# U.S. AIRLINE MERGERS – WHO BENEFITS FROM THEM AND WHEN?

Assessing The Impact of Mergers on Profitability


## ABSTRACT

In the wake of a potential wave of airline mergers post-Covid, which airlines should seek mergers and how to maximize bottom-line performance post-merger remain two important questions to be addressed. We aim to answer these two questions by studying the impact of mergers on airlines' profitability using eight recent U.S. airline mergers. Building a stacked regression in a staggered difference-in-difference design, we find that the impact of mergers on profitability in the U.S. airline industry is not instantaneous but long-term instead: airlines first experienced a U-shaped performance curve in the first two years following mergers before profitability starts to show synergy from the third year onwards. Therefore, minimizing the negative impact in the first two years post-merger is key to achieve better financial performance. In addition, low-performing acquirers benefit from profitability synergy while high-performing acquirers do not. Accordingly, we do not advise high-performing airlines to engage in mergers. Our study extends the understanding of two management concepts in the airline industry: the learning-by-doing concept and the law of diminishing synergy as well as provides a practical guidance for policy makers engaging in airline mergers.



Keywords: airline merger, staggered difference-in-difference, profitability, stacked regression

# U.S. AIRLINE MERGERS – WHO BENEFITS FROM THEM AND WHEN?

Assessing The Impact of Mergers on Profitability

## 1. INTRODUCTION

Since the deregulation in 1978, the U.S. airline industry has experienced numerous mergers (Singal 1996a, 1996b; Department of Transportation 2023). Accordingly, the impact of these mergers on carrier performance has drawn researchers' attention. Most research on U.S. airline mergers has focused on merger-induced revenue synergy (Kim and Singal 1993, Peters 2006, Luo 2014, Carlton et al. 2019) while some research has also investigated cost synergy (Lichtenberg and Kim 1989, Ryerson and Kim 2014, Schosser and Wittmer 2015). However, one important part of the equation – profitability synergy – has been largely ignored, especially given airline stakeholders have been increasingly focused on making airlines more profitable ever since the deregulation (Adrangi et al. 1997). Therefore, we aim to study the effect of mergers on profitability synergy in the U.S. airline industry.

In the airline merger literature, synergy is defined as "value enhancements that can derive either from less input factors needed to produce the same output or from a higher output with constant input factors" (Schosser and Wittmer 2015, p.143). Following airline revenue synergy literature, we use acquirer's post-merger profitability to proxy profitability synergy, which is accordingly achieved if an acquirer's post-merger profitability improved over time following the merger event. Our study is important as revenue synergy does not necessarily translate into profitability synergy due to increased operating expenses post-merger, which could have a negative impact on profit (Jordan 1988). In addition, despite reductions in airline fares, Delta airline just reported a record-high profit in the second quarter of 2023 (Holmes 2023). In the wake of a potential new wave of carrier consolidation post Covid-19 (Primack 2022), studying profitability synergy could also provide meaningful insights about profitability management for policymakers engaging in potential airline mergers.

The effect of a merger, as well as other policy interventions, can be studied from three different perspectives: instantaneous effect, average effect, and long-term effect (de Chaisemartin and D'Haultfoeuille 2020, Callaway and Sant'Anna 2021, Sun and Abraham 2021). Instantaneous effect captures the immediate effect of mergers in the time period (i.e., the month, the quarter, and etc.) when the merger event occurred. Average effect, by its very name, averages the effect of mergers within a specified time period following mergers, such as two years. Long-term effect, also referred to as dynamic effect (Baker et al. 2022), is more commonly known as event study in the operations management (OM) field and aims to examine how the effect of mergers evolves over time. Reviewing the 31 U.S. airline merger studies (Appendix A), we see that most of the studies explored the average effect of mergers while a few studies examined how merger effect changes over time. The instantaneous effect has not been investigated in any

of these studies. However, management literature considers merger a rare event that will likely have a more pronounced instantaneous effect than a long-term effect (Lampel et al. 2009, Zollo 2009). Therefore, investigating the instantaneous effect may provide more important insights to policymakers. Accordingly, our first research objective is to investigate the instantaneous effect of mergers on profitability.

In addition to the instantaneous effect, policymakers may also be interested in understanding the long-term effect of mergers. For example, if a merger has a positive effect on profitability, does the positive effect change over time or remain constant? If the positive effect changes over time, what is the recovery path of profitability – a steady increase quarter over quarter or with some fluctuations from time to time? These questions are important as they may help policymakers design more specific and detailed post-merger plans to be implemented at different stages. However, merely presenting an average effect will not address these questions. Accordingly, examining the long-term effect of mergers on profitability becomes our second research objective. Both our first and second research objective answer the practical question of *when* airlines can expect to benefit from mergers.

Airlines have long touted revenue synergy as the key benefit to engage in mergers (Carey 2005, Delta 2008) and cited revenue synergy as the primary motivation for seeking approval from regulators (Merkert and Morrell 2012, Hüschelrath and Müllera 2014, Manuela et al. 2016). However, Schosser and Wittmer (2015) found that North American airline mergers show very little evidence of revenue synergy. In addition, in various meta-analysis, scholars have also revealed that not every firm could benefit from mergers (King et al. 2004, Homberg et al. 2009) with one reason being that the acquirer's pre-merger performance is a strong indicator of post-merger success (Zott 2003, Haleblian et al. 2009). Therefore, we also investigate the differentiating effect of mergers on profitability by comparing high-performing acquirers and low-performing acquirers (i.e., measured by their respective pre-merger profitability level) to answer the practical question of *who* benefit from airline mergers.

Drawing on organizational learning framework, we predict that profitability will deteriorate immediately upon mergers – as for individual carriers, mergers can be considered as rare events which could have a short-term negative impact on firm performance (Lampel et al. 2009, Zollo 2009). We further propose that the negative impact on profitability will keep worsening before it eventually improves – a U-shaped performance curve predicted by the learning by doing concept (Levitt and March 1988). For the differentiating effect, we hypothesize that low-performing acquirers will experience more pronounced profitability synergy than high-performing acquirers – as explained by the law of diminishing synergy (Schmenner and Swink 1998).

To test our hypothesized relationships, we collect data from the Department of Transportation (DOT) and study eight mergers from 2004Q1 to 2019Q4. Our research is a natural/quasi experiment that draws statistical inference for merger effect (the treatment effect) by comparing acquirer's performance (the

treatment group) with non-acquirer's performance (the control group) before and after mergers – a Difference-in-Difference (DiD) design. As the eight mergers occurred at different times, our design is also known as a staggered DiD design. Two-way fixed-effects (TWFE) DiD regression is a prevalent approach adopted in the OM field to estimate the treatment effect in a staggered design. However, the latest advancement in econometrics has shown that TWFE DiD regression is prone to yield biased estimates and misleading inferences when there is variation in treatment timing (Goodman-Bacon 2021, Callaway and Sant'Anna 2021, Sun and Abraham 2021). Accordingly, the latest econometrics literature proposes alternatives to estimate the treatment effect in a staggered DiD design to avoid the biases associated with TWFE. We adopt one of the alternatives – stacked regression – to estimate merger effect due to the ease of incorporating covariates in the model. To this end, our methodological approach completes our third research objective – to remind researchers of the potential biases associated with TWFE DiD in a staggered design. We accordingly call for researcher to adopt alternative estimators to achieve better research rigor.

Our results indicate that mergers do not have an instantaneous effect on profitability. However, profitability does experience a U-shaped curve within the first two years post-merger before profitability starts to stabilize (for high-performing acquirers) or improve (for low-performing acquirers) from the third year. Low-performing acquirers indeed benefit from profitability synergy while high-performing acquirers do not benefit from it, as predicted by the law of diminishing synergy.

Our study makes several contributions to the OM literature. First, we contribute to enrich the understanding of U.S. airline merger literature by investigating the largely-ignored profitability synergy. In contrast to the popular findings of positive revenue synergy resulting from increased fares within the first two years post-merger (Appendix A), we find that within the first two years, profitability synergy experienced a U-shaped curve, reinforcing the fact that revenue synergy does not necessarily translate into profitability synergy, at least in the first two years post-merger. In contrast to Schosser and Wittmer (2015) whose findings show little evidence of revenue synergy for U.S. airline mergers in the long run, our study reveals that U.S. airline mergers do, however, benefit from profitability synergy – but not until the third year and beyond after the merger event. Second, the post-merger U-shaped performance curve corroborates the learning by doing concept which states that organizational changes normally worsen a firm's performance before eventually helping to improve it (Levitt and March 1988). The differentiating effect of low-performing and high-performing acquirers also validates the law of diminishing synergy in that if a firm is already at a high level of performance, further synergy would be challenging to achieve (Schmenner and Swink 1998). To this end, we contribute to broadening the understanding and application of these two theories in the OM field. Third, we contribute to enhancing research validity of DiD design in the OM field by introducing new alternatives to estimate treatment effect in a staggered DiD design, in contrast to the widely adopted TWFE DiD regression that is susceptible to biased and misleading results (Callaway and

Sant'Anna 2021, Baker et al. 2022). To this end, we call for researchers to at least use the alternatives to estimate treatment effect as a robustness test to ensure research rigor and stringency. In addition, the widespread practice of using event study to test the parallel trend assumption is also problematic in a staggered design as it can result in both Type I and Type II errors (cf. section 3.2.1 in Baker et al. 2022). We also recommend researchers use alternative estimators to test the parallel trend assumption as this assumption is paramount to the validity of all DiD designs.

Managerially, our study provides meaningful insights for airline policymakers to engage in potential airline mergers by answering the questions of who benefit from mergers and when. Our study is especially helpful in the post-Covid period when a potential new wave of carrier consolidation could happen (Primack 2022). First, who should seek a merger? In terms of profitability synergy, our results indicate that only low-performing acquirers benefit from mergers. Applying this finding to the recent Spirit merger proposal where both Frontier and JetBlue are the potential acquirers, we would advise both Frontier and JetBlue to proceed with caution as both of them can be classified as high performers based on their financial performance while high performers did not benefit from profitability synergy, as was revealed in our findings from the eight recent U.S. airline mergers. Especially, among the eight mergers studied, Frontier was the acquirer who suffered the most from profitability deterioration in its 2010 merger with Midwest. Most likely, merging with Spirit will just repeat the history for Frontier. Second, despite airlines have touted immediate revenue synergy as the key benefit to seek mergers, it takes time for profitability synergy to take effect. Policymakers and investors who seek a short term profit boost should reconsider potential airline mergers as profitability synergy does not materialize till the third year post-merger. In other words, policymakers engaging in airline mergers should look at long-term rather than short-term returns. Third, how can acquirers maximize profitability synergy? Our results show that after the U-shaped performance curve in the first two years post-merger, profitability improves starting from the third year. Therefore, if acquirers can manage to minimize the negative impact in the first two years post-merger, they can expect better profitability synergy to come. In other words, the first two years post-merger is the crucial stage to ensure greater success in the bottom-line performance.

We first review related literature to develop hypotheses. We then use the latest estimators in econometrics field to test our hypotheses, followed by various robustness tests. Our study concludes with contributions and call for more applications of the latest estimators in staggered DiD design in the OM community.

## 2. LITERATURE REVIEW – Performance Implications of Mergers in the U.S. Airline Industry

Performance implications of U.S. airline mergers have been studied in terms of fares, flight frequency, stock market response, on-time performance, and market competition effect (see Appendix A for a detailed breakdown). We focus on financial performance implication as the financial health of carriers in the airline

industry crucially contributes to a strong and robust economy, generating 10.9 million jobs and supporting $1.8 trillion in total economic activity in 2016 (Federal Aviation Administration 2020).

Regarding financial performance implication itself, researchers have investigated the impact of mergers on fares (Hüschelrath and Müllera 2014, 2015; Jain 2015), shareholder value (Singal 1996b, Gong and Firth 2006), revenue synergy (Schosser and Wittmer 2015), and profitability (Jordan 1988). Seventeen out of the 31 U.S. airline merger studies (Appendix A) examined merger impact on fares and 16 studies subsequently found that mergers substantially increased fares following mergers. Singal (1996b) examined 14 mergers between 1985 and 1988, concluding that merger announcements increased stock prices. Gong and Firth (2006) studied 15 mergers between 1985 and 2001 and found that both acquirer and target airlines experienced positive stock market responses in the wake of merger announcements. Considering impacts to revenue synergy, Schosser and Wittmer (2015) analyzed six large international airline mergers, including two in the U.S., and concluded that North American airline mergers showed "little evidence for revenue synergies" (p. 148). Regarding merger impact on profitability, Jordan (1988) examined 24 mergers between 1985 and 1987 and found that the profits of the merged airlines declined in the years following mergers due to increased operating expenses.

We focus on one specific aspect of financial performance – profitability – to investigate merger impact for the following two reasons. First, despite 16 out of the 31 U.S. airline merger studies found a substantial increase in fares following mergers (2.5% – 23% increase), higher fares do not necessarily translate into higher profits given the merged carrier's ability to streamline operating costs also impacts profitability (Jordon 1988). Moreover, Holmes (2023) reported that Delta witnessed a record high profit in the second quarter of 2023 despite falling ticket prices. Therefore, we elect to use the bottom-line measure profitability to measure financial performance impact, which is more pertinent to what shareholders' value and what C-level teams desire. Second, airline scholars have cited revenue synergy (i.e., "the sum of merging two firms is greater than their individual parts", King et al. 2004, p. 197) as a primary motivation for airlines to seek merger approval from regulators (Merkert and Morrell 2012, Ryerson and Kim 2014, Hüschelrath and Müllera 2014, Manuela et al. 2016). Against this motivation though, Schosser and Wittmer (2015) found that North American airlines failed to benefit from significant revenue synergies. Again, revenue synergy does not equal profitability synergy. We, therefore, investigate the bottom-line profitability synergy rather than the top-line revenue synergy.

When reviewing the airline merger literature, we observe that most studies are grounded in econometric analysis with little or no discussion of theoretical foundations. In the next section, we outline the theoretical foundation we use to develop our hypotheses to increase scientific understanding and to explain how and why specific relationships lead to specific events (Wacker 1998).

## 3. THEORY AND HYPOTHESES

### 3.1 Theoretical Foundation

In the current study, our definition of merger is consistent with current practice in the airline literature (Gong and Firth 2006, Gudmundsson et al. 2017) in that merger refers to both merger and acquisition. Mergers in the airline industry have influential and long-lasting impact on merged firms, involving IT system reconfiguration, human resource integration, and operational procedure redesign, which can take years to complete (Mouawad 2012). Consequently, acquirers' post-merger performances also differ. Understanding how and why merger events influence acquirers' performance is an important theoretical consideration.

Organizational learning is a useful theoretical lens to understand the nature of merger events and their consequent impact. Drawing on the seminal work of Cyert and March (1963), organizational learning has established itself as effective in explaining mergers (Zollo and Singh 2004, Zollo 2009), including airline mergers (Prince and Simon 2017). Organizational learning is defined as the "process of improving actions through better knowledge and understanding" (Fiol and Lyles 1985, p. 803). This definition indicates that 1) organizational learning is a *process*; 2) organizational learning involves *knowledge management*. We explain how these two perspectives relate to mergers in the airline industry.

At the process level, Argote and Miron-Spektor (2011) classified organizational learning into a three-step recursive process involving knowledge search, knowledge creation/transfer, and knowledge retention. In the airline industry, carriers' motivation to seek mergers and the subsequent post-merger integration can be illustrated within Argote and Miron-Spektor's (2011) framework. Since mergers are rare events, they can be viewed as a closed-loop learning cycle starting from knowledge search, transitioning to knowledge creation/transfer, and ending with knowledge retention. Because carriers are motivated to increase their market power over time (Hüschelrath and Müllera 2014), mergers are a quick means to achieve this goal (Schosser and Wittmer 2015, Chen and Gayle 2019). To identify potential opportunities, acquirers continuously evaluate competitor carriers, assessing expected synergies a merger might create. Within Argote and Miron-Spektor's (2011) organizational learning framework, this is the pre-merger knowledge search stage. In cases when a merger results, upon the official merger closure, acquirers and target carriers begin the integration process, where knowledge creation and knowledge transfer between the two organizations play an important role in determining if the merger will be successful (Azan and Sutter 2010). Lastly, after the integration process is completed, knowledge retention becomes essential and serves as a key step in the acquirers' sustained growth and success (Marsh and Stock 2006). In cases where acquirers experience multiple mergers, this cycle repeats itself with each new merger.

At the knowledge management level, various scholars have related the concept of knowledge to organizational learning (Nonaka 1994, Wang and Ahmed 2003). Specifically, Nonaka (1994) distinguished knowledge between explicit knowledge and tacit knowledge. Nonaka (1994) claimed that explicit

knowledge can be transferred in a formal and systematic language while tacit knowledge, having a personal quality, is hard to formalize or communicate. Nonaka's (1994) concept of explicit knowledge and tacit knowledge also helps to explain the impact of mergers on financial performance in the airline industry.

Achieving synergies in profitability requires two fundamentals: improving revenues and reducing operating costs, with the former involving more tacit knowledge while the latter requires more explicit knowledge. To achieve synergies in revenues, more tacit knowledge learning is necessary since the recipe to achieve better financial performance is normally not expressly documented in firms. Sound financial performance is usually the result of collective wisdom developed through individual experiences, where tacit knowledge resides (Huselid 1995, Katzenbach and Smith 2015). Therefore, the newly merged carrier will have to diligently extract tacit knowledge from individual experiences across the merged firms to achieve financial synergies (Crossan et al. 1999). On the other hand, reducing operating costs involves more explicit knowledge learning as, in the airline industry, almost all operational procedures are meticulously documented, including ground operations (Wenner and Drury 2000), flight-deck operations (Degani and Wiener 1998), and operations control centers (Clarke 1998). With nearly all operational procedures documented, the two carriers involved in a merger need to painstakingly make necessary adjustments to develop new operational procedures for post-merger operations.

In sum, achieving profitability synergy necessitates the acquirer to efficiently and effectively manage both explicit knowledge transfer and tacit knowledge retention. Missing any of these two links might negatively impact the outcome of an airline merger.

### 3.2 The Instantaneous Effect of Mergers on Financial Performance

Instantaneous effect refers to the immediate effect following a merger. Researchers can assign the time period where the merger event occurs to measure the instantaneous effect. The time period could be a month, a quarter, or a year. Following the current practice in the airline merger literature (Prince and Simon 2017), we use quarter to measure time period. Therefore, the instantaneous effect happens in the quarter where merger events took place in our study.

Once a merger is approved by the Department of Justice, merging carriers can immediately start to integrate operations where knowledge transfer between the two carriers begins. Both general organizational learning and specific explicit knowledge transfer activities could happen during this instantaneous stage. From a general organizational learning perspective, a merger can be considered a rare event, given mergers do not occur frequently to the same carrier. In analyzing the impact of rare events on organizational learning, Lampel et al. (2009) concluded that rare events can have a negative impact on company performance. Zollo (2009) explained that this is because rare event may be viewed as a potential threat by certain people and accordingly engender excessive caution that can paralyze an organization's ability to change. Once paralyzed, carriers may need some time to recover from the paralysis and it is highly unlikely that carriers

can recover within one single quarter. Paralysis at the organization level may lead to deteriorated performance. Accordingly, we expect that a merger may have an instantaneous negative impact on financial performance in the quarter when merger occurred.

From an explicit knowledge transfer perspective, post-merger operations integration relies heavily upon existing documentation to transfer explicit knowledge between the two carriers. Existing documentation covers a wide range of areas and encompasses "thousands of procedures used by pilots and flight dispatchers, gate agents, flight attendants and ground crew" (Mouawad 2012), which will take considerable amount of time to synchronize. We argue that it is highly unlikely that the two carriers can finish integrating these "thousands of procedures" within a quarter. Unsynchronized operations procedures may lead to duplicating operating expenses, which in turn will hurt the bottom-line performance. Combining the theorization from general organizational learning and explicit knowledge transfer perspective, we propose: *H1: Profitability will deteriorate immediately upon mergers.*

**3.3 The Long-term Effect of Mergers on Financial Performance**

Some decisionmakers may be interested to learn how the merger effect unravels over time. This is typically referred to as event study and also known as dynamic effect in the econometrics literature. In Argote and Miron-Spektor's (2011) organizational learning framework, transitioning from knowledge transfer to knowledge retention is crucial as process stabilization in the format of knowledge retention indicates an organizations' full recovery from the paralysis brought by the rare event of merger. However, a persistent problem carriers faced during mergers is the extended period of the frustrated knowledge transfer process due to the extreme complexity of the integration of the "thousands of procedures" discussed above. For example, United Airlines officially completed its merger with Continental Airlines on October 1st, 2010. Two years after the official completion on November 28th, 2012, United Airlines "still grapples with myriad problems in integrating the two airlines" and "the result has been hobbled operations" (Mouawad 2012): the worst operational record among the nation's top 15 airlines and the highest rate of delayed flights in summer 2018 and a 6% below industry average rate of on-time arrivals (77.5%) in 12 months ((Mouawad 2012).). United Airlines is not alone in dealing with post-merger struggles. American Airlines officially completed its merger with US Airways on December 9th, 2013. It took five years for American Airlines to finally finish its flight attendant integration with US Airways (Josephs 2018).

Based on these industry reports, we, accordingly, argue that the instantaneously worsened financial performance is likely to keep declining for an extended period of time due to the extended period of post-merger integration likely due to higher operating expenses at least for some time. Eventually the worsened financial performance will start to improve. We explain the rationale using the learning by doing concept (Levitt and March 1988). The learning by doing concept predicts that changes generally cause performance to deteriorate before it ultimately improves. In a merger, two carriers learn from each other during the

integration process, from which new solutions to solving problems should occur. When coping with new problems, managers may tend to refer to their prior solutions to address the post-merger problems. However, resorting to pre-existing solutions can result in negative learning outcomes (March et al. 1991), which may manifest itself as deteriorated financial performance. March et al. (1991) observed that after repeated negative learning outcomes which resulted from relying upon pre-existing solutions, managers begin to develop new integrative knowledge to correct their actions. Once developed, this new form of knowledge can be retained in the new organization's memory (i.e., knowledge retention) (Marsh and Stock 2006, Levy 2011, Argote and Miron-Spektor 2011), which can help to handle post-merger challenges more effectively. As such, financial performance should start to improve. Therefore, we expect the following long term effect of mergers:

*H2: Profitability will demonstrate a U-shaped curve with time elapsing following mergers.*

**3.4 Differentiating Effects of Mergers on Financial Performance**

While performance implications brought by mergers are expected to occur for all acquirers in both the short-term and the long-term, these effects may differ for some acquirers (King et al. 2004, Homberg et al. 2009). We develop our argument for these differentiating effects in the following section.

The performance difference can be explained from the following three perspectives. First, the concept of X-efficiency, originally proposed by Leibenstein (1966), was subsequently extended to explain airline merger performance by Gudmundsson et al. (2017). The key concept of X-efficiency, applied in airline mergers, is that low-performing acquirers can utilize their managerial capabilities to achieve superior synergy when engaged with mergers, especially when the two merged firms have similar or complementary resources (Gudmundsson et al. 2017). Carriers in the airline industry indeed share highly similar or complimentary resources (Gudmundsson et al. 2017). Second, low-performing acquirers are likely to be under distress and may have resorted to merger as a key means to improve performance (Schmidt 2016). In this case, we may expect that the distressed low-performing acquirers would be more motivated to work harder to improve performance. Third, the law of diminishing synergy (Schmenner and Swink 1998) proposes that when performance is at a higher level, further improvements will yield less noticeable synergy. Conversely, when performance is at a lower level, further improvements will yield more pronounced synergy. Combined, we propose:

*H3a: Low-performing acquirers will demonstrate greater improvement in profitability synergy right upon mergers, and vice versa.*

*H3b: Low-performing acquirers will demonstrate more pronounced profitability synergy following mergers, and vice versa.*

## 4. DATA

### 4.1 Sample

To test our hypotheses, we collect data from the Department of Transportation (DOT). DOT requires U.S. carriers to report financial performance on a quarterly basis if they have more than 1% domestic market share, measured by total scheduled domestic passenger revenues. To avoid the impact of 9/11, DOT report format change in 2003, and the global pandemic starting from 2020, we construct a panel data from 2004Q1 to 2019Q4, consisting of 64 quarters in total. After cleaning financial raw data and removing carriers with incomplete reporting over the sample timeframe, our data consists of 28 carriers from 2004Q1 to 2019Q4. Some carriers span the entire 64 quarters while others report fewer quarters either due to their revenues falling below the one percent reporting threshold or due to merger and acquisition. A detailed summary of airlines used in our analysis is presented in Table 1.

Table 1 Airlines in the Dataset

| No. | Airline | First quarter in the sample | Last quarter in the sample | Total quarters in the sample |
|---|---|---|---|---|
| 1 | AIRTRAN | 2004 Q1 | 2011 Q1 | 33 |
| 2 | ALASKA | 2004 Q1 | 2019 Q4 | 64 |
| 3 | ALEEGIANT | 2018 Q1 | 2019 Q4 | 8 |
| 4 | ALOHA | 2006 Q2 | 2008 Q1 | 8 |
| 5 | AMERICA WEST | 2004 Q1 | 2005 Q4 | 8 |
| 6 | AMERICAN | 2004 Q1 | 2019 Q4 | 64 |
| 7 | ATA | 2004 Q1 | 2006 Q4 | 12 |
| 8 | ATLANTIC SOUTHEAST | 2004 Q1 | 2011 Q4 | 32 |
| 9 | COMAIR | 2004 Q1 | 2010 Q4 | 28 |
| 10 | CONTINENTAL | 2004 Q1 | 2011 Q4 | 32 |
| 11 | DELTA | 2004 Q1 | 2019 Q4 | 64 |
| 12 | ENDEAVOR | 2010 Q4 | 2019 Q4 | 13 |
| 13 | ENVOY | 2004 Q1 | 2019 Q4 | 56 |
| 14 | EXPRESSJET | 2004 Q1 | 2019 Q4 | 62 |
| 15 | FRONTIER | 2005 Q2 | 2019 Q4 | 59 |
| 16 | HAWAIIAN | 2004 Q1 | 2019 Q4 | 64 |
| 17 | INDEPENDENCE | 2004 Q1 | 2005 Q4 | 8 |
| 18 | JETBLUE | 2004 Q1 | 2019 Q4 | 64 |
| 19 | MESA | 2006 Q1 | 2019 Q4 | 40 |
| 20 | NORTHWEST | 2004 Q1 | 2009 Q4 | 24 |
| 21 | PSA | 2018 Q1 | 2019 Q4 | 8 |
| 22 | REPLUBLIC | 2018 Q1 | 2019 Q4 | 8 |
| 23 | SKYWEST | 2004 Q1 | 2019 Q4 | 64 |
| 24 | SOUTHWEST | 2004 Q1 | 2019 Q4 | 64 |
| 25 | SPIRIT | 2015 Q1 | 2019 Q4 | 20 |
| 26 | UNITED | 2004 Q1 | 2019 Q4 | 64 |
| 27 | US AIRWAYS | 2004 Q1 | 2015 Q2 | 41 |
| 28 | VIRGIN AMERICA | 2012 Q1 | 2017 Q4 | 24 |

*Notes:*

1. *RU was used from October 2003 to June 2006 by DOT to code ExpressJet. Effective July 2006, ExpressJet changed in DOT report from RU to XE. In our dataset, RU was changed to XE.*
2. *American Eagle Airlines changed to Envoy effective April 2014 in DOT report. Both Envoy and American Eagle were coded as ENVOY in our data.*
3. *Atlantic Coast Airlines changed to Independence Airline since 2004 November in DOT report. Both airlines were coded as Independence in our data.*
4. *Endeavor Air, formerly Pinnacle Airlines, was ranked for the first time in January 2013 DOT report. Both Pinnacle and Endeavor were coded as Endeavor in our data.*
5. *Atlantic Southeast (EV) was acquired by ExpressJet and changed to XE since January 2012 in DOT report.*
6. *Low-Cost Carriers: Allegiant, Frontier, JetBlue, Southwest, Spirit, and Virgin America.*

To determine mergers and the associated dates in our data time window, we review various news bulletins (Appendix B), such as airline news releases and CNN, and identify eight mergers which occurred from 2005 to 2016 as shown in Table 2.

Table 2 List of Merged Carriers

| Carrier | Merger Announcement Date | Merger Completion Date | Merger Completion Quarter | Quarters Before Mergers | Quarters After Mergers |
|---|---|---|---|---|---|
| Alaska/Virgin America | April 4, 2016 | December 14, 2016 | 2016Q4 | 51 | 13 |
| American/US Airways | February 14, 2013 | December 9, 2013 | 2013Q4 | 39 | 25 |
| Delta/Northwest | April 14, 2008 | October 29, 2008 | 2008Q4 | 17 | 47 |
| ExpressJet/Atlantic Southeast | August 4, 2010 | December 31, 2011 | 2011Q4 | 31 | 28 |
| Frontier/Midwest | April 13, 2010 | October 1, 2010 | 2010Q4 | 27 | 37 |
| Southwest/Air Tran | September 27, 2010 | May 2, 2011 | 2011Q2 | 29 | 34 |
| United/Continental | May 3, 2010 | October 1, 2010 | 2010Q4 | 27 | 37 |
| US Airways-America West | May 19, 2005 | Sep 27, 2005 | 2005Q3 | 6 | 32 |

*Note: First carrier is the acquirer.*

### 4.2 Measures

#### Profitability Synergy

Our dependent variable is profitability synergy. We elect to use the profitability measure as discussed in Section 2.2. Profitability measures in the airline industry take three different forms in the extant airline literature: reported profitability (Kalemba and Campa-Planas 2017) and operating profit over operating revenue (Tsikriktsis 2007, Mellat-Parast et al. 2015). We elect to use operating profit over operating revenue (OPOR) instead of reported profitability as our financial performance measure for two reasons. First, profitability varies across years and is sometimes negative. When natural logarithms are calculated, those negative profitability values become missing data points, which is not a true reflection of airline financial status. Second, the excessive variance of profitability comes from size differences among carriers. Ratio measures like OPOR, in this case, can better account for the size differences among carriers in comparison to other financial measures (Dresner and Xu 1995). In addition, ratio measures can also overcome the difficulty in measures associated with carriers that own aircraft versus carriers that lease aircraft (Tsikriktsis 2007, Mellat-Parast et al. 2015). If an acquirer's post-merger profitability improved compared with its pre-merger level, profitability synergy is therefore achieved. Otherwise, profitability synergy was not achieved.

Operating profit and operating revenue data were retrieved from DOT Schedule P1.2. Note that financial performance data is only released quarterly by DOT, i.e., OPOR can only be computed as a quarterly measure. Therefore, all other relevant variables used in this study are subsequently converted into quarterly data to match OPOR.

#### Merger Event

Current airline merger research typically utilizes the officially released merger completion date to define pre- and post-merger time windows (Jain 2015, Prince and Simon 2017). However, after the official merger

completion date, carriers sometimes continue to report to DOT as two individual carriers, before eventually reporting to DOT as one single carrier. To address the issue of overlapping reporting, we aggregate the overlapped DOT records based on the official merger completion dates to align with current practices in the airline literature.

**Control Variables**

To address potential endogeneity bias, an area of challenge in airline research (Scotti and Dresner 2015), we control for factors that might influence the relationships between merger events and carriers' financial performance following current practices in the airline literature (Prince and Simon 2017, Alan and Lapré 2018). We list these control variables below.

*Load Factor*

Load factor measures how full the aircraft is based on passengers, which directly impacts revenue and, in turn, impacts the bottom-line performance. DOT's definition of load factor (i.e., revenue passenger miles divided by available seat miles) was adopted. Relevant data were retrieved from DOT Schedule T1. The data were in a monthly format and subsequently collapsed into a quarterly format.

*Yield*

Yield management in the airline industry refers to the techniques used to allocate limited seats to different customer segments by offering differential airfares (Netessine and Shumsky 2002). Using proper yield management systems, American Airlines credited a revenue increase of $500 million per year and Delta Airlines generated additional revenues of $300 million per year (Boyd 1998). Following current airline literature (Alan and Lapré 2018), we calculate yield as passenger revenues divided by revenue passenger miles. Yield data were obtained from DOT Schedule T1 and P1.2.

*Fleet Utilization*

Higher fleet utilization was considered a key driver to Southwest's success (Gittell 2003). Southwest Airline boasts an impressive 20-30 minutes turnaround time while the industry average is 1.5-2 hours (Mantin and Wang 2012). Belobaba (2009) also recommended airlines to improve fleet utilization to achieve better financial performance. Accordingly, we include fleet utilization as another control variable. We calculate fleet utilization as block hours divided by aircraft days following Alan and Lapré (2018). Relevant data were retrieved from DOT Schedule P52.

*Number of Enplaned Passengers*

The number of enplaned passengers was included as a control variable because carriers can expect increased revenue as the number of passengers increases, which may accordingly help boost profit. We retrieve the number of passengers from DOT Monthly Consumer Report.

*Fuel Cost*

Fuel is an important control variable for airline research given it is a key factor affecting airlines' financial performance. Typically, fuel has been either examined as a price or a cost (Ramdas et al. 2013, Dana and Orlov 2014). Since some carriers hedged their fuel requirements in order to leverage a fuel price below the market rate, we consider fuel cost to be a more appropriate control variable. To demonstrate this difference, the average fuel price fluctuated from $111.8 per barrel (2012) to $44.6 per barrel (2016) in our sample data period (IATA 2018). In the second half of 2005, Southwest hedged 85% of its fuel requirements at the equivalent of $26 while the industry average was $72.35 (Alexander 2006). Fuel cost was reported on a monthly basis in Schedule P12(a), which was aggregated to quarterly format.

*Firm Size*

As larger firms are expected to "have higher levels of resources and more developed market positions, it is important to control for the size of the firm" (Mishina et al. 2004, p. 1189). We use the number of employees as a proxy for firm size because more employees add more costs and negatively impact operating profit. Indeed, employee expenses are the second largest element impacting carrier's profit after fuel cost (IATA 2018). The number of employees was taken from DOT Schedule P1(a).

*Delays*

Delays, either airborne or ground, result in significant costs to carriers, impacting carriers' financial performance (Cook et al. 2012). As such, we include delays as a control variable in our models. Delay data were compiled from DOT Air Travel Monthly Consumer Report.

Table 3 summarizes the variables used in this study and Table 4 presents the observed means, standard deviations, and correlations for all variables in interest. Regarding carrier-specific control variables, most of the correlations are as expected, such as load factor and yield being positively correlated with OPOR.

Table 3 Variables Used in Analysis

| Variable | Formula or Definition | Data Source |
|---|---|---|
| OPOR | Operating profit divided by operating revenue at quarterly level | DOT Schedule P1.2 |
| Load Factor | Revenue passenger miles divided by available seat miles | DOT Schedule T1 |
| Yield | Passenger revenues divided by revenue passenger miles (RPMs) for each carrier | DOT Schedule T1 and P1.2 |
| Fleet Utilization | Block Aircraft Hours divided by Aircraft Days | DOT Schedule P52 |
| Number of Enplaned Passengers | Number of enplaned passengers | DOT Schedule T1 in Form 41 Air Carrier Summary Data |
| Fuel Cost | Quarterly fuel cost for each carrier | DOT Schedule P12(a) |
| Firm Size | Number of Full-Time Equivalent Employees | DOT Schedule P1(a) in Form 41 Financial Data |
| Delays | Quarterly sum of delays reported to DOT | DOT Air Travel Monthly Consumer Report |

Note: Monthly data was aggregated into quarterly data

Table 4 Summary Statistics and Correlation Matrix

| | | Panel A | | | Panel B | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | OPOR | 1050 | 0.039 | 0.148 | 1.000 | | | | | | | |
| 2 | Load Factor | 1061 | 0.805 | 0.070 | 0.369 | 1.000 | | | | | | |
| 3 | Yield | 1050 | -2.023 | 0.283 | 0.116 | -0.318 | 1.000 | | | | | |
| 4 | Fleet Utilization | 1050 | -1.100 | 0.275 | 0.008 | 0.196 | -0.347 | 1.000 | | | | |
| 5 | Enplaned Passengers | 1058 | 15.607 | 0.897 | 0.140 | 0.219 | 0.000 | 0.158 | 1.000 | | | |
| 6 | Fuel Cost | 974 | 18.912 | 1.587 | 0.087 | 0.238 | 0.168 | 0.400 | 0.650 | 1.000 | | |
| 7 | Number of Employees | 1050 | 10.514 | 1.117 | 0.022 | 0.189 | 0.079 | 0.185 | 0.933 | 0.719 | 1.000 | |
| 8 | Flight Delays | 1061 | 8.223 | 0.960 | 0.106 | -0.127 | 0.325 | -0.168 | 0.770 | 0.361 | 0.679 | 1.000 |

## 5. ANALYSIS

### 5.1 Research Design

From Table 1 and Table 2, we see that our data includes eight carriers that have been engaged in mergers and 20 carriers that have not experienced any mergers. This provides us a natural experiment for us to draw causal inference of merger effects. In econometrics terminology, the merger event is known as the "treatment"; carriers that have experienced mergers in our data are referred to as "treated" units or eventually-treated units while carriers that have never experienced mergers in our data are called "control" units or never-treated units; the impact of merger on the carriers who experienced merger events is accordingly called "treatment effect". Table 2 also shows that the eight mergers occurred at different times – commonly referred to as "staggered design", meaning there is variation in treatment timing (i.e., different merger dates) across multiple time periods (i.e., 64 quarters). Note that the current study is a single airline industry study (Tsikriktsis 2007, Alan and Lapré 2018) where all carriers are already highly comparable to each other and 28 carrier is the total available population size in this industry. Following existing airline research (Tsikriktsis 2007, Alan and Lapré 2018), therefore, we did not adopt the commonly used matching or synthetic control method to match samples.

Also known as quasi-experiment, natural experiment research design is widely applied in the OM field by using two way fixed effects (TWFE) Difference-in-difference (DiD) regression to draw causal inference for treatment effects (Dong et al. 2019, Cui et al. 2019, Li and Wu 2020, Cui et al. 2022). However, with the recent advancement in econometrics, TWFE DiD regression has been proven to produce biased estimates in a staggered design (Callaway and Sant'Anna 2021, Sun and Abraham 2021). Although the econometrics literature has settled on the drawbacks of TWFE DiD in estimating treatment effects in a staggered design, which alternative estimator is the best still remains debated. The advancement of econometrics may continue furthering this topic. However, at the time of writing, we recommend three alternatives that can be used to avoid the potential bias associated with TWFE in a staggered design: Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Stacked Regression (Deshpande and Li 2019, Cengiz et al. 2019). The commonality of these three alternatives is that they all modify those units

that can serve as effective controls (i.e., avoid using already-treated units as effective controls). The difference between the three alternatives is that neither Callaway and Sant'Anna (2021) nor Sun and Abraham (2021) requires data restructuring while stacked regression requires researchers to restructure the data. In addition, it is relatively easy to incorporate covariates when using Callaway and Sant'Anna (2021) and stacked regression. In this study, we elect to use stacked regression to estimate the effects of mergers as this alternative may be the most familiar to empirical researchers.

Unlike Callaway and Sant'Anna (2021) or Sun and Abraham (2021) that uses a completely new estimator, stacked regression per se is not an estimator as stacked regression also uses TWFE. The essence of stacked regression is to construct event-specific 2×2 datasets (i.e., avoid using already-treated units as effective controls in any of the 2×2 comparisons). Then, these clean 2×2 datasets were stacked together and TWFE can be readily applied on the stacked data (c.f. Deshpande and Li 2019 replication package for data construction details). Therefore, the name "stacked" regression.

**5.1 Instantaneous Effect of Mergers on Financial Performance**

Instantaneous effect occurs immediately following mergers, where "routines and expertise from the pre-change period are immediately transferred to the changed task" (Lang and Bliese, 2009, p. 415). A popular model to estimate instantaneous treatment effect is a TWFE regression model from social science (Bliese and Lang 2016). However, recent development in econometrics reveal that using TWFE regression to estimate instantaneous effect is problematic (de Chaisemartin and D'Haultfoeuille 2020). The reason is that when OLS computes the treatment effects by comparing all possible 2×2 comparisons in consecutive time periods, if any of the effective controls are treated at the time periods both before and after the treatment, their treatment effect will be differenced out at the second period, resulting in negative weights, leading to biased estimates for the treatment effect.

To cope with the negative weights issue, de Chaisemartin and D'Haultfoeuille (2020) proposed to estimate the following coefficient $\delta^S$ (Equation 1) using a different estimator $DID_M$ (cf. p. 2978 for more detailed interpretations). de Chaisemartin and D'Haultfoeuille (2020) showed that the new estimator $DID_M$ is an unbiased and consistent estimator that is robust to treatment effect heterogeneity across groups and time periods in both a single treatment design and a staggered design.

$$\delta^S = E\left[\frac{\mathbf{1}}{N_S}\sum_{(i,g,t):t\geq 2, D_{g,t}\neq D_{g,t-1}}\left[Y_{i,g,t}(\mathbf{1}) - Y_{i,gt}(\mathbf{0})\right]\right] \quad \text{Equation 1}$$

In a staggered design where there is variation in treatment timing, $\delta^S$ is the average of treatment effects at the time when all units start receiving treatment by comparing units whose treatment status has changed with those units whose treatment status is unchanged between $t-1$ and $t$ ($\left[Y_{i,g,t}(\mathbf{1}) - Y_{i,gt}(\mathbf{0})\right]$). Therefore, $\delta^S$ captures the average treatment effect among switchers right at the time period when they switch and avoids the negative weights and the biased estimates issue discussed above. In a staggered design as in our

current study, $\delta^S$ is the average of treatment effects at the time when all carriers start receiving treatments (i.e., started to merge with another carrier) by comparing carriers whose treatment status has changed with those carriers whose treatment status is unchanged between *t–1* and *t*. Therefore, $\delta^S$ captures the average treatment effect among carriers right at the time period when they merge. Table 5 illustrates how the data was coded to estimate the instantaneous effect using US Airways (Carriercode 27) as an example. US Airways merged with America West in 2005Q3 – the 7th occasion. Therefore, the indicator variable (IE_Indicator) was coded as 1 for occasion 7 and 0 otherwise. If a carrier has never experienced mergers (control group), then the indicator will be all 0 for this carrier.

Table 5 Illustrative Data to Estimate Instantaneous Effect

| Carriercode | Occasion | OPOR | IE_Indicator |
|---|---|---|---|
| 27 | 1 | -0.143 | 0 |
| 27 | 2 | -0.101 | 0 |
| 27 | ... | ... | 0 |
| 27 | 6 | -0.010 | 0 |
| **27** | **7** | **-0.075** | **1** |
| 27 | 8 | -0.046 | 0 |
| 27 | 9 | -0.010 | 0 |
| 27 | 10 | 0.083 | 0 |
| 27 | 11 | 0.056 | 0 |
| 27 | ... | ... | 0 |
| 27 | 40 | 0.015 | 0 |

We use the estimator $DID_M$ of de Chaisemartin and D'Haultfoeuille (2020) to estimate the instantaneous effect and report our result in Table 6. A prerequisite to use $DID_M$ estimator is the stable groups assumption, which holds in our case: between each pair of consecutive quarters, there are carriers whose treatment status does not change. The estimate for $DID_M$ in Table 6 ($DID_M$ =–0.006) is not statistically significant. Therefore, Hypothesis 1 was not supported, indicating that merger does not have an instantaneous effect on profitability.

Table 6 Instantaneous Effect of Mergers on Financial Performance

| | Estimate | Standard Error |
|---|---|---|
| $DID_M$ | -0.006 | 0.034 |
| *(Covariates included in model but omitted in reporting)* | | |

Note: This table reports the estimates for the instantaneous effect $DID_M$.
*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

### 5.2 Long Term Effect of Mergers on Financial Performance

Hypothesis 2 seeks to answer the question of how the effect of mergers evolves over time. This is commonly known as an event study methodology which allows treatment effects to vary over time non-parametrically. Researchers typically are interested in estimating the coefficients of relative time indicators after the treatment. These coefficients are interpreted as the average treatment effect at different lengths of exposure to the treatment. The baseline model of a standard event study is illustrated in Equation 2:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{l=-K}^{-1} \mu_l \boldsymbol{D_{it}^l} + \sum_{l=1}^{L} \mu_l \boldsymbol{D_{it}^l} + \beta X_{it} + \varepsilon_{it} \quad \text{Equation 2}$$

$\alpha_i$ and $\lambda t$ are vectors of carrier fixed effects and time fixed effects respectively. $\boldsymbol{X}_{it}$ is the vector of time-varying control variables and $\varepsilon_{it}$ is the error term. $D_{it}^l$ is a set of relative time indicators. $l$ = (–K, ... , L) represents the length of time periods relative to the time period when the merger occurred, such as (–2, –1, 0, 1, 2, 3) where 0 is when the merger started, –2 is two time periods before the merger, and 2 is two time periods after the merger. In an event study specification, it is necessary to exclude some relative time periods to avoid multi-collinearity. The most common practice is to exclude relative periods close to the initial treatment (Sun and Abraham 2021). When there are no never-treated units in a panel balanced data, at least two relative time periods need to be excluded (Sun and Abraham 2021). In the current study, time period of $t = 0$ was omitted to avoid multicollinearity. $\sum_{l=-K}^{-1} \mu_l D_{it}^l$, therefore, captures the time periods up to the first quarter before merger. $\sum_{l=1}^{L} \mu_l D_{it}^l$ represents the time periods after merger ($l$ = 1). The main parameters of interest in Equation 2 are the $\mu_l$s which captures the differences in the outcome $Y_{it}$ between treated and untreated carriers $l$ time periods apart from the merger (i.e., the quarter when the merger started).

Applying TWFE directly to Equation 2 to estimate $\mu_l$s has also been proven to be problematic. A decomposition of the regression coefficient $\sum_{l=0}^{L} \mu_l D_{it}^l$ on the relative time indicators reveals that these coefficients are a linear combination of the average treatment effect from its own time period as well as from other relative time periods (Sun and Abraham 2021). In addition, the weights associated with these coefficients are non-linear functions of the distribution of the groups, and these weights are still prone to the issue of negative weights as discussed in the previous sections (Sun and Abraham 2021). These two factors together contaminate the estimation of $\sum_{l=0}^{L} \mu_l D_{it}^l$. Among the several alternatives, we use stacked regression due to its ease to incorporate covariates. The only difference between applying TWFE directly on the original data and applying TWFE on a stacked data is that a stacked data defines main variables in a clean event-specific dataset, i.e., already-treated units are not used as effective controls. When TWFE estimates the DiD on each clean 2×2 dataset, the unit and time fixed effects ($\alpha_i$ and $\lambda t$) are saturated with dataset indicators. Then TWFE applies variance weighting to combine the treatment effects across groups and time periods (Deshpande and Li 2019, Cengiz et al. 2019). Thus, avoiding the contamination issue mentioned above.

We use stacked regression to estimate Equation 2. Following airline merger literature (Prince and Simon 2017), we estimate merger effect up to 20 quarters (5 years) post and 4 quarters before merger. Table 7 is an example of the stacked data used in our study. Carrier code 27 (US Airways) merged with America West on Sep 27, 2005 (2005Q3) – the 7th Occasion in our data. Therefore, we assign a data identifier (dt) 7 for US Airways. Because we elect to estimate the effect of 4 quarters prior and 20 quarters post-merger, the

Occasion column starts from 3 (4 quarters prior to event occasion 7) and ends at 27 (20 quarters after event occasion 7). Relative time indicators (rel_-1 .... rel_20) are the *l* time periods in Equation 2. The time period when merger occurred (rel_0) was omitted in estimation to avoid multicollinearity.

Table 7 Illustrative Stacked Data using US Airways

| Carriercode | Occasion | rel_occasion | OPOR | dt | rel_-1 | rel__-2 | ... | rel_20 |
|---|---|---|---|---|---|---|---|---|
| 27 | 3 | -4 | -0.143 | 7 | 0 | 0 | 0 | 0 |
| 27 | 4 | -3 | -0.101 | 7 | 0 | 0 | 0 | 0 |
| 27 | 5 | -2 | -0.139 | 7 | 0 | **1** | 0 | 0 |
| 27 | 6 | -1 | -0.010 | 7 | **1** | 0 | 0 | 0 |
| 27 | **7** | **0** | -0.075 | 7 | 0 | 0 | 0 | 0 |
| 27 | 8 | 1 | -0.046 | 7 | 0 | 0 | 0 | 0 |
| 27 | 9 | 2 | -0.010 | 7 | 0 | 0 | 0 | 0 |
| 27 | 10 | 3 | 0.083 | 7 | 0 | 0 | 0 | 0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 27 | 25 | 18 | 0.001 | 7 | 0 | 0 | 0 | 0 |
| 27 | 26 | 19 | 0.094 | 7 | 0 | 0 | **1** | 0 |
| 27 | 27 | 20 | 0.073 | 7 | 0 | 0 | 0 | **1** |

To use stacked regression, we first need to create a stacked data. Use data identifier 7 (US Airways) as an example, we first identify all possible clean 2×2 comparisons for data identifier 7 (US Airways). Carriers 2, 6, 14, 15, 24, 26 were identified as clean effective controls for data identifier 7. We then construct a similar copy of data as is shown in Table 7 for each effective control using all variables of interest and each effective control shares the same data identifier 7. Then, all these clean 2×2 comparisons for data identifier 7 are stacked together. Similar process repeats for all other carriers that have engaged in mergers. The final data is a stacked data consisting of all possible clean 2×2 comparisons for all the eight carriers engaged in mergers.

In our stacked regression, all statistical inferences use clustered bootstrapped standard errors at the carrier level to account for autocorrelation in the data (Wooldridge 2010). Figure 1 and Table 8 reports the results. In Figure 1 (and the rest of the Figures in this paper), red dots and red dashed lines represent point estimates and 95% confidence bands for pre-treatment periods. Blue dots and blue dashed lines represent point estimates and 95% confidence bands for post-treatment periods, allowing for clustering at the carrier level. Hypothesis 2 predicts a U-shaped curve for profitability following mergers. From Figure 1 and Table 9, we see that all post-merger coefficients, except for post-periods 4 and 5, are all statistically significant. Moreover, we see a clear U-shaped curve between post-period 1 and 8, after which the merger effect seems to be stabilized. Although the coefficients of post-periods 4 and 5 are in the predicted direction, these two coefficients are not statistically significant. However, we can see a clear downward-going trend before post-period 4 (all coefficients statistically significant) and a clear upward-going trend after post-period 5 (all coefficients statistically significant). Therefore, our Hypothesis 2 is partially supported.

Figure 1 Long Term Effect of Mergers – Estimated using Stacked Regression

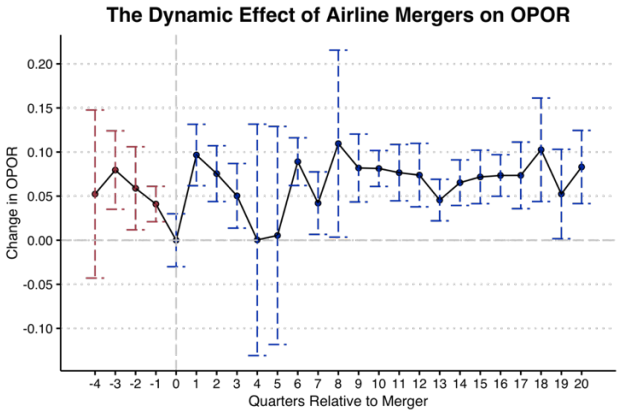



Table 8 Long Term Effect of Mergers – Estimated using Stacked Regression

| *l* Wave relative to mergers | $\widehat{\mu_l}$ | Standard Error |
|---|---|---|
| -4 | 0.052 | 0.060 |
| -3 | 0.080* | 0.036 |
| -2 | 0.059° | 0.037 |
| -1 | 0.041° | 0.024 |
| 1 | 0.097*** | 0.031 |
| 2 | 0.076** | 0.030 |
| 3 | 0.050* | 0.032 |
| 4 | 0.000 | 0.078 |
| 5 | 0.005 | 0.074 |
| 6 | 0.089*** | 0.027 |
| 7 | 0.042* | 0.031 |
| 8 | 0.109° | 0.065 |
| 9 | 0.082** | 0.033 |
| 10 | 0.081*** | 0.024 |
| 11 | 0.077** | 0.030 |
| 12 | 0.074** | 0.032 |
| 13 | 0.046* | 0.026 |
| 14 | 0.065** | 0.027 |
| 15 | 0.072** | 0.029 |
| 16 | 0.073*** | 0.026 |
| 17 | 0.074** | 0.033 |
| 18 | 0.103** | 0.043 |
| 19 | 0.052° | 0.039 |
| 20 | 0.083 *** | 0.034 |
| *(Covariates included in model but omitted in reporting)* | | |
| Adjusted $R^2 = 0.57$ | | |

*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

### 5.3 Differentiating Effects of Mergers on Financial Performance

Hypothesis 3 proposes a differentiating effect between high-performing acquirers and low-performing acquirers. First, we need to identify high and low performing acquirers. We compute the average OPOR in the respective pre-merger time periods for each acquirer. Then, we compare the acquirer's average OPOR to the industry average during the same time period. Table 9 summarizes the result. We see that among the eight acquirers, two of them perform better than industry average (Alaska and Southwest) while the rest six

perform below industry average. We accordingly label Alaska and Southwest as high-performing acquirers and the rest low-performing acquirers.

Table 9 Classify High Performing Acquirer and Low Performing Acquirer

| Acquirer | Pre-merger Period | | OPOR | |
|---|---|---|---|---|
| | Data Start | 1 Quarter Before Merger | Acquirer Average | Industry Average |
| **ALASKA** | **2004Q1** | **2016Q3** | **0.0725** | **0.0276** |
| AMERICAN | 2004Q1 | 2013Q2 | -0.0371 | 0.0115 |
| DELTA | 2004Q1 | 2008Q3 | -0.0103 | -0.0043 |
| EXPRESSJET | 2004Q1 | 2011Q3 | -0.0134 | 0.0033 |
| FRONTIER | 2004Q1 | 2010Q3 | -0.0078 | 0.0025 |
| **SOUTHWEST** | **2004Q1** | **2011Q1** | **0.0665** | **0.0115** |
| UNITED | 2004Q1 | 2010Q3 | -0.0256 | 0.0033 |
| US AIRWAYS | 2004Q1 | 2005Q2 | -0.0754 | -0.0469 |

If we estimate the differentiating effect using the original stacked data from the previous section consisting of both high-performing and low-performing acquirers, the stacked data effectively creates all possible 2×2 comparisons so long as the already-treated units are not used as controls. Accordingly, low-performing acquirer might be used as effective controls for high-performing acquirers and vice versa, contradicting to what we aim to achieve. To avoid this, we construct two separate "clean" datasets: one consists of never-treated carriers and high-performing acquirers and the other consists of never-treated carriers and low-performing acquirers. So all potential 2×2 comparisons will be low/high-performing VS never-treated, low-performing VS low-performing, and high-performing VS high-performing.

We then estimate instantaneous effect and long-term effect for high-performing and low-performing acquirers on the two clean separate datasets. Table 10 reports the instantaneous effect for high performers and lower performers. Neither of the instantaneous effects are statistically significant. Therefore, H3a is not supported (i.e., low performers did not experience greater improvement in OPOR in the quarter when the merger occurred). Figure 2 and Table 11 reports the results for testing the long-term effect for high performers and low performers. We see that none of the coefficients for the high performers is statistically significant, indicating that high performers did not benefit from profitability synergy at all. Most of the coefficients for low performers post-merger, however, are statistically significant (excluding post-occasion 4, 5, 7, 13) and show a slightly upward trend after year 3 (left graph in Figure 2). Therefore, H3b is supported, indicating that high performers demonstrate less pronounced improvement in OPOR compared to low performers, as predicted by the law of diminishing synergy (Schmenner and Swink 1998). Figure 2 also reveals that the post-merger trend of the partially supported Hypothesis 2 in Figure 1 is mostly driven by low performers.

Table 10 High Performing VS Low Performing Instantaneous Effect

| | Panel A High Performing Acquirers | | Panel B Low Performing Acquirers | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| $DID_M$ | 0.017 | 0.022 | -0.028 | 0.031 |
| *(Covariates included in model but omitted in reporting)* | | | | |

*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

Figure 2 High Performing VS Low Performing Long Term Effect

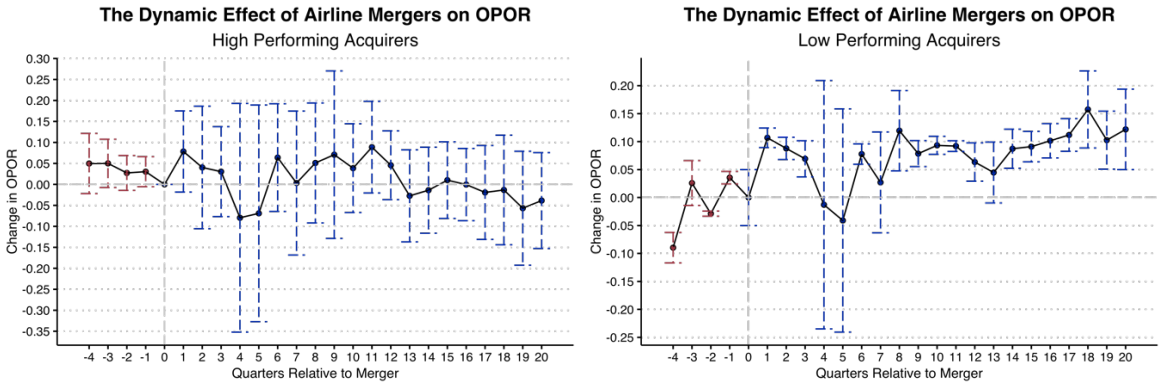



Table 11 High Performing VS Low Performing Long Term Effect

| | Panel A High-Performing Acquirers | | Panel B Low-Performing Acquirers | |
|---|---|---|---|---|
| $l$ Wave relative to mergers | $\widehat{\mu_l}$ | Standard Error | $\widehat{\mu_l}$ | Standard Error |
| -4 | 0.050 | 0.034 | (0.090)* | 0.037 |
| -3 | 0.050 | 0.027 | 0.026 | 0.043 |
| -2 | 0.027 | 0.020 | -0.029* | 0.022 |
| -1 | 0.030 | 0.017 | 0.036* | 0.029 |
| 1 | 0.078 | 0.045 | 0.107** | 0.032 |
| 2 | 0.040 | 0.069 | 0.088** | 0.033 |
| 3 | 0.030 | 0.050 | 0.069* | 0.039 |
| 4 | -0.080 | 0.128 | -0.013 | 0.130 |
| 5 | -0.069 | 0.121 | -0.041 | 0.119 |
| 6 | 0.064 | 0.060 | 0.078** | 0.033 |
| 7 | 0.003 | 0.080 | 0.027 | 0.067 |
| 8 | 0.051 | 0.067 | 0.120* | 0.058 |
| 9 | 0.071 | 0.094 | 0.078* | 0.035 |
| 10 | 0.039 | 0.050 | 0.093** | 0.032 |
| 11 | 0.089 | 0.051 | 0.092*** | 0.028 |
| 12 | 0.046 | 0.038 | 0.063° | 0.040 |
| 13 | -0.027 | 0.051 | 0.045 | 0.050 |
| 14 | -0.014 | 0.048 | 0.087** | 0.041 |
| 15 | 0.010 | 0.043 | 0.091** | 0.037 |
| 16 | -0.000 | 0.040 | 0.101*** | 0.039 |
| 17 | -0.019 | 0.053 | 0.112** | 0.038 |
| 18 | -0.013 | 0.061 | 0.157* | 0.057 |
| 19 | -0.057 | 0.064 | 0.103* | 0.049 |
| 20 | -0.039 | 0.054 | 0.122* | 0.037 |
| *(Covariates included in model but omitted in reporting)* | | | | |
| | Adjusted $R^2 = 0.524$ | | Adjusted $R^2 = 0.511$ | |

*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

### 5.4 Discussion

*5.4.1 When Does Merger Starts to Show Synergy in Bottom-line Performance?*

Airlines commonly tout revenue synergy as a key benefit of mergers. U.S. Airways in 2013 estimated an additional annual revenue of $150-200 million in their press release when merging with American West (Carey 2005). When merging with Northwest, Delta estimated a combined annual revenue of over 2 billion (Delta 2008), more than the sum of the two firms. However, revenue synergy does not equal to profitability synergy. Do airlines benefit from profitability synergy following mergers? If so, when?

Our research answers this long-ignored question and reveals that airlines do benefit from profitability synergy. However, profitability synergy only starts from approximately the third year post-merger and thereafter, as is witnessed by the upward-going trend Figure 3 (right graph). During the first two years post-merger, the merged carrier experienced a U-shaped performance curve, as predicted by the learning by doing concept (Levitt and March 1988) where changes generally cause performance to deteriorate before it ultimately improves. To this end, our findings also corroborate industry reports outlining the various conflicts following merger (Mouawad 2012) as conflicts result in reduced learning outcomes (Miller 1996). In addition, our results also echo the findings from Schosser and Wittmer (2015) who concluded that North American airlines engaged in mergers typically expect the full extraction of revenue synergies to happen over a 3-5 year period, as is witnessed in our case where profitability, instead of revenue, was studied. In sum, based on our results, profitability synergy does not happen immediately. It materializes starting from the third year following mergers.

*5.4.2 Who Benefit from Mergers in terms of Bottom-line Performance?*

King et al. (2004) conducted a meta-analysis of 93 merger studies and concluded that some subgroups of firms experienced significant positive returns while others not. Our third hypothesis and its related results reinforce the findings from King et al. (2014) in that when we split acquirers into high-performing and low-performing, only the subgroup of low-performing acquirers benefit from profitability synergy while high-performing acquirers do not benefit from profitability synergy despite that both have experienced a U-shaped performance curve in the first two years post-merger. Our findings also reinforce the law of diminishing synergy in OM field (Schmenner and Swink 1998) in that when performance is already at a high level, further synergy will become challenging to achieve, as is witnessed by the high-performing acquirers. However, when performance is at a lower level, further synergy should be easier to attain as is witnessed by the low-performing acquirers.

## 6. PARALLEL TREND ASSUMPTION AND ROBUSTNESS TEST

### 6.1 The Parallel Trend Assumption

The validity of all DiD design relies on the parallel trend assumption (i.e., the control group and the treatment group should trend at similar paces on the outcome variable before the treatment). There are two

common practices to test the parallel trend assumption in the OM field. The first practice is to compare the trendline of the average outcome on the control group and the treatment group before the treatment to visually detect if the two trendlines are trending parallelly. However, visually parallel trendlines do not equal to statistical significance. The second practice is to test the coefficients on the leads (i.e., relative time indicators leading to the treatment) using event study. If no statistical significance is found on the leads, researchers normally conclude that there is no statistical difference between the trend of the treatment group and the trend of the control group before the treatment. Hence, the parallel trend assumption holds. However, TWFE in an event study has been proved to produce biased estimates on relative time indicators of both leads and lags (Sun and Abraham 2021) in a staggered design, leading to either Type I or Type II errors (Baker et al. 2022) – meaning that the widespread practice of using event study to test for the parallel trend assumption is also problematic.

de Chaisemartin and D'Haultfoeuille (2020) accordingly proposed an alternative placebo estimator ($\text{DID}_{\text{M}}^{\text{pl}}$, from Equation 1) to test for pretrends which "essentially compares the outcome's evolution from *t − 2* to *t − 1*, in groups that switch and do not switch treatment between *t − 1* and *t*" (p. 2989). In our case, $\text{DID}_{\text{M}}^{\text{pl}}$ computes the change in OPOR between carriers experiencing mergers and those not experiencing mergers one time period before the merger event. We also extend $\text{DID}_{\text{M}}^{\text{pl}}$ and compute $\text{DID}_{\text{M}}^{\text{pl,2}}$, $\text{DID}_{\text{M}}^{\text{pl,3}}$, and $\text{DID}_{\text{M}}^{\text{pl,4}}$, which compare the change in OPOR two, three, and four periods before the merger event. We report the results in Table 12 and Figure 3. We see from Table 12 that the estimates of all the four placebo estimates (*t*=−4, −3, −2, −1) are not significantly different from 0, indicating that the parallel trend assumption was not violated: carriers who engaged in mergers do not experience different trends before the merger event compared with the carriers who did not experience mergers. We also separately test the parallel trend assumption for the two subgroups – high-performing and low-performing acquirers and report the results in Table 13 and Figure 4, from which we see that neither group has a different trend from the control group as none of the coefficients on the relative time indicators before merger is statistically significant. Therefore, we can draw the conclusion that the parallel trend assumption holds in our data.

Table 12 Pre-trend Test – All Carriers Included

| | Estimate | Standard Error |
|---|---|---|
| $\text{DID}_{\text{M}}^{\text{pl}}$ | -0.025 | 0.036 |
| $\text{DID}_{\text{M}}^{\text{pl,2}}$ | -0.002 | 0.024 |
| $\text{DID}_{\text{M}}^{\text{pl,3}}$ | 0.059 | 0.074 |
| $\text{DID}_{\text{M}}^{\text{pl,4}}$ | -0.044 | 0.067 |
| *(Covariates included in model but omitted in reporting)* | | |

Note: This table reports the estimates for the four placebos.
*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

Figure 3 Pre-trend Test – All Carriers Included

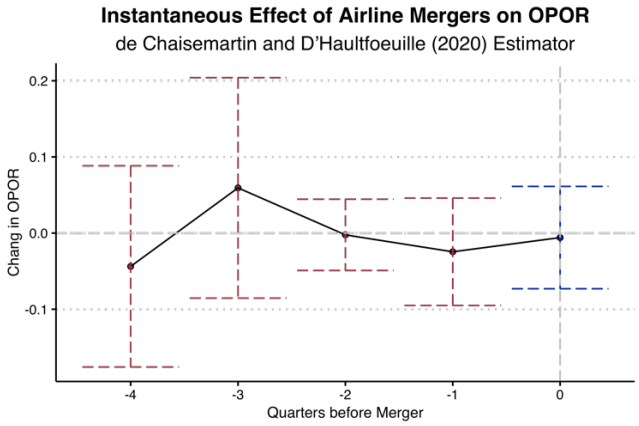
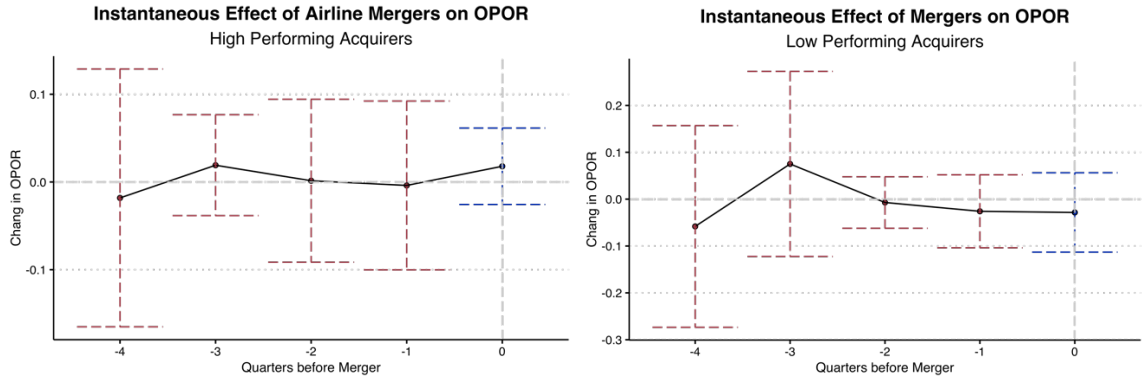



Table 13 High-performing VS Low-performing Pre-trend Test

| | Panel A (High-Performing Acquirers) | | Panel B (Low-Performing Acquirers) | |
|---|---|---|---|---|
| | Estimate | Standard Error | Estimate | Standard Error |
| $DID_M^{pl}$ | -0.004 | 0.049 | -0.026 | 0.036 |
| $DID_M^{pl,2}$ | 0.001 | 0.047 | -0.007 | 0.032 |
| $DID_M^{pl,3}$ | 0.019 | 0.029 | 0.075 | 0.075 |
| $DID_M^{pl,4}$ | -0.018 | 0.075 | -0.058 | 0.075 |
| *(Covariates included in model but omitted in reporting)* | | | | |

Note: This table reports the estimates for the four placebos.
*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

Figure 4 High Performing VS Low Performing Pre-rend Test




## 6.2 Placebo Test

From Table 2, we see that the time gap between merger announcement date and merger completion date ranges from 1-3 quarters. Carriers may start integration after the merger announcement date, potentially triggering a treatment effect before the official merger completion date that was used in our analysis. Therefore, we use merger announcement date as the "treatment" date to run a placebo test. We report Hypothesis 1 placebo test results in Table 14 and Figure 5, from which we see that 1) there is no instantaneous effect following merger announcement date ($DID_M$ is not statistically significant); 2) the

parallel trend assumption still holds (the four placebos $DID_M^p$ are not statistically significant). We also report Hypothesis 2 placebo test results in Table 15 and Figure 6, from which we see that none of the coefficients, either leads or lags, is statistically significant, indicating that merger announcement date does not have a long-term effect on OPOR. We also apply the placebo test on Hypothesis 3 to test the differentiating effect of high-performing and low-performing acquirers and all our Hypothesis 3 results stay the same.

Table 14 Using Merger Date as Placebo – Instantaneous Effect

| | Estimate | Standard Error |
|---|---|---|
| $DID_M$ | -0.013 | 0.033 |
| $DID_M^{pl}$ | 0.007 | 0.022 |
| $DID_M^{pl,2}$ | 0.045 | 0.082 |
| $DID_M^{pl,3}$ | -0.033 | 0.079 |
| $DID_M^{pl,4}$ | 0.021 | 0.040 |
| *(Covariates included in model but omitted in reporting)* | | |

Note: This table reports the estimates for both the instantaneous effect $DID_M$ and the four placebos.
*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

Figure 5 Using Merger Date as Placebo to Test the Instantaneous Effect

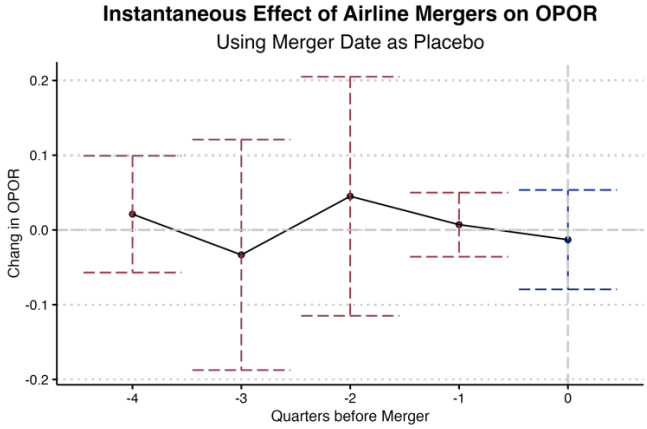



Figure 6 Using Merger Date as Placebo to Test the Long-term Effect

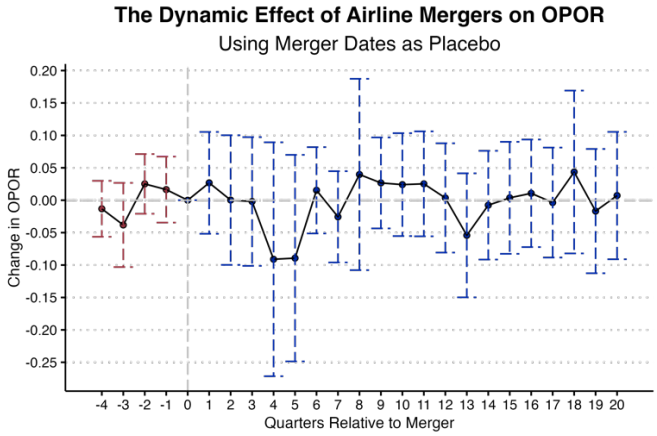


Table 15 Using Merger Date as Placebo – Long Term Effect of Mergers

| $l$ Wave relative to mergers | $\widehat{\mu_l}$ | Standard Error |
|---|---|---|
| -4 | -0.013 | 0.021 |
| -3 | -0.038 | 0.031 |
| -2 | 0.025 | 0.022 |
| -1 | 0.016 | 0.024 |
| 1 | 0.027 | 0.038 |
| 2 | 0.000 | 0.048 |
| 3 | -0.002 | 0.048 |
| 4 | -0.091 | 0.087 |
| 5 | -0.089 | 0.077 |
| 6 | 0.015 | 0.032 |
| 7 | -0.026 | 0.034 |
| 8 | 0.040 | 0.071 |
| 9 | 0.027 | 0.034 |
| 10 | 0.024 | 0.038 |
| 11 | 0.025 | 0.039 |
| 12 | 0.004 | 0.041 |
| 13 | -0.054 | 0.046 |
| 14 | -0.008 | 0.040 |
| 15 | 0.004 | 0.042 |
| 16 | 0.011 | 0.040 |
| 17 | -0.004 | 0.041 |
| 18 | 0.044 | 0.060 |
| 19 | -0.017 | 0.046 |
| 20 | 0.007 | 0.047 |
| *(Covariates included in model but omitted in reporting)* | | |
| Adjusted $R^2 = 0.515$ | | |

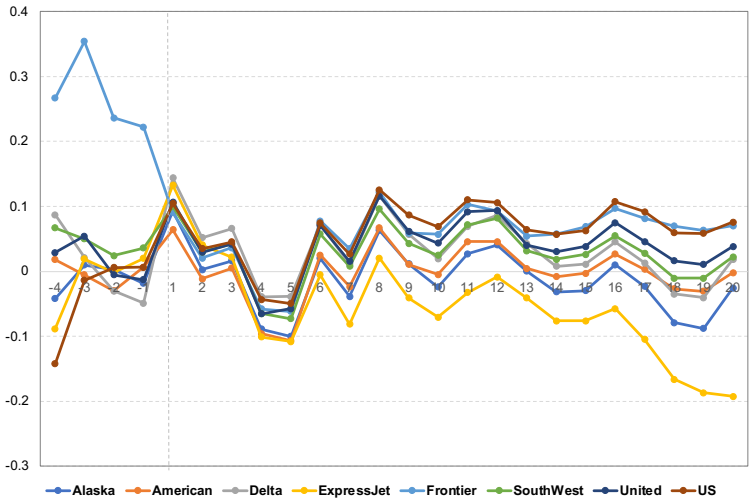
*P value: *** 0.001, ** 0.01, * 0.05, °0.1*

### 6.3 Heterogeneous Effect Among Carriers

In Hypothesis 3, we test the differentiating effect of mergers for high-performing and low-performing acquirers and accordingly find that low-performing acquirers benefit from profitability synergy while high-performing acquirers did not. In this section, we further break down the differentiating effect of mergers at individual acquirer level as a post-hoc analysis, aiming to provide more nuanced findings. We construct separate clean 2×2 stacked datasets (i.e., to avoid already-treated acquirers being used as effective controls) for each individual acquirer then estimate the effect using stacked regression. Figure 7 reports the results (note: the table output is omitted to conserve space). We see that each individual carrier shares roughly the same recovery path post-merger as was tested in Hypothesis 2: in the first two years, the learning-by-doing process leads to a U-shaped performance curve. Carrier's performance then stabilized from the third year onwards. Most of the acquirers share a similar recovery path with the exception of ExpressJet whose performance seems to further deteriorate after the third year. Most acquirers also share a similar pre-trend with the exception of Frontier whose pre-merger performance was much better than its post-merger performance. In sum, breaking down the long term effect into individual carrier level also supports our Hypothesis 2, indicating that acquirers first experienced a U-shaped curve before financial performance eventually starts to stabilize or improve. At individual carrier level, Frontier and ExpressJet did not seem to benefit from profitability synergy. We also tested the instantaneous effect at individual carrier level. With

the exception of Frontier whose performance plunged at about 50% from *t - 1* to *t* right in the quarter of merger (statistically significant), other acquirers did not experience any statistically significant instantaneous changes right upon merger.

Figure 7 Heterogeneous Effect of Merger Among Carriers




## 7. CONTRIBUTION

### Theoretical Contribution

Our study makes several contributions to the OM literature. First, our study extends scientific understanding of the law of diminishing synergy (Schmenner and Swink 1998) in the OM field. Adapted from microeconomic theory, Schmenner and Swink (1998) proposes that the strength of synergies achieved diminishes as firm moves towards its asset frontier (i.e., higher financial performance in our case). Using eight mergers, we were able to distinguish between high-performing and low-performing acquirers to empirically test this law. Our results indicate that acquirers with better than industry-average OPOR do not achieve profitability synergy while acquirers with lower than industry-average OPOR achieves profitability synergy. To this end, our research verifies the law of diminishing synergy in the airline industry. Our study is also among the very few studies to empirically test the law of diminishing synergy in the OM field.

Second, we contribute to further the understanding of merger impact in the airline industry. Among the 31 studies investigating U.S. airline mergers (Appendix A), 19 studies investigate how merger can help increase revenue synergy through increased fare and expanded market power. However, revenue synergy does not necessarily translate into profitability synergy as operating cost is another key factor in the equation (Jordon 1988). In addition, studying two airline mergers (Delta-Northwest and United-Continental), Schosser and Wittmer (2015) concluded that "North American airline mergers show only little evidence for revenue synergies" (p. 148). With the increasing attention focused on profitability in the airline industry (Ziady 2023), we instead examine profitability synergy of mergers. In contrast to the findings from the 19 studies confirming increased fare (2.5% – 23% increase) within the time period of one quarter to two years

post-merger, our result show that from a profitability perspective, acquirers experienced a U-shaped performance curve within the first two years following mergers. Then, profitability synergy starts to stabilize/improve from the third year following mergers. In addition, only low-performing acquirers benefit from mergers in terms of profitability synergy.

Third, we contribute to expand the understanding of the different effects of mergers in the airline industry. Treatment effects can be classified into three different categories: instantaneous effect, average effect, and dynamic/long-term effect. Most of the 31 studies of U.S. airline mergers examined average effect, such as how on-time performance has changed in the first two years following mergers. However, average effect, due to its aggregated nature, obscures the understanding of what happens immediately following merger (instantaneous effect) and how the effect of mergers evolves over time (long-term effect). Our study investigates these two overlooked merger effects and our findings indicate that merger does not have an immediate effect on acquirers but merger does demonstrate a long-term effect, especially for low-performing acquirers.

Fourth, we also contribute to enrich the merger literature across different disciplines. King et al. (2004) conducted a meta-analysis of 93 merger studies and concluded that the impact of merger on acquirer's financial performance is mixed, i.e., some acquirers experience significant and positive returns while others not. King et al. (2004) argued that one of the reasons for the mixed financial performance is due to the event time window, i.e., researchers have largely adopted a short event window to study merger effect while the actual impact of mergers may be more prolonged. Our study serves as a perfect example for this. Looking at Table 8 and Figure 1, if the event window was set as 4 quarters post-merger, the conclusion will be that the impact of merger is negative and statistically significant; if the event window was set as 8 quarters post-merger, the conclusion will be that the effect is a U-shaped curve. By setting the event window as 20 quarters post-merger, we are able to see that the effect stabilizes (Figure 1)/improves (Figure 2) from the third year following the U-curve in the first two years. To this end, we call for researchers to use a longer event window, if allowed, to capture the potential prolonged effect of policy impact in OM research.

Lastly, our research contributes to advancing research validity of DiD research in OM field, especially for staggered DiD design where TWFE produces biased estimates for both average treatment effect and event study effects (Callaway and Sant'Anna 2021, Sun and Abraham 2021). In a staggered DiD design, we call for researchers to adopt the latest estimators and alternatives to estimate treatment effects to increase research rigor. In addition, using an event study to test for the parallel trend assumption has a long track record in OM research, which is also problematic as in a staggered design, using an event study to test the parallel trend assumption can result in both Type I and Type II errors (see section 3.2.1 in Baker et al. 2022). Therefore, we also call for researchers to use alternative estimators to test the parallel trend assumption as this assumption is paramount to the validity of DiD designs.

**Managerial Contribution**

Our study also contributes to policymakers in the airline industry to consider various aspects of airline mergers. Despite the struggles and diminished prospects in the airline industry during the Covid-19 Pandemic, "interest of mergers shows no sign of flagging" (Louge 2021). A recent example is that since 2022, Spirit airline has been actively seeking mergers which could "kick off a new wave of carrier consolidation" (Primack 2022). Spirit first negotiated with Frontier (LeBeau and Josephs 2022) then abandoned the proposal and worked with JetBlue (Tyko 2022). For JetBlue to proceed with the proposal, JetBlue will have to terminate its alliance with American airline following the Department of Justice's orders (Kelleher 2023). In the wake of potential mergers post-Pandemic, our study helps to answer the following three practical questions for policymakers engaging in airline mergers.

First, who should seek a merger? Based on the eight recent U.S. airline mergers, our results indicate that not every acquirer benefit from mergers in terms of profitability. Only low performers benefit from mergers. If a carrier enjoys better than industry-average financial performance, then most likely this carrier will not benefit from a pure financial perspective from mergers. Take the most recent Spirit merger proposal as an example. Both Frontier and JetBlue are potential acquirers. Looking at their respective average financial performance (OPOR) from 2015 to 2019, we see that both Frontier (0.128 VS 0.109 industry average) and JetBlue (0.119 VS 0.109 industry average) can be classified as high performers. Based on our results from Section 5.3, we can predict that neither Frontier nor JetBlue will likely benefit financially from this merger, especially given Frontier already suffered from a worsened financial performance in its previous merger with Midwest in 2010 (Figure 7).

Second, when does profitability synergy start to take effect? Instead of an immediate boost of revenue synergy (Carey 2005), profit does not increase immediately following mergers. Acquirers wishing for an immediate boost in profit should take caution before engaging in mergers. In addition, unlike revenue synergy which is expected to witness a steady increase through increased fare in the first two years post-merger (Appendix A), profitability synergy will first experience a U-shaped curve in the first two years following mergers before eventually starting to improve from the third year. Our analysis show that all the eight U.S. airline mergers share the same U-shaped recovery pattern (Figure 7). Therefore, future policymakers engaging in airline mergers can draw from our research result and justify the potentially deteriorated financial performance in the first two years, which eventually will improve approximately from the third year. In other words, investors and shareholders need not to panic if the post-merger performance plunged in the first two years.

Third, how can acquirers benefit more from profitability synergy? We see from our analysis that the hurdle on profitability synergy is the U-shaped performance curve. We attribute the U-shaped curve to the complicated integration of "thousands of procedures used by pilots and flight dispatchers, gate agents, flight

attendants and ground crew" (Mouawad 2012, p. 3), which may lead to duplicated operating expenses and higher costs (Jordan 1988, Schosser and Wittmer 2015). Therefore, we encourage policymakers, when designing the post-merger integration strategy, focus on mobilizing all possible resources to reduce costs in the first four quarters following mergers as financial performance sharply decreased in the first four quarters before climbing up from the 6$^{th}$ quarter. In other words, the first year post-merger is crucial to improve profitability synergy. We highly recommend policymakers, operations managers, and all related parties to design all plans possible to speed up the process of reducing duplicated operating expenses in the first year post-merger, which, if achieved, can help to improve the overall profitability synergy.

## 8. FUTURE RESEARCH AND CONCLUSION

Like all research, ours has limitations as well as associated future research avenues. First, we rely on organizational learning framework to theorize our hypothesized relationships. This framework involves three recursive stages from knowledge search to knowledge creation/transfer, and to knowledge retention (Argote and Miron-Spektor 2011). Due to data limitations, however, we could not directly test the knowledge transfer and knowledge retention activities. Scholars have developed variable constructs and models to investigate knowledge transfer between organizations (Reus et al. 2016). Future research can build on the knowledge transfer literature to investigate knowledge transfer and retention activities during mergers to examine the impact of knowledge transfer/retention on merger outcomes.

Second, to test the hypothesized differentiating effect of mergers, we construct high-performing and low-performing acquirers by comparing the averaged OPOR and with that of the whole industry during the pre-merger period. Such a comparison is a cross-sectional comparison which loses its ability to further examine how longitudinal changes in financial performance impact merger outcomes. We encourage future researchers to leverage the nature of longitudinal data to investigate this relationship.

Third, for the differentiating effect, we find support that acquirer's pre-merger financial performance impacts acquirer's post-merger performance. However, this does not necessarily mean that pre-merger financial performance is the main, or the only, factor that might impact merger outcomes. There are other acquirer-specific characteristics worth investigating that could have significant influences on merger outcomes, such as method of payment, firm characteristics, and environmental factors (King et al. 2004, Haleblian et al. 2009). Future research can explore these acquirer idiosyncratic characteristics to investigate their impact on mergers.

Finally, this study is a single industry study in the U.S. airline industry. Single industry study has its advantages in that it provides researchers with deeper understanding of the industry and accordingly allows researchers to directly compare performance differences between firms where the determinants of superior performance can be precisely identified (Garvin 1988). However, a single industry study also has its limitations in that its ability to test contextual factors may be hindered (Hale et al. 2016). Thus, we also call

for future study to examine other similar research questions utilizing data from various industries where the impact of different contextual factors on merger outcomes can also be modeled.

In sum, by investigating eight recent mergers in the U.S. airline industry using the latest advancement in econometrics, we demonstrate both a U-shaped performance curve and the differentiating effects of airline mergers. We call for researchers to consider using the latest estimators to increase research validity in drawing treatment effect in a staggered DiD design, or at least use these new estimators as a robustness if researchers continue to report traditional TWFE DiD results. In addition, the track-record practice of using event study to test the parallel trend assumption in a staggered DiD design is also problematic (Baker et al. 2022). We, therefore, also call for researchers to adopt alternatives to test this assumption to increase research validity.

**REFERENCES**


Adrangi, B., Chow, G. & Raffiee, K. 1997. Airline deregulation, safety, and profitability in the US. *Transportation Journal*, *36*(4), 44-52.

Alan, Y., & Lapré, M.A. 2018. Investigating operational predictors of future financial distress in the US airline industry. *Production and Operations Management*, *27*(4), 734-755.

Alexander, K. 2006. Southwest Pushes Fares Higher, Blaming Fuel Costs. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2006/07/05/AR2006070501721.html. (Accessed on July 30, 2023).

Argote, L., & Miron-Spektor, E. 2011. Organizational learning: From experience to knowledge. *Organization Science, 22*(5), 1123–1137.

Azan, W., & Huber Sutter, I. 2010. Knowledge transfer in post–merger integration management: Case study of a multinational healthcare company in Greece. *Knowledge Management Research & Practice, 8*(4), 307–321.

Baker, A.C., Larcker, D.F. & Wang, C.C. 2022. How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, *144*(2), 370-395.

Belobaba, P. 2009. Overview of airline economics, markets and demand. *The Global Airline Industry*, 47-72.

Bliese P.D., & Lang J.W. 2016. Understanding relative and absolute change in discontinuous growth models: Coding alternatives and implications for hypothesis testing. *Organational Research Methods*, *19*(4), 562–592.

Callaway B., Sant'Anna P.H. (2021) Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200-230.

Carey, S. 2005. 'Revenue Synergy' is key for US Air and America West. Retrieved from https://www.wsj.com/articles/SB111680696611140277. Accessed on July 30, 2023.

Carlton, D., Israel, M., MacSwain, I., & Orlov, E. 2019. Are legacy airline mergers pro–or anti–competitive? Evidence from recent US airline mergers. *International Journal of Industrial Organization, 62*, 58–95.

Cengiz D., Dube A., Lindner A., & Zipperer B. 2019. The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics, 134*(3), 1405-1454.

Chen, Y., & Gayle, P.G. 2019. Mergers and product quality: Evidence from the airline industry. *International Journal of Industrial Organization, 62,* 96–135.

Clarke, M.D.D. 1998. Irregular airline operations: A review of the state–of–the–practice in airline operations control centers. *Journal of Air Transport Management, 4*(2), 67–76.

Cook, A., Tanner, G. and Lawes, A. 2012. The hidden cost of airline unpunctuality. *Journal of Transport Economics and Policy*, *46*(2), 157-173.

Crossan, M.M., Lane, H.W. & White, R.E. 1999. An organizational learning framework: From intuition to institution. *Academy of Management Review, 24*(3), 522–537.

Cui R., Ding H., Zhu F. 2022. Gender inequality in research productivity during the COVID-19 pandemic. *Manufacturing and Service Operations Management, 24*(2), 707-726.

Cui R., Zhang D.J., & Bassamboo A. 2019. Learning from inventory availability information: Evidence from field experiments on Amazon. *Management Science, 65*(3), 1216-1235.

Cyert, R.M. & March, J.G. 1963. A behavioral theory of the firm. *Englewood Cliffs, NJ, 2*(4), 169–187.

Dana Jr, J.D., & Orlov, E. 2014. Internet penetration and capacity utilization in the US airline industry. *American Economic Journal: Microeconomics*, *6*(4), 106-37.

de Chaisemartin C., & D'Haultfoeuille X. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review, 110*(9), 2964-96.

Degani, A. & Wiener, E.L 1998. Design and operational aspects of flight–deck procedures. International Air Transport Association (IATA): Human Factors.

Delta Airlines. 2008. Delta and Northwest merge, creating premier global airline. https://news.delta.com/delta–and–northwest–merge–creating–premier–global–airline. Accessed on July 30, 2023.

Department of Transportation. 2023. Data directory: list of databases. https://www.transtats.bts.gov/DataIndex.asp. Accessed on July 30, 2023.

Deshpande, M., & Li, Y. 2019. Who is screened out? Application costs and the targeting of disability programs. *American Economic Journal: Economic Policy, 11*(4), 213-248.

Dong, Y., Chung, M., Zhou, C., & Venkataraman, S. 2019. Banking on "Mobile Money": The Implications of Mobile Money Services on the Value Chain. *Manufacturing and Service Operations Management, 21*(2), 290-307.

Dresner, M., & Xu, K. 1995. Customer service, customer satisfaction, and corporate performance. *Journal of Business logistics*, *16*(1), 23-40.

Federal Aviation Administration, January 2020. The Economic Impact Report of Civil Aviation on the U.S. Economy. https://www.faa.gov/sites/faa.gov/files/about/plans_reports/ (Accessed on July 30, 2023).

Fiol, C., & Lyles, M. 1985. Organizational learning. *Academy of Management Review, 10*(4), 803–813.

Garvin, D. 1988. Managing quality: The strategic and competitive edge. Simon and Schuster.

Gittell, J. H. 2003. The Southwest airlines Way: Using the power of relationships of high performance. New York: McGraw-Hill.

Gong, S., & Firth, M. 2006. The competitive effects of airline mergers and acquisitions: More capital market evidence. *Journal of Air Transportation, 11*(1), 88–105.

Goodman-Bacon, A. 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*(2):254-277.

Gudmundsson, S., Merkert, R. & Redondi, R. (2017). Cost functions and determinants of unit cost effects in horizontal airline M&As. *Transportation Research Part A: Policy and Practice*, *103*, 444–454.

Hale Jr, D., Ployhart, R. & Shepherd, W. 2016. A two–phase longitudinal model of a turnover event: Disruption, recovery rates, and moderators of collective performance. *Academy of Management Journal, 59*(3), 906–929.

Haleblian, J., Devers, C., McNamara, G., Carpenter, M.A. & Davison, R.B. 2009. Taking stock of what we know about mergers and acquisitions: A review and research agenda. *Journal of Management, 35*(3), 469–502.

Holmes, F. 2023. Lower Airfares, Higher Profits? The Hidden Connections In The Airline Industry. https://www.forbes.com/sites/greatspeculations/2023/07/17/lower-airfares-higher-profits-the-hidden-connections-in-the-airline-industry/?sh=28ffe41444c6. Accesses on July 30, 2023.

Homberg, F., Rost, K., & Osterloh, M. 2009. Do synergies exist in related acquisitions? A meta-analysis of acquisition studies. *Review of Managerial Science, 3*(2), 75–116.

Hüschelrath, K. & Müller, K. 2014. Airline networks, mergers, and consumer welfare. *Journal of Transport Economics and Policy, 48*(3), 385–407.

Hüschelrath, K. & Müller, K. 2015. Market power, efficiencies, and entry evidence from an airline merger. *Managerial and Decision Economics, 36*(4), 239–255.

Huselid, M.A. 1995. The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal, 38*(3), 635–672.

IATA. 2018. Fuel Fact Sheet. Retrieved from https://www.iata.org/contentassets/ebdba50e57194019930d72722413edd4/fact-sheet-fuel.pdf. (Accessed on July 30 2023).

Jain, V. 2015. What did the wave bring?: Short–term price effect of the us airline merger wave (2009–2012). *Journal of Economic Policy and Research, 10*(2), p.49–68.

Jordan, W.A. 1988. Problems stemming from airline mergers and acquisitions. *Transportation Journal, 27*(4), 9–30.

Josephs, L. 2018. Years after airline mergers, flight attendants are finally flying together. Retrieved from https://www.cnbc.com/2018/10/01/years–after–airline–mergers–flight–attendants–start–to–fly–together.html. Accessed on July 30, 2023.

Kalemba, N., & Campa-Planas, F. 2017. An overview of the quality concept in the air transportation business: a systematic literature review. *International Journal for Quality Research*, *11*(1), 51-70.

Katzenbach, J.R., & Smith, D.K. 2015. The wisdom of teams: Creating the high–performance organization. Harvard Business Review Press.

Kelleher, S.R. 2023. JetBlue Will Dump Alliance With American To Save Spirit Deal. https://www.forbes.com/sites/suzannerowankelleher/2023/07/06/jetblue-dump-alliance-american-spirit-merger/?sh=2d1b0d7d2188. Accessed on July 30, 2023.

Kim, E.H., & Singal, V. 1993. Mergers and market power: Evidence from the airline industry. *The American Economic Review, 83*(3), 549–569.

King, D.R., Dalton, D.R., Daily, C.M. & Covin, J.G. 2004. Meta-analyses of post-acquisition performance: Indications of unidentified moderators. *Strategic Management Journal, 25*(2), 187–200.

Lampel, J., Shamsie, J. & Shapira, Z. 2009. Experiencing the improbable: Rare events and organizational learning. *Organization Science, 20*(5), 835–845.

Lang, J.W. & Bliese, P.D., (2009). General mental ability and two types of adaptation to unforeseen change: Applying discontinuous growth models to the task–change paradigm. *Journal of Applied Psychology, 94*(2), 411–428.

LeBeau, P., & Josephs, L. 2022. Frontier and Spirit to merge, creating fifth-largest airline in U.S. in $6.6 billion deal. https://www.cnbc.com/2022/02/07/frontier-and-spirit-to-merge-creating-5th-largest-airline-in-us.html. Accessed on July 30, 2023.

Leibenstein, H. 1966. Allocative efficiency vs. X-efficiency. *The American Economic Review*, *56*(3), 392–415.

Levitt, B. & March, J.G. 1988. Organizational learning. *Annual Review of Sociology, 14*(1), 319–338.

Levy, M. 2011. Knowledge retention: minimizing organizational business loss. *Journal of Knowledge Management*, *15*(4), 582-600.

Li, J., Wu, D. 2020. Do corporate social responsibility engagements lead to real environmental, social, and governance impact? *Management Science, 66*(6):2564-2588.

Lichtenberg, F.R., & Kim, M. 1989. The effects of mergers on prices, costs, and capacity utilization in the US air transportation industry, 1970–84 (No. w3197). National Bureau of Economic Research.

Louge, F. 2021. How the pandemic is changing the outlook for airline mergers. https://www.frontier-economics.com/uk/en/news-and-articles/articles/article-i8918-how-the-pandemic-is-changing-the-outlook-for-airline-mergers/. Accessed on July 30, 2023.

Luo, D. 2014. The price effects of the Delta/Northwest airline merger. *Review of Industrial Organization, 44*(1), 27–48.

Mantin, B., & Wang, J.H.E. 2012. Determinants of profitability and recovery from system-wide shocks: The case of the airline industry. *Journal of Airline and Airport Management*, 2(1), 1-21.

Manuela Jr, W.S., Rhoades, D.L., & Curtis, T. 2016. The US Airways Group: A post–merger analysis. *Journal of Air Transport Management, 56*, 138–150.

March, J.G., Sproull, L.S., & Tamuz, M. 1991. Learning from samples of one or fewer. Organization science, *2*(1), 1-13.

Marsh, S.J., & Stock, G.N. 2006. Creating dynamic capability: The role of intertemporal integration, knowledge retention, and interpretation. *Journal of Product Innovation Management, 23*(5), 422–436.

Mellat-Parast, M., Golmohammadi, D., McFadden, K.L., & Miller, J.W. 2015. Linking business strategy to service failures and financial performance: Empirical evidence from the US domestic airline industry. *Journal of Operations Management*, 38, 14-24.

Merkert, R. & Morrell, P.S. 2012. Mergers and acquisitions in aviation –Management and economic perspectives on the size of airlines. *Transportation Research Part E: Logistics and Transportation Review, 48*(4), 853–862.

Miller, D. 1996. A preliminary typology of organizational learning: Synthesizing the literature. *Journal of Management, 22*(3), 485–505.

Mishina, Y., Pollock, T.G., & Porac, J.F. 2004. Are more resources always better for growth? Resource stickiness in market and product expansion. *Strategic Management Journal*, 25(12), 1179-1197.

Mouawad, J. 2012. For United, Big Problems at Biggest Airline. https://www.nytimes.com/2012/11/29/business/united-is-struggling-two-years-after-its-merger-with-continental.html. Accessed on July 30, 2023.

Netessine, S., & Shumsky, R. 2002. Introduction to the theory and practice of yield management. *INFORMS Transactions on Education*, *3*(1), 34-44.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37.

Peters, C. 2006. Evaluating the performance of merger simulation: Evidence from the US airline industry. *The Journal of Law and Economics, 49*(2), 627–649.

Primack, D. 2022. Spirit Airlines takeover fight heads to a vote. https://www.axios.com/2022/06/28/spirit-airlines-takeover-fight-heads-to-a-vote. Accesses on July 30, 2023.

Prince, J.T., & Simon, D.H. 2017. The impact of mergers on quality provision: Evidence from the airline industry. *The Journal of Industrial Economics, 65*(2), 336–362.

Ramdas, K., Williams, J., & Lipson, M. 2013. Can financial markets inform operational improvement efforts? Evidence from the airline industry. *Manufacturing and Service Operations Management*, *15*(3), 405-422.

Reus, T.H., Lamont, B.T., & Ellis, K.M. 2016. A darker side of knowledge transfer following international acquisitions. *Strategic Management Journal, 37*(5), 932–944.

Ryerson, M.S., & Kim, H. 2014. The impact of airline mergers and hub reorganization on aviation fuel consumption. *Journal of Cleaner Production, 85*, 395–407.

Schmenner, R.W., & Swink, M.L. 1998. On theory in operations management. *Journal of Operations Management*, *17*(1), 97-113.

Schosser, M., & Wittmer, A. 2015. Cost and revenue synergies in airline mergers–Examining geographical differences. *Journal of Air Transport Management, 47*, 142–153.

Scotti, D., & Dresner, M. 2015. The impact of baggage fees on passenger demand on US air routes. *Transport Policy, 43*, 4-10.

Singal, V. 1996a. Airline mergers and multimarket contact. *Managerial and Decision Economics, 17*(6), 559–574.

Singal, V. 1996b. Airline mergers and competition: an integration of stock and product price effects. *Journal of Business, 69*(2), 233–268.

Sun L., & Abraham, S. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225(*2), 175-199.

Tsikriktsis, N. 2007. The effect of operational performance and focus on profitability: A longitudinal study of the US airline industry. *Manufacturing and Service Operations Management*, 9(4), 506-517.

Tyko, K. 2022. Spirit-JetBlue merger takes flight after Frontier deal terminated. https://www.axios.com/2022/07/28/spirit-airlines-jetblue-merger. Accessed on July 30, 2023.

Wacker, J.G. 1998. A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of operations management*, *16*(4), 361-385.

Wang, C. L., & Ahmed, P. K. 2003. Organisational learning: a critical review. *The learning organization*, *10*(1), 8-17.

Wenner, C.A., & Drury, C.G. 2000. Analyzing human error in aircraft ground damage incidents. *International Journal of Industrial Ergonomics, 26*(2), 177–199.

Wooldridge, J.M. 2010. Econometric analysis of cross section and panel data. MIT press.

Ziady, H. 2023. Airlines expect to make $10 billion this year despite economic slowdown. https://www.cnn.com/2023/06/05/business/iata-airlines-profits-post-pandemic/index.html. Accessed on July 30, 2023.

Zollo, M., & Singh, H. 2004. Deliberate learning in corporate acquisitions: post-acquisition strategies and integration capability in US bank mergers. *Strategic management journal*, 25(13), 1233–1256.

Zollo, M. 2009. Superstitious learning with rare strategic decisions: Theory and evidence from corporate acquisitions. *Organization Science*, 20(5), 894–908.

Zott, C. 2003. Dynamic capabilities and the emergence of intra–industry differential firm performance: Insights from a simulation study. *Strategic Management Journal*, 24(2), 97–125.