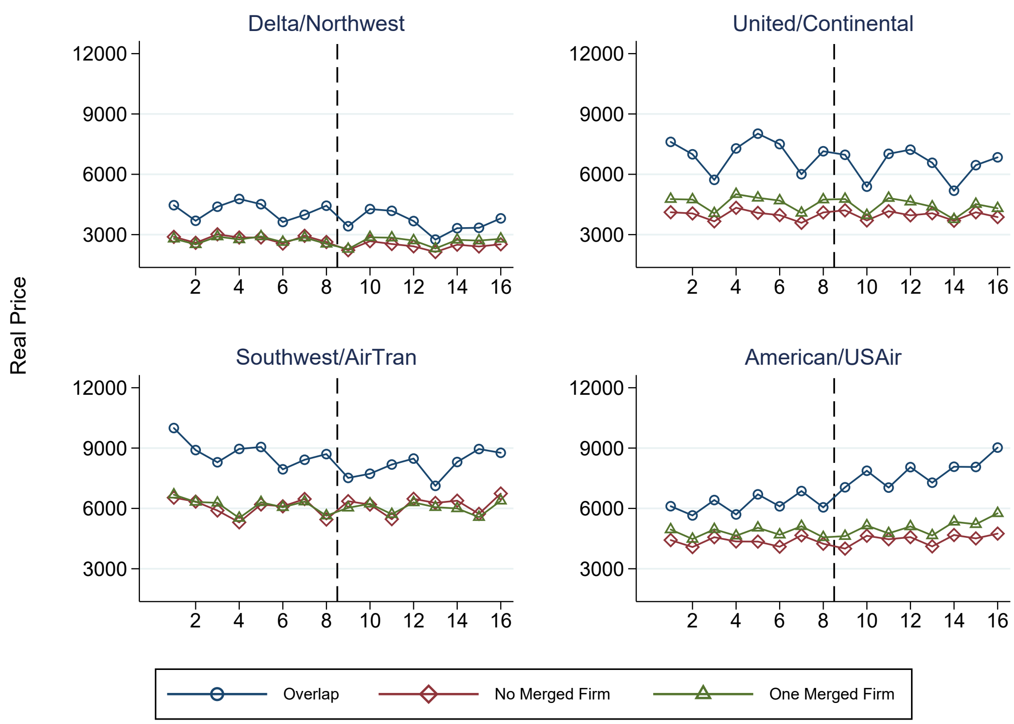
Notes:

Figure 3: Nearest Neighbor Price Trends



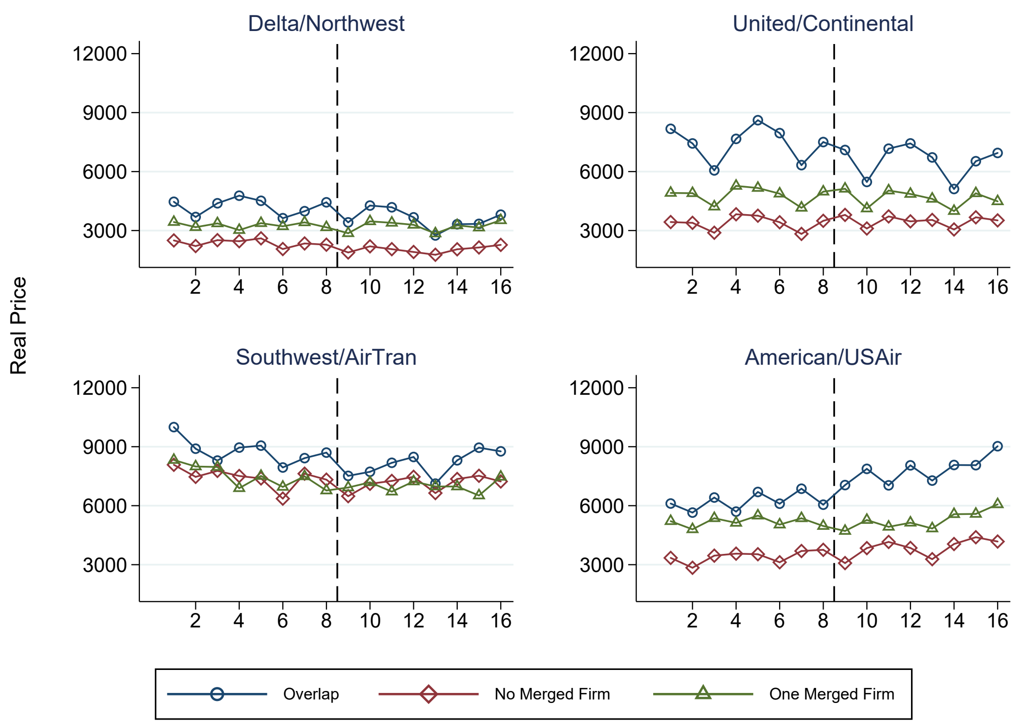
Notes:

Figure 4: Passengers Before and After Mergers



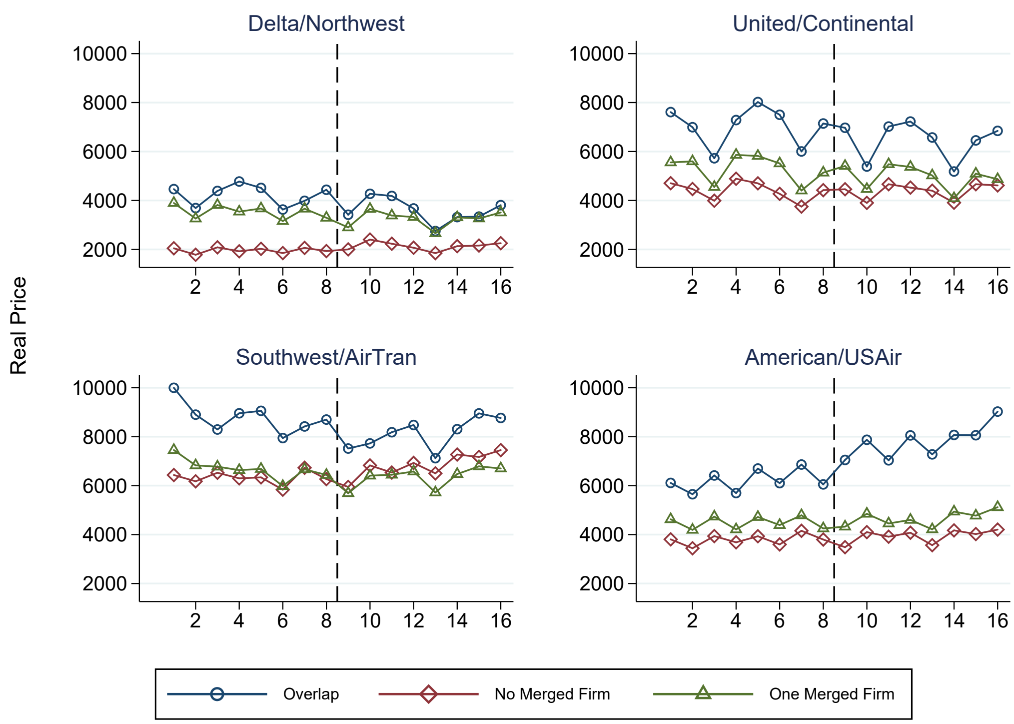
Notes:

Figure 5: Passengers on Synthetic Controls



Notes:

Figure 6: Passengers on Nearest Neighbor Routes



Notes:

9 Tables

Table 1: Route Summary Statistics

	DL/NW	UA/CO	WN/FL	AA/US
Overlap Routes				
Total Routes	5	7	13	10
Legacy Carriers	2.00	2.43	0.08	2.40
Pre-Merger HHI	4323	5052	5590	4865
Δ HHI	3371	3021	3775	3693
Control - No Merging Firms				
Total Routes	678	609	207	387
Legacy Carriers	0.70	0.60	1.39	0.39
Pre-Merger HHI	8572	8130	7073	8277
Control - One Merging Firm				
Total Routes	176	153	250	161
Legacy Carriers	1.31	1.51	0.57	1.42
Pre-Merger HHI	7117	6292	7296	6458

Notes:

Table 2: Price and Passenger Summary Statistics

	DL/NW	UA/CO	WN/FL	AA/US
Overlap Routes				
Pre-Merger Price	175	180	126	202
Post-Merger Price	169	210	147	189
Pre-Merger Passengers	4217	6995	8765	6186
Post-Merger Passengers	3565	6414	8110	7781
Control - No Merging Firms				
Pre-Merger Price	164	153	173	164
Post-Merger Price	164	166	185	167
Pre-Merger Passengers	2789	3982	6021	4343
Post-Merger Passengers	2422	3967	6190	4458
Control - One Merging Firm				
Pre-Merger Price	196	179	141	186
Post-Merger Price	181	196	153	192
Pre-Merger Passengers	2735	4604	6132	4799
Post-Merger Passengers	2648	4381	6029	5062

Notes:

Table 3: Synthetic Control Routes Summary

	DL/NW	UA/CO	WN/FL	AA/US
Overlap Routes				
Total Routes	5	6	13	10
Pre-Merger Price	175	194	126	202
Post-Merger Price	169	221	147	189
Pre-Merger Passengers	4217	7420	8765	6186
Post-Merger Passengers	3565	6510	8110	7781
Legacy Carriers	2.00	2.33	0.08	2.40
Pre-Merger HHI	4323	5307	5590	4865
Control - No Merging Firms				
Pre-Merger Price	175	194	126	202
Post-Merger Price	165	210	143	196
Pre-Merger Passengers	2365	3364	7426	3403
Post-Merger Passengers	2028	3475	7127	3829
Legacy Carriers	0.91	1.08	0.65	0.81
Pre-Merger HHI	8442	7860	8570	7488
Control - One Merging Firm				
Pre-Merger Price	175	194	126	202
Post-Merger Price	160	211	145	205
Pre-Merger Passengers	3268	4794	7470	5166
Post-Merger Passengers	3217	4625	6995	5246
Legacy Carriers	1.17	1.66	0.73	1.49
Pre-Merger HHI	7282	5988	6989	6602

Notes:

Table 4: Nearest Neighbor Routes Summary

	DL/NW	UA/CO	WN/FL	AA/US
Overlap Routes				
Total Routes	5	7	13	10
Pre-Merger Price	175	180	126	202
Post-Merger Price	169	210	147	189
Pre-Merger Passengers	4217	6995	8765	6186
Post-Merger Passengers	3565	6414	8110	7781
Legacy Carriers	2.00	2.43	0.08	2.40
Pre-Merger HHI	4323	5052	5590	4865
Control - No Merging Firms				
Pre-Merger Price	179	178	155	191
Post-Merger Price	161	195	168	193
Pre-Merger Passengers	1961	4384	6321	3788
Post-Merger Passengers	2131	4382	6814	3935
Legacy Carriers	1.04	0.81	1.29	0.41
Pre-Merger HHI	8312	7699	7455	8698
Control - One Merging Firm				
Pre-Merger Price	174	182	139	199
Post-Merger Price	154	204	154	204
Pre-Merger Passengers	3529	5279	6672	4483
Post-Merger Passengers	3236	4953	6339	4650
Legacy Carriers	1.13	1.54	0.65	1.32
Pre-Merger HHI	6125	6825	6992	6901

Notes:

9.1 Delta-Northwest Tables

Table 5: DL-NW Merger: Price Regression

	(1)	(2)	(3)	(4)
Merger X Post	-0.035		0.043	
	(0.03)		(0.04)	
Merger X Post1		-0.049***		0.031*
		(0.01)		(0.02)
Merger X Post2		-0.021		0.054
		(0.06)		(0.06)
Constant	5.099***	5.099***	5.234***	5.234***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	10928	10928	2896	2896
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No DL/NW	No DL/NW	DL/NW: No Overlap	DL/NW: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the quantity-weighted, route-level average price. The data include observations for two years before and after the merger. The "No DL/NW" control group includes only routes with no DL or NW direct flights. The "DL/NW: No Overlap" control group includes only routes with either DL or NW direct flights, but not both. Standard errors are clustered at the route level.

Table 6: DL-NW Merger: Price Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.038	0.462	0.067	0.216	-0.002	0.380	-0.004	0.387
T+2	-0.049	0.096	-0.028	0.267	-0.040	0.258	-0.041	0.239
T+3	-0.033	0.173	-0.085	0.061	0.055	0.308	0.054	0.328
T+4	0.050	0.266	-0.001	0.411	0.061	0.273	0.060	0.277
T+5	0.081	0.236	0.052	0.386	0.060	0.268	0.059	0.284
T+6	0.056	0.379	0.001	0.311	0.087	0.186	0.086	0.203
T+7	0.057	0.413	-0.034	0.174	0.110	0.143	0.109	0.144
T+8	0.002	0.311	-0.081	0.054	0.081	0.175	0.079	0.186
Treated Routes	5		4		5		5	

Notes:

Table 7: DL-NW Merger: Price Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.048	0.219	0.043	0.300	0.062	0.304	0.072	0.195
T+2	-0.015	0.602	-0.027	0.461	0.018	0.716	0.007	0.898
T+3	-0.020	0.770	-0.005	0.927	0.069	0.004	0.081	0.124
T+4	0.037	0.507	0.049	0.503	0.070	0.000	0.095	0.019
T+5	0.019	0.747	0.028	0.620	0.081	0.000	0.058	0.149
T+6	-0.016	0.859	0.033	0.667	0.114	0.012	0.051	0.349
T+7	0.007	0.956	0.026	0.811	0.144	0.021	0.114	0.097
T+8	-0.009	0.924	-0.065	0.389	0.099	0.004	0.126	0.049
Treated Routes	5		5		5		5	

Notes:

9.2 United-Continental Tables

Table 8: UA-CO Merger: Price Regression

	(1)	(2)	(3)	(4)
Merger X Post	0.069*		0.062	
	(0.04)		(0.04)	
Merger X Post1		0.066**		0.068**
		(0.03)		(0.03)
Merger X Post2		0.072		0.055
		(0.05)		(0.06)
Constant	5.074***	5.074***	5.233***	5.233***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	9856	9856	2560	2560
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No UA/CO	No UA/CO	UA/CO: No Overlap	UA/CO: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the quantity-weighted, route-level average price. The data include observations for two years before and after the merger. The "No UA/CO" control group includes only routes with no UA or CO direct flights. The "UA/CO: No Overlap" control group includes only routes with either UA or CO direct flights, but not both. Standard errors are clustered at the route level.

Table 9: UA-CO Merger: Price Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.092	0.007	0.088	0.012	0.048	0.119	0.050	0.072
T+2	0.057	0.110	0.064	0.077	0.016	0.323	0.017	0.309
T+3	0.042	0.134	0.057	0.087	0.073	0.029	0.066	0.056
T+4	0.051	0.138	0.079	0.066	0.079	0.053	0.073	0.119
T+5	0.019	0.367	0.044	0.234	0.019	0.461	0.012	0.473
T+6	0.050	0.249	0.096	0.084	0.112	0.038	0.105	0.073
T+7	0.032	0.243	0.069	0.111	0.075	0.113	0.068	0.181
T+8	0.074	0.115	0.118	0.043	0.084	0.109	0.081	0.138
Treated Routes	6		5		7		7	

Notes:

Table 10: UA-CO Merger: Price Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.119	0.000	0.104	0.005	0.078	0.010	0.098	0.014
T+2	0.116	0.026	0.083	0.152	0.057	0.148	0.070	0.166
T+3	0.114	0.042	0.097	0.081	0.046	0.197	0.037	0.276
T+4	0.139	0.006	0.107	0.033	0.026	0.276	0.009	0.777
T+5	0.110	0.019	0.069	0.124	-0.006	0.833	-0.038	0.213
T+6	0.140	0.068	0.140	0.039	0.007	0.857	0.007	0.890
T+7	0.086	0.159	0.061	0.252	-0.041	0.401	-0.120	0.025
T+8	0.129	0.013	0.109	0.045	0.002	0.966	-0.066	0.310
Treated Routes	7		7		7		7	

Notes:

9.3 Southwest-Airtran Tables

Table 11: WN-FL Merger: Price Regression

	(1)	(2)	(3)	(4)
Merger X Post	0.093***		0.075***	
	(0.02)		(0.02)	
Merger X Post1		0.097***		0.082***
		(0.02)		(0.01)
Merger X Post2		0.090***		0.067**
		(0.03)		(0.03)
Constant	5.170***	5.170***	4.984***	4.984***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	3520	3520	4208	4208
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No WN/FL	No WN/FL	WN/FL: No Overlap	WN/FL: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the quantity-weighted, route-level average price. The data include observations for two years before and after the merger. The "No WN/FL" control group includes only routes with no WN or FL direct flights. The "WN/FL: No Overlap" control group includes only routes with either WN or FL direct flights, but not both. Standard errors are clustered at the route level.

Table 12: WN-FL Merger: Price Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.004	0.227	0.004	0.326	-0.012	0.138	-0.017	0.127
T+2	0.014	0.378	0.031	0.432	0.017	0.037	0.023	0.057
T+3	0.042	0.444	0.057	0.316	0.016	0.129	0.012	0.240
T+4	0.053	0.299	0.063	0.213	0.023	0.045	0.022	0.134
T+5	0.041	0.404	0.032	0.493	0.028	0.087	0.026	0.182
T+6	0.029	0.482	0.026	0.423	0.030	0.056	0.025	0.137
T+7	0.035	0.494	0.041	0.471	0.038	0.017	0.049	0.014
T+8	0.012	0.321	0.012	0.313	0.012	0.222	0.024	0.184
Treated Routes	13		12		13		10	

Notes:

Table 13: WN-FL Merger: Price Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.190	0.000	0.088	0.056	-0.009	0.687	-0.025	0.409
T+2	-0.143	0.000	0.108	0.030	0.027	0.220	0.129	0.000
T+3	-0.165	0.000	0.298	0.000	0.040	0.096	0.105	0.000
T+4	-0.092	0.011	0.173	0.001	0.053	0.057	0.088	0.007
T+5	-0.027	0.640	0.126	0.011	0.047	0.179	0.238	0.000
T+6	-0.008	0.888	-0.055	0.263	0.065	0.114	0.311	0.000
T+7	-0.072	0.221	-0.008	0.872	0.055	0.130	0.216	0.000
T+8	-0.028	0.655	-0.019	0.706	0.029	0.498	0.208	0.000
Treated Routes	13		13		13		13	

Notes:

9.4 American Airlines-USAir Tables

Table 14: AA-US Merger: Price Regression

	(1)	(2)	(3)	(4)
Merger X Post	-0.089**		-0.097***	
	(0.04)		(0.04)	
Merger X Post1		-0.069**		-0.092***
		(0.03)		(0.03)
Merger X Post2		-0.109**		-0.101**
		(0.05)		(0.05)
Constant	5.115***	5.115***	5.246***	5.246***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	6352	6352	2736	2736
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No AA/US	No AA/US	AA/US: No Overlap	AA/US: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the quantity-weighted, route-level average price. The data include observations for two years before and after the merger. The "No AA/US" control group includes only routes with no AA or US direct flights. The "AA/US: No Overlap" control group includes only routes with either AA or US direct flights, but not both. Standard errors are clustered at the route level.

Table 15: AA-US Merger: Price Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.059	0.016	-0.066	0.017	-0.072	0.019	-0.070	0.025
T+2	-0.032	0.106	-0.021	0.225	-0.072	0.013	-0.071	0.027
T+3	-0.022	0.199	0.012	0.388	-0.084	0.011	-0.083	0.009
T+4	-0.035	0.106	-0.018	0.249	-0.071	0.064	-0.070	0.068
T+5	-0.025	0.199	0.017	0.428	-0.071	0.134	-0.068	0.182
T+6	0.017	0.458	0.064	0.102	-0.064	0.267	-0.061	0.295
T+7	-0.048	0.083	0.013	0.461	-0.100	0.194	-0.096	0.204
T+8	-0.094	0.008	-0.043	0.125	-0.111	0.103	-0.106	0.122
Treated Routes	10		9		10		10	

Notes:

Table 16: AA-US Merger: Price Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Prices		Prices + Controls		Prices		Prices + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.102	0.011	-0.054	0.255	-0.085	0.055	-0.091	0.037
T+2	-0.068	0.068	-0.023	0.530	-0.068	0.054	-0.093	0.017
T+3	-0.066	0.076	0.034	0.490	-0.086	0.019	-0.109	0.014
T+4	-0.077	0.092	0.040	0.415	-0.084	0.030	-0.081	0.114
T+5	-0.104	0.050	0.018	0.801	-0.082	0.089	-0.062	0.436
T+6	-0.060	0.267	0.071	0.189	-0.065	0.172	-0.036	0.688
T+7	-0.141	0.086	0.032	0.710	-0.107	0.192	-0.093	0.487
T+8	-0.168	0.019	0.032	0.675	-0.126	0.079	-0.106	0.406
Treated Routes	10		10		10		10	

Notes:

10 Appendix - Passenger Results

10.1 Delta/Northwest

Table 17: DL-NW Merger: Passengers Regression

	(1)	(2)	(3)	(4)
Merger X Post	-0.027		-0.136**	
	(0.05)		(0.06)	
Merger X Post1		0.041*		-0.060**
		(0.02)		(0.03)
Merger X Post2		-0.094		-0.211*
		(0.11)		(0.11)
Constant	7.866***	7.866***	7.910***	7.910***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	10928	10928	2896	2896
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No DL/NW	No DL/NW	DL/NW: No Overlap	DL/NW: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the total non-stop passengers. The data include observations for two years before and after the merger. The "No DL/NW" control group includes only routes with no DL or NW direct flights. The "DL/NW: No Overlap" control group includes only routes with either DL or NW direct flights, but not both. Standard errors are clustered at the route level.

Table 18: DL-NW Merger: Passengers Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.080	0.141	0.076	0.139	-0.017	0.441	-0.013	0.422
T+2	0.046	0.181	0.043	0.164	-0.048	0.443	-0.040	0.464
T+3	0.017	0.248	0.020	0.220	-0.131	0.216	-0.124	0.246
T+4	0.011	0.256	0.010	0.273	-0.227	0.101	-0.209	0.116
T+5	-0.119	0.316	-0.110	0.320	-0.268	0.090	-0.247	0.095
T+6	-0.135	0.268	-0.127	0.267	-0.315	0.055	-0.293	0.058
T+7	-0.206	0.144	-0.194	0.129	-0.369	0.037	-0.355	0.035
T+8	0.013	0.289	0.019	0.316	-0.211	0.188	-0.198	0.204
Treated Routes	5		5		5		5	

Notes:

Table 19: DL-NW Merger: Passenger Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.162	0.014	0.159	0.088	-0.087	0.381	0.019	0.795
T+2	0.210	0.000	0.147	0.212	-0.026	0.580	-0.073	0.320
T+3	0.166	0.023	0.155	0.001	-0.061	0.389	-0.183	0.182
T+4	0.107	0.068	0.259	0.013	-0.171	0.023	-0.282	0.044
T+5	0.044	0.096	0.114	0.176	-0.243	0.057	-0.267	0.143
T+6	0.036	0.725	-0.002	0.985	-0.229	0.113	-0.234	0.225
T+7	-0.036	0.825	-0.035	0.830	-0.342	0.147	-0.251	0.244
T+8	0.129	0.079	0.134	0.077	-0.139	0.167	-0.073	0.604
Treated Routes	5		5		5		5	

Notes:

10.2 United/Continental

Table 20: UA-CO Merger: Passengers Regression

	(1)	(2)	(3)	(4)
Merger X Post	-0.083		-0.037	
	(0.06)		(0.07)	
Merger X Post1		-0.063		-0.042
		(0.04)		(0.05)
Merger X Post2		-0.103		-0.032
		(0.09)		(0.09)
Constant	8.294***	8.294***	8.428***	8.428***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	9856	9856	2560	2560
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No UA/CO	No UA/CO	UA/CO: No Overlap	UA/CO: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the total non-stop passengers. The data include observations for two years before and after the merger. The "No UA/CO" control group includes only routes with no UA or CO direct flights. The "UA/CO: No Overlap" control group includes only routes with either UA or CO direct flights, but not both. Standard errors are clustered at the route level.

Table 21: UA-CO Merger: Passengers Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.005	0.337	-0.020	0.498	-0.028	0.261	-0.065	0.118
T+2	-0.087	0.230	-0.121	0.125	-0.150	0.044	-0.154	0.044
T+3	0.006	0.240	-0.027	0.431	-0.047	0.274	-0.075	0.187
T+4	0.037	0.181	-0.002	0.373	-0.043	0.302	-0.058	0.268
T+5	0.009	0.365	-0.022	0.495	-0.052	0.284	-0.046	0.340
T+6	-0.116	0.210	-0.187	0.083	-0.127	0.154	-0.140	0.136
T+7	-0.056	0.398	-0.102	0.287	-0.117	0.165	-0.087	0.238
T+8	0.018	0.382	-0.031	0.461	-0.100	0.232	-0.026	0.441
Treated Routes	7		6		6		7	

Notes:

Table 22: UA-CO Merger: Passenger Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.121	0.008	-0.052	0.482	-0.050	0.445	0.051	0.603
T+2	-0.268	0.000	-0.092	0.252	-0.095	0.397	0.100	0.494
T+3	-0.092	0.171	0.040	0.612	-0.099	0.193	0.014	0.896
T+4	-0.091	0.272	0.114	0.213	-0.164	0.107	-0.057	0.654
T+5	-0.155	0.030	0.038	0.707	-0.097	0.393	-0.004	0.979
T+6	-0.174	0.029	-0.058	0.533	-0.105	0.419	-0.003	0.985
T+7	-0.105	0.193	0.025	0.788	-0.144	0.345	-0.067	0.722
T+8	-0.010	0.932	-0.008	0.938	-0.137	0.456	-0.060	0.763
Treated Routes	7		7		7		7	

Notes:

10.3 Soutwest/AirTrain

Table 23: WN-FL Merger: Passengers Regression

	(1)	(2)	(3)	(4)
Merger X Post	-0.105***		-0.061**	
	(0.03)		(0.03)	
Merger X Post1		-0.110***		-0.084***
		(0.02)		(0.02)
Merger X Post2		-0.101**		-0.037
		(0.05)		(0.05)
Constant	8.739***	8.739***	8.730***	8.730***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	3520	3520	4208	4208
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No WN/FL	No WN/FL	WN/FL: No Overlap	WN/FL: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the total non-stop passengers. The data include observations for two years before and after the merger. The "No WN/FL" control group includes only routes with no WN or FL direct flights. The "WN/FL: No Overlap" control group includes only routes with either WN or FL direct flights, but not both. Standard errors are clustered at the route level.

Table 24: WN-FL Merger: Passengers Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.188	0.000	-0.198	0.000	-0.119	0.002	-0.122	0.000
T+2	-0.103	0.025	-0.098	0.033	-0.109	0.000	-0.098	0.001
T+3	-0.030	0.453	-0.020	0.475	-0.045	0.190	-0.039	0.250
T+4	-0.124	0.059	-0.116	0.069	-0.023	0.423	-0.039	0.373
T+5	-0.222	0.007	-0.243	0.001	-0.119	0.009	-0.135	0.001
T+6	-0.059	0.197	-0.096	0.056	0.007	0.318	-0.005	0.438
T+7	-0.010	0.496	-0.058	0.183	0.041	0.104	0.052	0.091
T+8	-0.079	0.139	-0.128	0.013	-0.013	0.399	-0.027	0.499
Treated Routes	13		12		11		12	

Notes:

Table 25: WN-FL Merger: Passenger Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	-0.118	0.020	-0.027	0.740	-0.074	0.146	-0.103	0.081
T+2	-0.143	0.009	0.113	0.195	-0.156	0.004	-0.336	0.000
T+3	-0.091	0.089	0.262	0.002	-0.143	0.002	-0.293	0.000
T+4	-0.111	0.049	0.239	0.009	-0.138	0.006	-0.283	0.000
T+5	-0.101	0.149	0.284	0.015	-0.189	0.000	-0.297	0.000
T+6	0.004	0.955	0.577	0.000	-0.120	0.077	-0.305	0.000
T+7	0.019	0.792	0.696	0.000	-0.086	0.259	-0.194	0.039
T+8	-0.017	0.827	0.761	0.000	-0.106	0.144	-0.175	0.044
Treated Routes	13		13		13		13	

Notes:

10.4 American/USAir

Table 26: AA-US Merger: Passengers Regression

	(1)	(2)	(3)	(4)
Merger X Post	0.203***		0.176***	
	(0.06)		(0.06)	
Merger X Post1		0.175***		0.170***
		(0.05)		(0.05)
Merger X Post2		0.231***		0.183***
		(0.07)		(0.07)
Constant	8.398***	8.398***	8.518***	8.518***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	6352	6352	2736	2736
Route FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Control Group	No AA/US	No AA/US	AA/US: No Overlap	AA/US: No Overlap

Notes: Data is at the route-quarter level. The dependent variable is the log of the total non-stop passengers. The data include observations for two years before and after the merger. The "No AA/US" control group includes only routes with no AA or US direct flights. The "AA/US: No Overlap" control group includes only routes with either AA or US direct flights, but not both. Standard errors are clustered at the route level.

Table 27: AA-US Merger: Passengers Synthetic Control

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.094	0.023	0.092	0.026	0.107	0.011	0.118	0.014
T+2	0.098	0.014	0.105	0.020	0.100	0.010	0.103	0.018
T+3	0.092	0.034	0.101	0.031	0.087	0.060	0.087	0.045
T+4	0.101	0.031	0.121	0.027	0.131	0.015	0.136	0.009
T+5	0.101	0.044	0.112	0.033	0.153	0.023	0.162	0.021
T+6	0.106	0.029	0.119	0.028	0.117	0.104	0.127	0.075
T+7	0.162	0.013	0.197	0.007	0.119	0.147	0.126	0.102
T+8	0.147	0.013	0.158	0.007	0.116	0.123	0.116	0.091
Treated Routes	10		9		10		10	

Notes:

Table 28: AA-US Merger: Passenger Nearest Neighbor

Control Group:	No Merging Airlines				One Merging Airline			
Match On:	Passengers		Passengers + Controls		Passengers		Passengers + Controls	
	Effect	P-Value	Effect	P-Value	Effect	P-Value	Effect	P-Value
T+1	0.117	0.063	0.134	0.171	0.113	0.073	0.196	0.035
T+2	0.100	0.127	0.069	0.516	0.133	0.031	0.190	0.021
T+3	0.071	0.317	-0.046	0.674	0.131	0.046	0.144	0.078
T+4	0.088	0.154	-0.087	0.359	0.143	0.012	0.108	0.107
T+5	0.071	0.326	-0.115	0.243	0.127	0.076	0.101	0.250
T+6	0.087	0.341	-0.183	0.149	0.121	0.157	0.101	0.258
T+7	0.141	0.259	-0.090	0.557	0.110	0.267	0.081	0.443
T+8	0.140	0.181	-0.153	0.219	0.061	0.403	0.132	0.187
Treated Routes	10		10		10		10	

Notes:

10.5 Passenger Comparisons

Table 29: RMSE of Passenger Trends using Price Control Groups

Method:	OLS		Nearest Neighbor		Synthetic Control	
Control Group:	No Merging Airlines	One Merging Airline	No Merging Airline	One Merging Airline	No Merging Airlines	One Merging Airline
UACO	2553	2103	2117	1518	3129	1978
AAUS	3371	2747	3889	3170	3985	2557
DLNW	1217	997	1497	435	1608	481
WNFL	2061	2205	1358	1801	1051	1311
Average:	2300.5	2013	2215.25	1731	2443.25	1581.75

Notes:

Table 30: Correlation of Passenger Trends using Price Control Groups

Method:	OLS		Nearest Neighbor		Synthetic Control	
Control Group:	No Merging Airlines	One Merging Airline	No Merging Airlines	One Merging Airline	No Merging Airlines	One Merging Airline
UACO	0.80	0.89	0.82	0.90	0.86	0.87
AAUS	0.69	0.89	0.89	0.90	0.52	0.84
DLNW	0.58	0.47	0.53	0.59	0.81	-0.05
WNFL	0.23	0.35	0.25	0.85	0.68	0.50
Average:	0.58	0.65	0.62	0.81	0.72	0.54
Method Average:	0.61		0.72		0.63	

Notes: