BAGGAGE FEES AND AIRLINE PERFORMANCE

ABSTRACT

Leveraging on cognitive appraisal theory, this manuscript investigates how the implementation of the new baggage fee policy impacts carrier's financial performance, on-time performance, and consumer complaints. Utilizing the latest event study methodology from econometrics field, our analysis shows that the effect of this policy is twofold. Financial performance dropped immediately upon the policy implementation and continued to deteriorate before eventually started to improve in the long run. On-time performance improved immediately upon the policy implementation and kept improving, demonstrating an inverted U-shaped curve. Consumer complaint increased immediately upon the policy implementation and continued to increase at a diminishing rate. Our different findings of the baggage fee policy impact extend the current stream of airline literature as well as providing new managerial implications for airline managers.

Keywords: Baggage fees, on-time performance, operating profit, consumer complaint, US airline industry

1. INTRODUCTION

Amid the economic recession and soaring fuel prices of the late 2000s, major US airlines started to charge baggage fees as they struggled with costs associated with baggage operations (McCartney 2008a, 2008c). These major airlines claimed that the additional revenue generated from baggage fees allowed them to alleviate the impact of the economic recession and high fuel costs (McCartney 2010b); in the meantime, airlines also claimed that they were able to improve baggage operations and overall operational performance due to fewer checked bags (McCartney 2008b, 2008c, 2010a, 2010b).

However, evidence regarding the effect of baggage fees is mixed in the literature. With regard to financial performance, both airline annual reports (Department of Transportation 2024) and airline research (Garrow et al. 2012, Schumann and Singh 2014) confirmed that baggage fees increase revenue. Nevertheless, Yazdi et al. (2017) contended that the net effect of baggage fees is unknown, as it both increases revenue directly through extra fees and decreases revenue indirectly as a result of reduced consumer demand. With regard to operational performance, such as on-time arrivals, related literature also presents mixed findings concerning the post-policy impact: there is evidence of a positive impact (Scotti et al. 2016, Nicolae et al. 2017), a negative impact (McCartney 2008a, 2010b, 2012), and a mixed impact, where an initial deterioration was followed by an improvement (Yazdi et al. 2017). With regard to consumer response, similar mixed results appear. While there has been tremendous coverage in the media of consumer outrage around new baggage fees (McCartney 2010b, Tuttle 2014, Elliott 2015), Scott et al. (2016) found no significant relationship between the new baggage fee policy and consumer response, when measured as consumer complaints. These mixed findings draw our attention. In this research, we re-investigate the relationship between baggage fees and three outcome variables: financial performance, operational performance, and consumer complaints.

Different from the extant literature, we do not use revenue to measure financial performance. Instead, we focus on investigating the bottom-line performance – profit. The reason is that airlines are reported to yield modest profits when revenues are record-high (Garcia 2024) as well as record-high profits when revenues are low (Holmes 2023). Investigating the bottom-line performance, therefore, may more accurately represent the true picture of financial performance. When investigating operational performance, all extant airline studies have adopted DOT's method of calculating on-time performance (OTP). However, DOT's OTP calculation includes five categories of delays that have impacted on-time arrivals, such as weather delays, which can account for as high as 45.5% of all delays (Nicolae et al. 2017). Weather delays, nonetheless, should not have any direct interaction with any policy implementation. Therefore, in contrast to extant literature, we only study carrier-induced delays that may best represent any potential policy impact on OTP. Lastly, different from extant airline research that only investigated consumer complaints about mishandled bags, we elect to investigate total number of consumer complaints as we argue that consumer complaints may be triggered by the policy itself due to the potential halo effect following policy implementation (Halstead et al. 1996).

We ground our reasoning on appraisal theory and the relevant literature; then we test our theorizing using the latest development in event study methodology, which allows us to assess the relationship both immediately upon policy implementation and in the long term. We argue that upon the implementation of the baggage fee policy, airlines will see a decrease in financial performance, an improvement in operational performance, and an increase in consumer complaints. Further, we hypothesize different polynomial relationships regarding the long-term impact on the three outcome variables. Our hypotheses testing results support our predictions of the immediate impact on the three outcome variables; we also show that over time, financial performance, operational performance, and consumer complaints all demonstrate different curvilinear relationships.

Our results contribute to airline research in operations management field in the following ways. First, in contrast to extant research showing an improvement in revenues following the baggage feel policy implementation, our results indicate that when financial performance is measured by profit, carriers who charged baggage fees suffered from a decrease in financial performance both instantaneously and in the long term. Our study, therefore, extends the related airline literature. Second, different from related research that shows either a positive or negative impact of baggage fee policy on OTP, we show that when OTP was measured by carrier-induced delays only, OTP improved immediately upon the policy implementation and continued to improve before the effect started to taper off. Our different findings regarding the impact on OTP, therefore, also extend the related stream of literature. Third, we find that consumers started to complain more even before the policy implementation and this trend continued after the policy implementation, corroborating our theoretical reasoning of the halo effect. In lieu of a full-length analysis of consumer behavior and the halo effect, our findings provide useful evidence for future operations research in this domain. Lastly, two-way fixed-effects (TWFE) are generally used to estimate event studies in operations management. However, latest advancement in econometrics show that TWFE produces biased estimates when used to estimate treatment effect in a situation where there are multiple time periods and variation in treatment timing (Callaway and Sant'Anna 2021, Baker et al. 2022). We accordingly adopt the latest recommendation from econometrics field to avoid this bias by constructing a clean control group. To this end, we contribute to increase research rigor when conducting event study research in operations management and accordingly call for operations researchers to apply the latest development from econometrics field on future event studies.

2. HYPOTHESES DEVELOPMENT

In this section, we build our hypotheses by drawing on the findings from previous literature as well as relevant theories and concepts to explain the hypothesized relationships. To study policy impact, we leverage relevant concept from social sciences field that examines policy impact at two different stages: the transition stage and the recovery stage (Lang and Bliese 2009). The transition stage is defined as "the degree to which routines and expertise from the pre-change period are immediately transferred to the changed task" (Lang and Bliese 2009, p. 415). The recovery stage refers to the process of recovering following the immediate impact after a change and reflects "individuals' ability to fundamentally re-evaluate the applicability of already acquired expertise" (Lang and Bliese 2009, p. 415). Accordingly, we develop our hypotheses at two different stages: the immediate impact of baggage fees at the transition stage and the long-term impact of baggage fees at the recovery stage.

2.1 New Baggage Fee Policy and Carrier Financial Performance

The relationship between the baggage fee policy and carriers' financial performance, especially in the form of revenue, has been widely researched in the airline literature, the vast majority of which demonstrates a positive relationship. Garrow et al. (2012) studied the de-bundling of baggage fees among major US airlines from 2007 to 2009 and found that baggage fees contributed directly to the increase of ancillary revenues for both legacy carriers and low-cost carriers (excluding Southwest and JetBlue). In examining the impact of baggage fees on ticket price, Henrickson and Scott (2012) also observed that baggage fees have successfully helped airlines increase their revenues. Using DOT data from 2006 to 2010, Schumann and Singh (2014) also concluded that those carriers who charge baggage fees benefit from extra revenue growth compared with those carriers who choose to adopt a "No Fee" policy. Scotti and Dresner (2015), analyzing DOT data from 2007 to 2010, concluded that "the ancillary fees provide a means to increase revenues" (p. 9). Yazdi et al. (2017), compiling panel data from 2003 to 2014, claimed that baggage fee implementation is a ''floor wax and dessert topping" which "raises revenues for airlines" (p. 96).

Despite these findings regarding revenue increase following baggage fee policy implementation, there is still a need to investigate baggage fee policy impact on financial performance for two reasons. First, revenue increase does not equal to profit increase. On one hand, Delta airlines reported a record-high profit in Q2 2023 despite a record-low revenue (Holmes 2023). On the other hand, despite a record-high revenue, airlines are only reported to yield a modest profit (Garcia 2024). Therefore, investigating the impact of baggage fee policy on profit, which has not been examined yet, might yield different and interesting result. Second, Yazdi et al. (2017) also called for more refined future research to investigate the impact of baggage fees on financial performance, citing that "the net effect of baggage fees, (which increases revenue directly, but also decreases revenues through lower demand) is not known a priori" (p. 96).

To further investigate baggage fee impact on financial performance, we focus on cognitive theory—specifically, the concept of cognitive appraisal to develop our hypothesis (Lazarus 1991, Scherer et al. 2001). According to cognitive theory as it is applied in airline literature, "negative emotions arise from the cognitive appraisal of ancillary fees that lead to coping behaviors" (Tuzovic et al. 2014, p. 99). When a baggage fee policy is first implemented, we predict that consumers will develop cognitive appraisal of the new policy. Given checked baggage used to be a free service, consumers may develop negative emotions toward the new policy (Lazarus 1991). Related literature has observed that consumers have strong negative emotions of betrayal regarding baggage fees (Tuzovic et al. 2014). These emotions of betrayal may trigger consumers to develop coping behaviors, such as avoiding travelling by air. The avoidance of air travel can be explained from two perspectives. First, consumers may perceive baggage fees as unfair (Yazdi et al. 2017). When consumers perceive prices as unfair, they tend to avoid purchases (Xia et al. 2004). Second, the avoidance of air travel is exacerbated in the north-eastern US, where alternatives, such as rail travel, are readily available (Morrison and Winston 2005). Moreover, Scotti and Dresner (2015) confirmed that a one dollar increase in baggage fees leads to a loss of 0.7 passengers. Loss of passengers may translate to loss in financial performance. Combining these arguments leads us to postulate that airlines are most likely to suffer from a loss in financial performance immediately upon policy implementation as a result of consumer's coping behaviors.

H1a: Financial performance will decrease immediately in the transition stage of implementing baggage fees.

To assess the long-term impact, we continue using cognitive appraisal theory to explain the expected relationships. As discussed, immediately upon policy implementation, consumers will develop coping behaviors based on their appraisals of the policy (Lazarus and Folkman 1984, Lazarus 1991, Scherer et al. 2001). This is what we designate as the initial stage appraisal. Refusal to purchase is one of the expected results in the initial stage appraisal. However, as more airlines gradually implement baggage fee policies, the progressive nature of appraisals (Lazarus 1991) predicts that "continuous appraisal of the significance of the event sometimes results in an interruption of ongoing plans and actions" (Trabasso and Stein 1993, p. 328). Therefore, it is expected that, as consumers keep evaluating the policy, they will realize that refusal to purchase is futile and may affect their own interests, such as missing important client meetings. Consumers will adjust their actions by rationalizing their travelling utilities (Ben-Akiva and Lerman 1985, Suzuki 2004). In other words, consumers will make travelling choices that give them the most benefit. This is what we designate as the second stage appraisal. To rationalize their trip utilities (Suzuki 2000) and coping behaviors at this stage, consumers are more likely to start (i) travelling by air and (ii) paying for baggage fees. Such processes are not expected to happen overnight, given the continuous nature of appraisals. Consequently, as these gradual appraisals result in consumers starting to repurchase air tickets and pay for baggage fees, we predict that—after the initial financial performance plunge and continuous decrease—airlines will see a gradual improvement in the long run.

H1b: Financial performance will demonstrate a U-shaped curve in the recovery stage.

2.2 New Baggage Fee Policy and On-Time Performance

Among the various service dimensions in air travelling, on-time performance (OTP) is one of the service characteristics most valued by passengers (Mitra 2001) because OTP is "a strong indicator of whether [a] trip will encounter delays and disruptions" for travellers (McCartney 2010a). A flight is considered "on-time" if it operated within 15 minutes of the scheduled operating time shown in a carrier's Computerized Reservations Systems (DOT 2024). OTP is a frequently researched construct in airline research, though with varying measures. One stream of research strictly adopts DOT's definition (Suzuki 2000, Steven et al. 2012, Peterson et al. 2013, Kalemba and Campa-Planas 2017), while another stream of research measures arrival delay (Forbes 2008, Prince and Simon 2009, Cook et al. 2012, Ramdas et al. 2013, Mellat-Parast et al. 2015, Yimga 2017).

After carefully studying DOT's method to calculate OTP, we elect to only use carrier-induced delays to proxy on-time performance. The reason is that OTP was calculated as total number of flights minus arrival delayed flights. However, arrival delayed flights include flights that were delayed due to five reasons: carrier delay + weather delay + national aviation system (NAS) delay + security delay + late arriving aircraft delay. Implementing baggage fee or not, OTP should have nothing to do with weather delay or NAS delay, especially given weather delay accounts for 42% of all delays in our data. Nicolae et al. (2017) also reported that weather delay accounts for 45.5% of all delays in their data. Including the 40%+ weather delay to assess baggage fee policy impact on OTP would most likely yield biased causal inferences. Therefore, to isolate baggage fee policy impact on OTP from other irrelevant impact on OTP caused by weather delay and other delays, we only use carrier delay to proxy OTP, hoping to provide different and new findings.

The relationship between baggage fee policies and OTP has been investigated in airline literature with mixed findings. One stream of literature observes that OTP deteriorates after policy implementation, while other scholars find that OTP actually improves following baggage fee implementation. Still other research claims that OTP initially deteriorates before it improves. We briefly review these three streams of literature in the following paragraphs.

In industry reports, McCartney (2008a, 2010a, 2012) contends that OTP suffers due to a baggage fee policy implementation for several reasons. First, after the policy implementation, more consumers carry on their luggage rather than checking it for a fee, resulting in an increased total number of carry-ons. Consequently, consumers fight for cabin storage space in the "boarding stampede" (McCartney 2012), which prolongs boarding time, delays flight departures, and subsequently impacts on-time arrivals. Second, consumers not only avoid checked bags, but they also begin to pack more into their carry-ons. As the cabin fills up or a consumer's carry-on is too bulky to fit in the cabin, flight attendants often "find themselves taking bags off planes" to check them (McCartney 2008a). This further prolongs the boarding process and potentially affects on-time arrivals. Third, when airlines realize that "bin battles" (McCartney 2008a) can delay flights, they station airline workers to screen bags at boarding gates. These workers identify bulky bags that might not fit in the cabin, consuming human resources and adding five to six additional minutes to the boarding process (McCartney 2012). This contributes further to worsened OTP. In airline research, Jiang and Zheng (2020) also found both checked bags and carry-ons will cause airport congestion and subsequently prolong boarding time, ground handling time, and eventually impact OTP.

The second stream of literature uncovers the opposite findings, revealing that OTP actually improved following the policy implementation. Nicolae et al. (2017) investigated the impact of the new baggage fee policy on departure delays at route-flight level. They found that airlines who charge baggage fees witnessed a significant improvement in their departure performance following policy implementation, despite the increased number of carry-on bags, which might be expected to have a detrimental effect on departures. Nicolae et al. (2017) claimed that this is because the below cabin effect (ground handling of checked bags) outweighs the above cabin effect (cabin handling of carry-ons). Nicolae et al.'s (2017) findings also indicate that the improved departures contribute to better on-time arrivals at destinations. Scotti et al. (2016), compiling panel data from 2004 to 2012, investigated the relationship between baggage fee policies and on-time arrivals at the carrier level and found that increases in baggage fees lead to an improvement in on-time arrivals. Combined, Nicolae et al. (2017) and Scotti et al. (2016) found that charging baggage fees is associated with improvement in airline OTP.

Other findings also exist regarding the relationship between baggage fee policies and OTP. Yazdi et al. (2017) examined the impact of the new baggage fee policy on gate arrivals at carrier-route level. Yazdi et al. (2017) found that up to 2014Q4, on-time arrival performance initially deteriorated, then subsequently improved. Yazdi et al. (2017) further explained that in the initial stage of baggage fee implementation, the below cabin effect, which resulted from a reduced number of checked bags and which resulted in shorter ground handling time, was not significant enough to offset the negative above cabin effect brought by an increased number of carry-ons, leading to boarding delays. As a result, the initial OTP deteriorated. But as time elapsed, "the reduction in checked-bags may have been large enough to lead to improvements in airport-side operations" (Yazdi et al. 2017, p. 94). Yazdi et al.'s (2017) research suggests that a gradual improvement in OTP follows the initial deterioration in OTP.

Our first hypothesis examines the immediate impact of the policy, building on the following three steps of reasoning. First, McCartney (2008a) reported that major airlines had been struggling with their baggage operations due to the overwhelming quantity of checked bags before the implementation of the new baggage fee policy. For example, American Airlines mishandled one bag for every 141 passengers in the first four months of 2008 (McCartney 2008a). Exacerbating the issue, American Airlines was ranked worst among all US airlines in terms of "on-time dependability" (McCartney 2008a). Taken together, these suggest the struggle to efficiently handle checked-in baggage may be a factor of poor OTP. Second, appraisal theory posits that consumers are likely to accumulate negative emotions toward baggage fee policies and accordingly develop coping behaviors, such as carrying on luggage to avoid baggage fees. Both industry reports and airline research observe that passengers began to "dodge" baggage fees by carrying on their luggage after policy implementation (Higgens 2010, McCartney 2008a; 2010a, Cho and Dresner 2018). As a result, a significant decline in checked baggage is expected after policy implementation (Cho and Dresner 2018). Third, given airlines had been struggling with handling too many bags and given our theoretical prediction that the number of checked bags will decline immediately after policy implementation, it logically follows that airlines' ground handling operations will be relieved from the influx of bags after policy implementation. Under this situation, the same number of ground handling staff should yield a higher efficiency due to the reduced number of bags. This should result in faster ground handing time and contribute to OTP improvement.

H2a: OTP will improve immediately in the transition stage of implementing baggage fees.

To build our arguments regarding the long-term effect of the new baggage fee policy on OTP, we continue to apply appraisal theory. First, as discussed, upon policy implementation, consumers will initially attempt to avoid the fees (Lazarus and Folkman 1984, Lazarus 1991, Scherer et al. 2001). However, as most airlines gradually adopt similar policies, consumers will realize that baggage fees are standard airline practice during their second stage appraisal. As a result, consumers are likely to change their behaviors to maximize their utilities for each trip (Suzuki 2000, 2004) by checking in luggage instead of struggling with the unpleasant boarding stampede (McCartney 2012). Consequently, we expect the number of checked bags to increase with time. Second, if the number of bags gradually increases, while the number of ground handling staff stays fixed in the near term (Bruno et al. 2019, Zeng et al. 2019), then the ground operations will have to handle more bags. As a result, the pace of the improvement of OTP may slow down and eventually begin to decline when ground handling staff are unable to keep up with the increase in checked bags. Therefore:

H2b: OTP will demonstrate an inverted U-shaped curve in the recovery stage.

2.3 New Baggage Fee Policy and Consumer Complaints

Although not widely researched, the relationship between a baggage fee policy and consumer complaints also demonstrates mixed findings in the literature. Scotti et al. (2016) collected data from DOT for the period of 2004 to 2012 in order to investigate how charging baggage fees impacted consumer complaints. Using the number of complaints about baggage-related issues as the outcome variable and the fee charged as the predictor, Scotti et al. (2016) were unable to find any significant relationship. Tuzovic et al. (2014) used survey data to conduct related research examining consumer response to airline price de-bundling. They found that consumers felt the strongest sense of betrayal about the baggage fee de-bundling, which had a direct impact on consumer complaints.

As with previous hypotheses, we delineate the relationship between baggage fees and consumer complaints into an immediate impact upon policy implementation and a long-term impact over time. To explore the immediate impact upon policy implementation, we leverage appraisal theory as well as the halo effect of consumer complaint behavior (Halstead et al. 1996) to explain the expected relationship. Appraisal theory suggests that once baggage fees have been de-bundled from an airline's ticket price, consumers will start evaluating the service attributes associated with baggage fees, because they have to pay extra. The appraisal process will most likely result in the following three outcomes. First, consumers will assess the new baggage fees as unfair, given that the airlines charged the fees without adding any additional value to the existing baggage service (Tuzovic et al. 2014, Yazdi et al. 2017). The perceived unfairness will then engender complaints against the service associated with fees that were appraised as unfair (Zaltman et al. 1978, Tuzovic et al. 2014). Second, when airlines start to charge extra for baggage, consumers will accordingly develop expectations of receiving higher quality in baggage-related service (Forbes 2008), especially for those consumers who stick with their carriers and pay extra for checked bags (Nicolae et al. 2016). However, given that airline ground operating procedures are normally standardized (Bazargan 2016), even with the extra fees gleaned from consumers, airlines are unlikely to redesign their operating procedures to proactively improve their service quality. The service quality associated with baggage service, therefore, is not likely to improve coincident with the increase of baggage fees. Thus, the expectation of higher service quality will not be met. Forbes (2008) found that consumer complaints in the airline industry were driven by the gap between expectations and experienced quality levels. As a result, consumers might complain more on baggage-related issues when "they would have expected to receive higher quality" (Forbes 2008, p. 191). Third, complaining behavior will increase when the social climate (i.e., a social environment shared by a group of people) is favorable for complaining (Landon 1977, Halstead et al. 1996). This is especially true with baggage fee policy implementation, which has caused millions of consumers to develop a collective negative feeling (McCartney 2008a). This collective negative feeling will in turn trigger consumers to complain more, as predicted by appraisal theory (Lazarus 1991, Scherer et al. 2001). Halstead et al.'s (1996) halo effect of consumer complaint behavior further predicts that "complaints may beget more complaints" (p. 109). If we apply this halo effect to complaints on baggage-fee related issues, we expect that such complaints may also trigger complaints on other airline service-related issues.

Moreover, we expect that the implementation of a baggage fee policy might impact not only the ground handling of baggage, which is reflected in consumer complaints regarding baggage-related issues, but also the whole travel experience. If the implementation of a baggage fee policy impacts security check in, gate check in, boarding, and deplaning, it is also likely to impact consumer complaints in other service-related issues. Keeping these potential outcomes in mind, we expect that the total number of consumer complaints will increase immediately after the implementation of a baggage fee policy.

H3a: Consumer complaints will increase immediately in the transition stage of implementing baggage fees.

We also investigate the long-term impact of the baggage fee policy on consumer complaints by referring to the previously discussed two-stage appraisals. We argue that the initial stage of appraisal happens immediately after policy implementation, when consumers are triggered to complain more. The perceived feelings of unfairness (Yazidi et al. 2017) and betrayal due to baggage fees (Tuzovic et al. 2014) will motivate consumers to keep complaining. However, as more and more airlines implement a baggage fee policy, it is expected that consumers will realize that baggage fees have become a standard part of airline pricing mechanisms. Accordingly, consumers will gradually develop the second stage appraisal of baggage fees and rationalize their travelling utilities, accepting the fee. Thus, we posit that the positive complaint trend against baggage-related issues will continue, but its rate will gradually diminish as consumers adapt. We further argue that the collective negative emotions, developed in the initial stage of appraisal immediately after policy implementation, will also diminish gradually in line with consumers' gradual shift to a second stage of appraisal. As a result, the social climate and halo effect that engendered more complaints in other service areas will also be ameliorated, leading to fewer complaints regarding other service-related issues.

H3b: Consumer complaints will continue to increase at a diminishing rate in the recovery stage.

3. DATA

To test our hypotheses, we collected data from the U.S. Department of Transportation (DOT). Airlines with at least 1% of total domestic scheduled service passenger revenues are required to report their financial and operational performance to DOT. Other carriers may also report their financial and operational performance voluntarily to DOT. At the time of accessing the DOT database, the relevant data were available from

1998Q1. To reflect DOT's data format change in October 2003, as well as to avoid the external shock of 9/11 and Covid-19, we choose 2004Q1 as our data starting point and 2019Q4 as our data ending point, resulting in 16 years of data. Dependent variables and control variables were drawn from different data sources and were handled differently, depending on their data formats. This will be discussed in greater detail in this section.

3.1 Airlines

A three-step data cleaning process was conducted to finalize the airline list. First, we tracked the name changes of airlines through the years and kept the most recent brand names to identify the unique airlines in our dataset. These airlines include Envoy (American Eagle until March 2003) and Endeavor (Pinnacle until December 2012). Other airlines, such as United Airlines and United Express, who simultaneously promoted two brands were grouped together as one carrier based on DOT report records. Second, we removed some ad hoc airlines who reported to DOT for only a short period of time during which they met the threshold of 1% of domestic passenger revenues. Third, we adopted DOT's reporting as our method to classify airlines during the grace period following airline merger and acquisition. These airlines, the acquirer and target airlines, were still treated as two separate airlines until they officially reported jointly as one carrier to DOT. This three-step data cleaning process yielded a total of 27 airlines in our data set, with reporting records between 8 quarters and 64 quarters. A summary of the airline list appears 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	AMERICA WEST	2004 Q1	2005 Q4	8
5	AMERICAN	2004 Q1	2019 Q4	64
6	ATA	2004 Q1	2006 Q4	12
7	ATLANTIC SOUTHEAST	2004 Q1	2011 Q4	32
8	COMAIR	2004 Q1	2010 Q4	28
9	CONTINENTAL	2004 Q1	2011 Q4	32
10	DELTA	2004 Q1	2019 Q4	64
11	ENDEAVOR	2010 Q4	2019 Q4	13
12	ENVOY	2004 Q1	2019 Q4	56
13	EXPRESSJET	2004 Q1	2019 Q4	62
14	FRONTIER	2005 Q2	2019 Q4	59
15	HAWAIIAN	2004 Q1	2019 Q4	64
16	INDEPENDENCE	2004 Q1	2005 Q4	8
17	JETBLUE	2004 Q1	2019 Q4	64
18	MESA	2006 Q1	2019 Q4	40
19	NORTHWEST	2004 Q1	2009 Q4	24
20	PSA	2018 Q1	2019 Q4	8
21	REPLUBLIC	2018 Q1	2019 Q4	8
22	SKYWEST	2004 Q1	2019 Q4	64
23	SOUTHWEST	2004 Q1	2019 Q4	64
24	SPIRIT	2015 Q1	2019 Q4	20
25	UNITED	2004 Q1	2019 Q4	64
26	US AIRWAYS	2004 Q1	2015 Q2	41
27	VIRGIN AMERICA	2012 Q1	2017 Q4	24

3.2 Dependent Variables

The immediate impact of implementing the new baggage fee policy is reflected in the potential immediate increase in revenue and subsequently in profit – our first dependent variable. Following current airline research (Dresner and Xu 1995, Tsikriktsis 2007, Mellat-Parast et al. 2015), we elect to use OPOR (operating profit/operating revenue) to reflect carrier's financial performance to avoid the absolute size differences among carriers. The relevant measures were taken from DOT Schedule P12. Revenue and profit are reported on a quarterly basis by DOT and are log transformed following current econometrics research practice (Wooldridge 2010).

Our second dependent variable is OTP. As discussed in Section 2.2, we elect to use carrier-induced delays to estimate baggage fee policy impact on OTP to isolate other compounding effect caused by weather delays and other non-carrier controllable delays. To match financial measures, which are only available at quarterly level, carrier-controllable delays were collapsed into quarterly level.

Consumer complaints, our third dependent variable, have been studied in airline research in the form of complaint per 1,000 passengers (Steven et al. 2012) and total number of complaints toward flight, baggage, and overbooking (Halstead et al.1996). Since March 2002, DOT has classified consumer complaints into the following 12 categories: flight problem, over-sales, reservation/ticketing/boarding, fares, refunds, baggage, customer service, disability, advertising, discrimination, animals, and other. We elect to use the total number of consumer complaints to capture the direct as well as the indirect/halo effect of the baggage fee policy on consumer complaints. Consumer complaints were taken from the DOT Monthly Consumer Report. Consumer complaints were log transformed in our analysis following existing practices in airline research (Lapré and Tsikriktsis 2006, Steven et al. 2012).

3.3 Control Variables

We briefly discuss control variables in this section. Some of the covariates may impact all the three outcome variables while other covariates may only impact only one or two outcome variables.

Load factor and Yield

Load factor and yield are the two classic variables in the airline literature used to investigate their respective impact on financial performance (Alan and Lapré 2018). Tsikriktsis (2007) concluded that a 1% increase in passenger load factor resulted in a 0.63% increase in OPOR. Zou and Chen (2017) also found that higher load factor had positive effects on carrier's profitability. Using proper yield management systems, American Airlines credited a revenue increase of $500 million per year while Delta Airlines generated additional revenues of $300 million per year (Boyd 1998). In addition, it is reasonable to expect that higher load factor will potentially impact OTP due to the longer boarding and deboarding time. Therefore, we include load factor and yield as our covariates.

Operational Metric Measures

Airline fleet structure and airline network structure have been found to directly impact airline operational performance in the airline literature (Lapré and Scudder 2004, Manchiraju et al. 2023). Following extant airline research (Lapré and Scudder 2004, Manchiraju et al. 2023), we compile three operational metrics: fleet utilization, fleet heterogeneity, and network sparsity to specifically evaluate their impact on operational performance as well as their potential impact on OPOR and consumer complaints.

Enplaned Passengers

The number of enplaned passengers is a commonly used control variable in baggage fee and airline related research. Prince and Simon (2009) used the monthly total number of passengers at route level to control for market demand. Nicolae et al. (2017) calculated the expected averaged number of passengers at flight level as a control for consumer demand. Yimga (2017) aggregated the number of passengers to itinerary level to control for market demand. Because our unit of analysis is at carrier level, we use the total monthly enplaned passengers of each carrier, as reported by DOT. This variable is taken from the DOT Air Travel Monthly Consumer Report.

We include the number of passengers as a control variable for two main reasons. First, the number of passengers has a direct impact on financial performance, as greater numbers of passengers translate to more revenue. Second, the number of passengers is expected to have an indirect impact on OTP and consumer complaints. A greater number of passengers imposes greater operational challenges, such as checking in, boarding, deplaning, and baggage handling, all of which could lead to potential delays in OTP as well as more consumer complaints.

Number of Employees

As part of airline scheduling practices, airline staff are assigned to each airport to service ground operations, gate operations, and flight operations (Ernst et al. 2004). The number of staff scheduled directly impacts our three outcome variables. First, employees, as the human capital of a firm, play a significant role in a firm's revenue productivity (Chowdhury et al. 2014). Second, how efficiently the ground operating staff handles baggage and other operations directly impacts OTP. Third, line personnel, such as check-in agents, gate agents, and flight attendants, play important roles in shaping consumer complaint behavior through their employee-customer interactions (Anderson et al. 2009). DOT only reports the total number of full-time employees, but not every employee has a direct impact on the three outcome variables. Those employees who have direct impacts on the three outcome variables, such as pilots, flight attendants, check-in and gate agents, and ground operating agents, account for 85% of an airline's employees (DOT 2024). Hence, we take 85% of the full-time employees reported by DOT as a control variable in our model. This variable is taken from Schedule P1(a) in DOT's Form 41 Financial Data Report.

Market Share

Market share is a commonly used control variable in airline research to control for a carrier's market power. Market share also has relationships with our three outcome variables. First, the relationship between market share and firm financial performance has long been established (Buzzell et al. 1975). Second, Suzuki (2000) found that carriers with better on-time performance enjoy greater market share. To account for the potential reverse causality, we use market share to control for its impact on OTP. Lastly, greater market share indicates greater number of passengers. With everything else being equal, greater number of passengers is likely to result in more complaints. Therefore, we included market share as a control variable.

Market share has been operationalized in several ways in airline research. Rupp et al. (2006) operationalized a carrier's market share as the number of total scheduled flights by a carrier divided by the total number of scheduled flights on that route. Prince and Simon (2009) defined market share as the carrier's number of enplanements on the route divided by the total number of enplanements. Shaffer et al. (2000) calculated individual carrier market share as the ratio of a carrier's monthly revenue passenger miles to the monthly sum of all carriers' revenue passenger miles. We elect to follow Shaffer et al. (2000) to operationalize market share, because this operationalization calculates market share at carrier level, consistent with our unit of analysis. The related variables to construct market share are taken from Schedule T1 in DOT Form 41 Air Carrier Summary Data.

Other Macro-Economic Control Variables

One of the most challenging problems empirical studies face in airline research is endogeneity bias (Scotti and Dresner 2015, Yazdi et al. 2017). The changes in OPOR, OTP, and consumer complaints are also likely to be driven by other macro-economic variables. As such, endogeneity bias could yield a positive correlation between the variables measuring the changes in our three outcome variables and the error term. Therefore, we also include selected macro-economic variables that could affect consumer travelling behavior to further address the concern of endogeneity bias. The first variable is percentage change of GDP (Bureau of Economic Analysis 2024), retrieved from Federal Reserve Economic Data. The second variable is airline's fare, retrieved from DOT. Lastly, we control for carrier fixed effect, year fixed effect, and quarter fixed effect (as our data is at a quarterly level) to further control for all the other effects that are not captured by our selected variables.

3.4 Summary Statistics

Table 2 provides a list of variables used in our analysis, their definitions, and their data sources. Those variables that are reported at a monthly level to DOT are subsequently collapsed into a quarterly level to match financial measures that are only available at a quarterly level. Appendix A shows the summary statistics and correlation matrix between the three outcome variables and other variables. We take the natural logarithm for relevant variables following current airline literature (Garrow et al. 2012, Scotti and Dresner 2015, Yazdi et al. 2017) as well as the standard econometric practice (Wooldridge 2010).

Table 2 Variables Used in Analysis

Variable	Definition	Data Source
OPOR	Airline's operating profit divided by operating revenue as reported each quarter.	DOT Schedule P1.2 in Form 41 Financial Data
Carrier Delays	The cause of the cancellation or delay was due to circumstances within the airline's control (e.g. maintenance or crew problems, etc.)	DOT On-time Reporting On-time Performance
Consumer Complaint	Total number of consumer complaints each month.	Table 3 in DOT Air Travel Monthly Consumer Complaint Report
Load Factor	Revenue passenger miles divided by available seat miles	DOT Schedule T1
Yield	Passenger revenues divided by revenue passenger miles (RPMs)	DOT Schedule T1 and P1.2
Fleet Utilization	Block Aircraft Hours divided by Aircraft Days	DOT Schedule P52
Fleet Heterogeneity	Blau index of different aircraft types within an airline's fleet in each quarter	DOT Schedule T100
Network Sparsity	Sum of squared proportions of flights originating from each airport in an airline's network in each quarter	DOT Schedule T100
Enplaned Passengers	Number of enplaned passengers	DOT Schedule T1 in Form 41
Full-time Employees	Number of Full-Time Equivalent Employees	DOT Schedule P1(a) in Form 41
Market Share	The ratio of a carrier's quarterly revenue passenger miles to the sum of revenue passenger miles of the total 7 carriers in that quarter	DOT Schedule T1 in Form 41
GDP % Change	Quarter over quarter change in GDP	Federal Reserve Economic Data
Fare	Airline fares	DOT DB1B

* Monthly data was aggregated into quarterly data to match the quarterly financial measures.

4. METHOD

4.1 Baggage Fee Implementation Dates

As our main research objective is to investigate the impact of baggage fee implementation on the three outcome variables, we first compile baggage fee implementation dates for relevant carriers. We collect baggage fee implementation dates from related literature (Barone et al. 2012) and accordingly verify the dates with relevant industry reports such as airline news release and other web sources. Table 3 summarizes the dates when relevant airlines officially started charging fees for *all* checked bags. Note that there is a short period of time when several airlines only charged fees for the second checked bags before eventually started charging fees for all checked bags. However, the time gap between these two events is so short that the two events will fall into the same quarter or just a quarter apart. Therefore, we only investigate the dates when all related airlines started to charge fees for all checked bags. Also, after the initial implementation of charging fees for all checked bags for $15, airlines have increased the fees from $15 to $20 or $25. In our current study, we only focus on the initial implementation of charging $15 for all checked bags and leave the multiple implementations as a future research avenue.

Table 3 Implementation Date of Charging Baggage Fees for All Checked Bags

Airlines	Implementation Date	Occasion
American	15-Jun-08	Q2 2008
United	13-Jun-08	Q2 2008
US Airways	9-Jul-08	Q2 2008
Northwest	10-Jul-08	Q3 2008
Continental	7-Oct-08	Q3 2008
Frontier	13-Sep-08	Q3 2008
Delta	5-Nov-08	Q4 2008
AirTran	12-Nov-08	Q4 2008
Alaska Air	1-May-09	Q2 2009

4.2 Event Study Methodology

Event study methodology is a popular method adopted to estimate policy impact in operations management field. We accordingly adopt this method to evaluate the impact of baggage fee implementation on our three outcome variables. The baseline model of an event study can be written in Equation 1:

$$Y_{it} = \alpha_i + \lambda t + \sum_{l=-K}^{-2} \mu_l \boldsymbol{D_{it}^l} + \sum_{l=0}^{L} \mu_l \boldsymbol{D_{it}^l} + \beta X_{it} + \varepsilon_{it} \qquad \text{Equation 1}$$

α_i and λt are vectors of carrier fixed effects and time fixed effects respectively. $\boldsymbol{X_{it}}$ is the vector of time-varying covariate from Section 3.3 and ε_{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 baggage fees were charged, such as (–2, –1, 0, 1, 2, 3) where 0 is when the baggage fee policy started, –2 is two time periods before the policy, and 2 is two time periods after the policy. 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). As we intend to investigate both the immediate effect ($t = 0$) and the long-term effect ($t >= 1$), we elect to omit $t = -1$ to avoid multicollinearity and while still being able to estimate the immediate effect when $t = 0$. $\sum_{l=-K}^{-2} \mu_l D_{it}^l$, therefore, captures the time periods up to the second quarter before the policy implementation. $\sum_{l=1}^{L} \mu_l D_{it}^l$ represents the time periods after the policy implementation (l = 1). The main parameters of interest in Equation 1 are the μ_ls which captures the differences in the outcome Y_{it} between carriers who charge baggage fees and carriers who do not charge baggage fees at l time periods apart from the time when the policy started.

4.3 Model Set Up

The first task in setting up the model to estimate Equation 1 is to select an appropriate event window for the *l*. Reviewing the related literature, different post-event windows were used to estimate baggage fee policy impact, ranging from 4 quarters (Nicolae et al. 2017), 10 quarters (Scotti et al. 2016), to 12 quarters (Yazdi et al. 2017). As previously mentioned, airlines have increased their baggage fees from $15 to $20 or $25, approximately six quarters after the initial implementation of $15. Therefore, we elect to choose 6-quarter as our post-event window to isolate the potential multiple implementation effect. Accordingly, we also use 6 quarters as our pre-event window to compare the before-and-after effect.

The second task in model set-up is to select an appropriate estimator. Equation 1 is generally estimated using Two-way Fixed-effects (TWFE) in traditional operations management studies by comparing carriers who charged baggage fees (treatment group) VS carriers who did not charge baggage fees (control group). However, the recent advancement in econometrics reveals major pitfalls of TWFE regression in estimating policy impact when the policy implementation dates are different, such as in our case (Callaway and Sant'Anna 2021, Baker et al. 2022). The main reason is that TWFE uses already-treated units as a control group to estimate the treatment effect while these "already-treated" units are not clean controls (Baker et al. 2022). For example, in Table 3, when estimating the effect of implementing baggage fees for Alaska Airline who implemented the policy in Q2 2009, apart from using other airlines from our 27 airlines listed in Table 1, TWFE estimator also uses all the other airlines in Table 3 as a control group. However, all the other airlines in Table 3 have already implemented baggage fees before Alaska and their corresponding OPOR, OTP, and consumer complaints may have already changed due to the policy implementation itself. Hence, these already-treated carriers could not serve as a clean control group. If already-treated units are used as the control group, the corresponding estimates are proven to be biased (Baker et al. 2022).

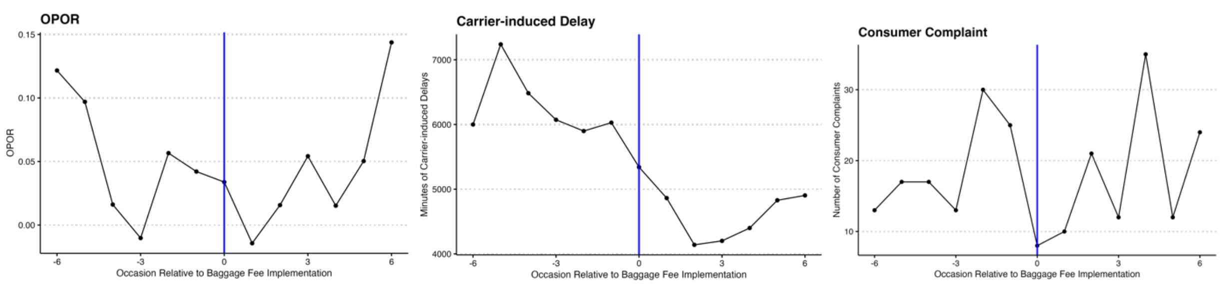
To solve this issue associated with TWFE estimator when drawing causal inference on treatment effect using the original data, we follow the recommendation from Baker et al. (2022) to reconstruct our data into a stacked data where each single carrier in Table 3 is only being compared to other carriers who did not implement baggage fee policy during the event window we selected (–6 quarters to 6 quarters). For example, for Alaska Airline who implemented baggage fees in Q2 2009, only carriers who did not charge baggage fees 6 quarters before and 6 quarters after Q2 2009 were used as the effective controls for Alaska Airline. This process is repeated for all the treated carriers in Table 3. Then, TWFE estimator is applied on this clean data set to estimate the treatment effect.

5. ANALYSIS AND RESULTS

5.1 Model Free Evidence

Before we conduct our statistical analysis, we first present model-free plots to test if our data indeed presents the hypothesized relationships. We use the event window of – 6 quarters to 6 quarters to cut our data. Then, we average the values of OPOR, carrier-induced delays, and number of consumer complaints for all treated carriers as is shown in Table 3. Finally, we plot these averaged values in Figure 1. Due to the vastly different scales, we plot three plots for each outcome variable. From Figure 1, we see a clear instant drop in OPOR as well as a U-shaped curve for OPOR, supporting H1a and H1b; a clear instant drop in carrier-induced delays (the reverse of OTP) and a U-shaped curve for carrier-induced delays, supporting H2a and H2b. However, for consumer complaints, the plot shows a seasonal plot and does not provide any support to H3a and H3b. Model-free evidence, anyways, does not equal to statistical significance. Therefore, we present our statistical analysis results in the next section.

Figure 1 Model Free Evidence



5.2 Statistical Results from Event Study

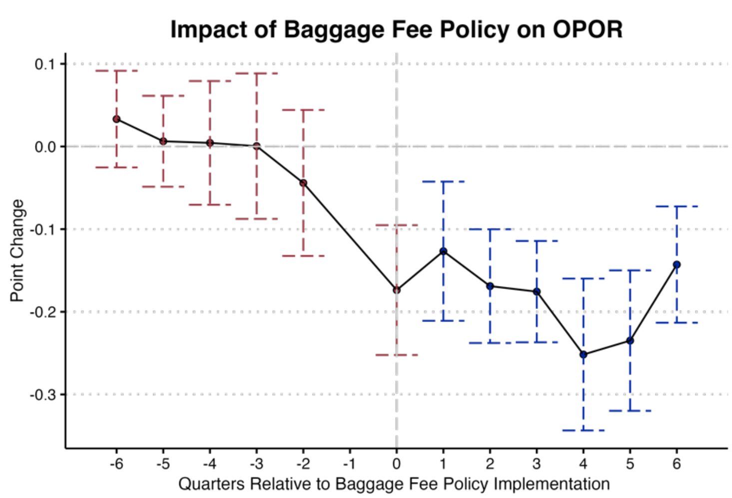
Our first set of hypotheses predicts that financial performance will decrease immediately upon policy implementation (H1a), but over time, financial performance will demonstrate a U-shaped curve (H1b). We present our results in Table 4. The immediate impact is tested by the coefficient of $l = 0$ with a parameter estimate of –0.17 ($p = 0.000$), supporting H1a and indicating that compared with carriers who did not implement baggage fees, carriers who implemented baggage fees suffer from a 0.17 unit decrease in OPOR (OPOR is in its original format) immediately upon policy implementation. The coefficients of $l = 1 \sim 6$ test the long term effect measured as six quarters following the policy implementation. We see that all the coefficients are statistically significant. To better detect the trend, we plot the six coefficients in Figure 2. The horizontal axis represents the six post-event quarters. The vertical axis represents OPOR. From Figure 2, we see a clear U-shaped curve, supporting H1b, indicating that after initially suffering from a deteriorated financial performance, carriers gradually recovered from it, aligning with our theoretical reasoning that consumers gradually resigned from rejecting the new policy.

Table 4 Effect of Baggage Fee Policy on OPOR – Estimated using Stacked Regression

Dependent variable is OPOR			
l Wave relative to mergers	$\widehat{\mu_l}$	Standard Error	t-Value
-6	0.02	0.03	0.77
-5	0.002	0.03	0.08
-4	0.006	0.03	0.17
-3	0.008	0.04	0.21
-2	-0.04	0.04	-1.14
0	**-0.17*****	**0.02**	**-7.81**
1	-0.13**	0.04	-3.13
2	-0.18***	0.04	-4.28
3	-0.20**	0.04	-4.08
4	-0.27***	0.06	-4.42
5	-0.24***	0.04	-5.41
6	-0.14**	0.03	-4.02
Load factor	1.62***	0.38	4.19
Yield	1.05**	0.28	3.77
Fleet Utilization	0.06	0.08	0.75
Fleet Heterogeneity	-0.001	0.01	-0.01
Network Sparsity	0.26	0.21	1.21
# of Enplaned Passengers	-0.16	0.22	-0.73
# of Employees	-0.04	0.29	-0.13
Market Share	0.48	0.29	1.66
% GDP change	-4.56***	0.87	-5.24
Fare	-1.06*	0.46	-2.33
Fixed Effects Included: Carrier, Year, Quarter			
R^2 = 0.633 RMSE = 0.122			

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

Figure 2 Hypothesis 1 OPOR Graph



Our second set of hypotheses posits that OTP will improve immediately upon policy implementation (H2a), but over time, OTP will demonstrate an inverted U-shaped curve (H2b). Remember that to isolate the irrelevant factors that were included to calculate OTP by DOT, such as the 40%+ weather delays, we use carrier-induced delays to proxy OTP. Therefore, an immediate decrease in carrier-induced delays and a U-shaped curve in carrier-induced delays will support out hypotheses. Test results are reported in Table 5 and

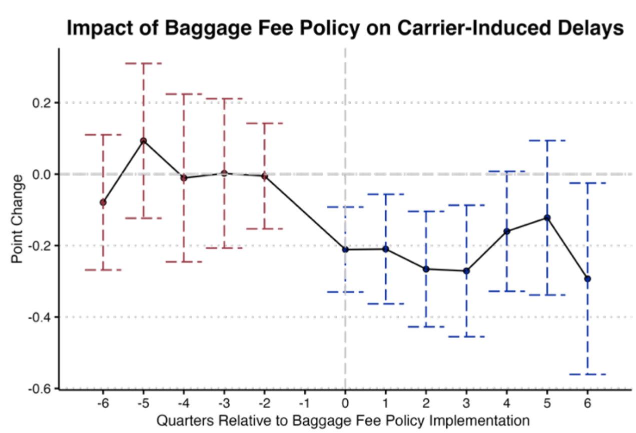
Figure 3. Again, the coefficient of *l = 0* tests the immediate effect (–0.21, $p = 0.004$), indicating that in the quarter when the baggage fees were charged, carrier-induced delays (log-transformed) were reduced by about 20% (1 – exponential (–0.21)) for carriers who charged fees, compared with the quarter prior to the policy implementation and other carriers who did not charge fees, supporting H2a. The effect in the recovery stage was tested by the six post-event time indicators $l = 1 \sim 6$. With the exception of the coefficient on $l = 5$, which is non-significant, all the other coefficients on the other five time indicators are statistically significant. In addition, we observe a U-shaped curve on the post-event trend in Figure 3. Therefore, we can conclude that H2b is partially supported.

Table 5 Effect of Baggage Fee Policy on Carrier Delays – Estimated using Stacked Regression

Dependent variable is carrier-controllable delays			
l Wave relative to mergers	$\widehat{\mu}_l$	Standard Error	t-Value
-6	-0.08	0.09	-0.87
-5	0.09	0.10	0.90
-4	-0.01	0.11	-0.09
-3	0.002	0.10	0.02
-2	-0.005	0.07	-0.07
0	**-0.21****	**0.06**	**-3.73**
1	-0.21*	0.07	-2.88
2	-0.27**	0.08	-3.45
3	-0.27**	0.08	-3.09
4	-0.16*	0.10	-2.00
5	-0.12	0.13	-1.18
6	-0.29*	0.03	-2.29
Load factor	-1.05 °	0.55	-1.94
Yield	0.28	0.24	1.14
Fleet Utilization	-0.05	0.11	-0.47
Fleet Heterogeneity	0.06**	0.02	3.39
Network Sparsity	-0.47	0.25	-1.88
# of Enplaned Passengers	1.29*	0.49	2.60
# of Employees	-1.41	0.84	-1.67
Market Share	0.26	0.40	0.64
Fixed Effects Included: Carrier, Year, Quarter			
$R^2 = 0.938$ *RMSE = 0.181*			

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

Figure 3 Hypothesis 2 OTP Graph



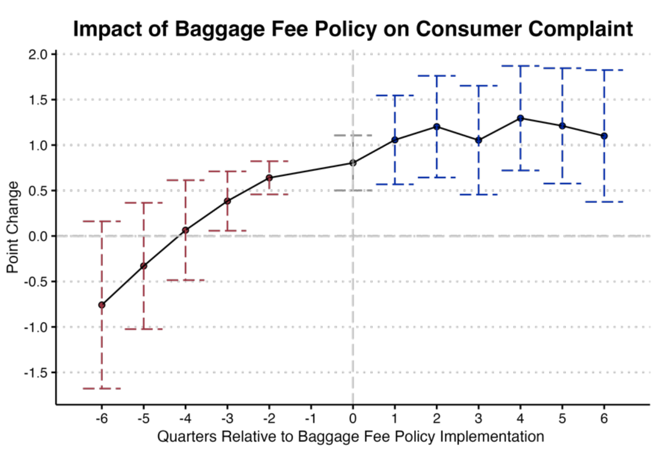
Turning our attention to the third set of hypotheses, consumer complaints are predicted to increase immediately upon policy implementation (H3a), but over time, complaints are expected to continue increasing at a diminishing rate (H3b). As with our previous hypotheses, the immediate impact is tested by the coefficient of $l = 0$ and the long-term impact is jointly tested by the coefficients of $l = 1 \sim 6$. In Table 6, the coefficient of $l = 0$ is 0.80 ($p = 0.000$). So, H3a was supported. The coefficients of $l = 1 \sim 6$ are all statistically significant. To clearly detect the trend, we also plot the event study coefficients in Figure 4, from which we see that the rate of consumer complaint increases at a tapered rate, supporting H3b.

Table 6 Effect of Baggage Fee Policy on Consumer Complaints – Estimated using Stacked Regression

Dependent variable is consumer complaints			
l Wave relative to mergers	$\widehat{\mu}_l$	Standard Error	t-Value
-6	-0.76	0.44	-1.73
-5	-0.33	0.33	-0.99
-4	0.06	0.26	0.24
-3	0.38*	0.15	2.47
-2	0.64**	0.08	7.36
0	0.80***	0.14	5.57
1	1.06***	0.23	4.54
2	1.20***	0.26	4.51
3	1.05**	0.28	3.69
4	1.29***	0.27	4.73
5	1.21**	0.30	3.18
6	1.10***	0.35	4.02
Load factor	2.36 °	1.13	2.07
Carrier Delay	0.48***	0.10	4.44
Mishandled Bags	0.53*	0.19	2.70
Fleet Utilization	-0.07	0.15	-0.49
Fleet Heterogeneity	-0.06*	0.02	-2.27
Network Sparsity	0.62	0.43	1.44
# of Enplaned Passengers	-0.82	0.62	-1.31
# of Employees	0.15	0.63	0.23
Market Share	0.95	0.61	1.55
Fixed Effects Included: Carrier, Year, Quarter			
$R^2 = 0.942$ RMSE = 0.243			

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

Figure 4 Hypothesis 3 Complaint Graph



6. PARALLEL TREND ASSUMPTION, ROBUSTNESS TEST, AND POST HOC ANALYSIS

6.1 The Parallel Trend Assumption

As we adopt event study methodology to test our hypotheses, there arises a need to test the parallel assumption, i.e., between airlines who implemented baggage fees (the treated group) and airlines who did not implement baggage fees (the control group), the trend of OPOR, carrier-induced delays, and consumer complaints should be parallel between these two groups before the policy implementation such that the pattern observed during the post-event window (Figure 2, 3, and 4) cannot simply be explained by a linear violations of parallel trends (Rambachan and Roth 2023).

In operations management field, there are traditionally two different methods to test the parallel trend assumption. The first method is to plot the prior-treatment trends of the outcome variable between the treated group and the control group to observe if the two trendlines are parallel. However, graphic evidence does not equal to statistical significance – leading to the second popular method where statistical test is conducted to test the coefficients on the time indicators prior to the treatment, or the leads. However, the second method is only valid on a clean dataset where no already-treated units are used as a control group (Baker et al. 2022, Rambachan and Roth 2023). The reason is that applying TWFE to estimate the coefficients on the leads and lags (post-event time indicators) on the original dataset will result in biased estimates of the coefficients (Sun and Abraham 2021, Callaway and Sant'Anna 2021, Baker et al. 2022). However, as we have followed the recommendation of the recent econometrics development to construct a clean dataset, we can test the parallel trend assumption using the coefficients on the leads. Looking at the coefficients on the leads of OPOR (Table 4 and Figure 2) and OTP (Table 5 and Figure 3), we see that none of the coefficients on the leads is statistically significant, supporting the parallel trend assumption. Turning to the coefficients on the leads in Table 6 and Figure 4 for consumer complaints, we observe an interesting phenomenon: the coefficients on the leads start to become statistically significant three quarters before the policy implementation. Given the time-gap between announcement and implementation of baggage fees for the initial baggage fee policy implementation, this pattern seems to be aligned with the appraisal theory in that consumers' physiological response – in the form of raising more complaints – have started at the time when they first heard the news. We leave this as an interesting future research topic for researchers who conduct consumer behavior research in operations management.

6.2 Robustness Test

Several tests were conducted to assess the robustness of our findings. First, we exclude airlines with fewer measurement occasions from our analysis to see if this affected our results. Seven airlines with fewer than 20 quarters of observations were removed from the original data. The same analysis was conducted on this new dataset for all the three sets of hypotheses testing. Except for a slight change in parameter estimations, the significance levels and signs of all variables remained the same. We also extracted a smaller sample, which consists only of five major airlines that span the entire 64 quarters (Alaska, American, Delta, SkyWest, and United), to conduct the same analysis. The parameter estimates for all three outcome variables still had the same signs, although the associated significance levels differed slightly. In addition, we also trimmed down our data to four quarters before and four quarters after policy implementation and tested the models again. All parameter estimations for all the three outcome variables, although slightly different, were still statistically significant with the same sign. Overall, our robustness tests show that our findings are neither affected by changing data structure nor affected by changing model specifications.

6.3 Post Hoc Analysis – The Moderating Effect

Reviewing the related airline literature, we see that all related literature only investigated the direct effect of implementing baggage fees on different outcome variables. However, it is reasonable to expect that there may exist certain indirect effect regarding the impact of baggage fees on various outcome variables. For example, realizing that charging baggage fees may result in lower customer satisfaction, airlines who charged baggage fees may subsequently offer lower ticket price to keep their customers happy. What will be the impact of these two variables together on financial performance? Therefore, in this section, we investigate the moderating effect during the baggage fee implementation process. We elect to study the moderating effect of fare on financial performance, the moderating effect of load factor on operational performance, and the moderating effect of mishandled bags on consumer complaints. Note that this is not an exhaustive list. Researchers can continue investigating other moderating effects in future research.

To test the moderating effect, we create a dummy variable “treated” with 1 being coded for all the six post-event quarters for the treated group and 0 for the control group. Then, different interaction terms were created between treated and the intended variables to test the moderating effect. In other words, we test a *static* moderating effect using the average treatment effect during the six post-event quarters to avoid the modeling challenges of interacting each time indicator in Equation 1 with the intended variables (Wooldridge 2010).

Table 7, Table 8, and Table 9 report the moderating effect on the three outcome variables. In Table 7, consistent with our first set of hypotheses testing results of policy impact on OPOR, the coefficient on “Treated” is negative, indicating that during the six quarters following baggage fee implementation, carriers who charged baggage fees did worse than carriers who did not charge baggage fees in terms of OPOR by 1.36 point. Although the coefficient on “Fare” is not statistically significant, the negative sign is expected as increasing fares may deter consumer demand. As we intend to use “Fare” as the moderator, the coefficient on the interaction term “Treated×Fare” represents the expected difference in the effect of one additional percent of increase in fare for the treated group versus the control group. The coefficient is 0.23 and statistically significant, meaning the moderator variable “Fare” makes the effect of “Treated” on OPOR less negative. In other words, for each one percent increase in fare (“Fare” was log-transformed) during the six post-event quarters, carriers who charged baggage fees witnessed about 0.2% increase in OPOR (OPOR not log-transformed), which is expected as increasing fares and charging extra baggage fees at the same time should help boost financial performance. The implication from this is that as carriers suffer from deteriorated financial performance following charging ancillary fees, one way to boost the deteriorated financial performance is to increase fare. However, increasing fare and charging ancillary fees simultaneously may further anger customers and make customer walk away. Carriers may carefully select implementation based on short-haul and long-haul flights, leisure traveling and business traveling, and etc. to avoid potential customer deflection (Warnock-Smith et al. 2017).

Table 7 The Moderating Effect of Fare on OPOR

Variables	*Coefficient*	*Standard Error*	*t-Value*
Dependent variable is OPOR			
Treated	-1.36**	0.41	-3.29
Fare	-0.13	0.09	-1.45
Treated*Fare	**0.23****	**0.07**	**3.34**
Load Factor	0.65**	0.21	3.04
Fleet Utilization	0.07*	0.03	2.58
Fleet Heterogeneity	-0.02**	0.008	-3.04
Network Sparsity	-0.03	0.02	-1.37
# of Enplaned Passengers	0.01	0.06	0.15
# of Employees	-0.03	0.02	-1.37
Market Share	0.03	0.04	0.77
% GDP Change	0.19	0.21	0.91
Fixed Effects Included: Carrier, Year, Quarter			
R^2 = 0.439 RMSE = 0.100			

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

Turning to the moderation effect of load factor on carrier-induced delays in Table 8, we see that the average treatment effect during the six post-event quarters is negative and statistically significant (– 0.17, $p = 0.08$), corroborating our second set of hypotheses testing results, indicating that the treatment group witnessed less delays potentially due to fewer checked bags. The coefficient on “Load Factor” is positive and statistically significant, indicating that higher load factor leads to more delays, as expected. However, the coefficient on the interaction term “Treated×Load Factor” is positive and statistically significant, meaning the moderator “Load Factor” makes the effect of “Treated” less negative. In other words, compared with carriers who did not charge baggage fees in the six post-policy quarters, for each 1% increase in load factor, carriers who charged baggage fees witnessed a 1.65% less carrier-induced delays (both load factor and delays are log-transformed). Therefore, for carriers who charged baggage fees, higher load factor has actually led to less carrier-induced delays. One reasonable explanation is that for those carriers who charged baggage fees, passengers still stay with their carriers but started to check fewer bags, which accordingly helped reduce delays, corroborating the findings from Cho and Dresner (2018) and Nicolae et al. (2016).

Table 8 The Moderating Effect of Load Factor on Carrier Delays

Variables	*δ^{DD}*	*Standard Error*	*t-Value*
Dependent variable is carrier-induced delays			
Treated	-0.17 °	0.09	1.87
Load Factor	1.43***	0.32	-4.54
Treated*Load Factor	**1.65****	**0.54**	**3.07**
Fleet Utilization	0.20**	0.06	2.92
Fleet Heterogeneity	0.09*	0.03	2.53
Network Sparsity	-0.14**	0.04	-3.10
# of Enplaned Passengers	0.80***	0.13	5.96
# of Employees	0.32***	0.08	3.99
Market Share	-0.27*	0.13	-2.00
Fixed Effects Included: Carrier, Year, Quarter			
R^2 = 0.900 RMSE = 0.298			

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

Table 9 reports the moderating effect of mishandled bags on consumer complaint. Mishandled bags are measured by DOT as mishandled bags per 1000 passengers. More mishandled bags should lead to more consumer complaints, as is demonstrated by the coefficient of “Mishandled Bags” (0.56, p = 0.001). As expected, the coefficient of “Treated” is positive, indicating that consumer complaints have increased in the six post-policy quarters for carriers who charged baggage fees, compared with those who did not charge baggage fees. The coefficient on the interaction term “Treated×Mishandled Bags” is negative and statistically significant, meaning mishandled bags as the moderator makes the effect of “Treated” less positive for carriers who charged baggage fees. In other words, for carriers who charged baggage fees, more mishandled bags did not increase total consumer complaints. One potential explanation is that due to the halo effect associated with consumer complaint behavior towards baggage fees, the driver of consumer compliant is not mishandled bags anymore. Rather, consumers tend to complain more on all dimensions after feeling betrayed by the event of charging baggage fees (Tuzovic et al. 2014).

Table 9 The Moderating Effect of Mishandled Bags on Consumer Complaint

Variables	*Coefficient*	*Standard Error*	*t-Value*
Dependent variable is consumer complaint			
Treated	0.87 °	0.50	1.73
Mishandled Bags	0.56***	0.13	4.29
Treated*Mishandled Bags	**-0.33 °**	**0.17**	**-1.86**
Carrier Delay	0.53**	0.15	3.51
Fleet Utilization	-0.01	0.09	-0.13
Fleet Heterogeneity	-0.05*	0.03	-2.10
Network Sparsity	0.57**	0.18	3.14
# of Enplaned Passengers	-1.15	0.94	-1.23
# of Employees	0.42	0.56	0.76
Market Share	0.99	0.85	1.16
Fixed Effects Included: Carrier, Year, Quarter			
R^2 = 0.937 RMSE = 0.268			

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

7. CONTRIBUTION

7.1 Theoretical Contribution

Our research contributes to knowledge accumulation in the following ways. First, in contrast to findings in extant airline literature showing revenue increase following baggage fee implementation (Henrickson and Scott 2012, Schumann and Singh 2014), our findings suggest that when OPOR (reported by DOT) was used as the outcome variable, airlines actually suffered from an immediate loss in financial performance. In addition, the deterioration in financial performance continues for the six quarters in our selected event window, corroborating the findings from Garrow et al. (2012) in that suboptimal outcomes were observed in the early stages of implementing baggage fees. Therefore, our findings reinforce the fact that increase in revenue does not necessarily translate into increase in the bottom line performance. Airline industry reports also show that record-high revenues yield modest profits (Garcia 2024) while lower revenues could also generate record-high profits (Holmes 2023). To this end, we call for researchers to start using bottom-line measures, such as profit, to evaluate policy impact as increased revenues and fares do not necessarily guarantee increased profit.

Second, in all related airline research that have investigated baggage fee policy impact on operational performance, weather delays were unfortunately included to operationalize operational performance – such as on-time arrivals and arrival delays. However, it is easy to see that the 40%+ weather delays in DOT data should not have anything to do with the effect of any policy implementation. Therefore, we call for future researchers to exclude weather delays and other unrelated delays from operational performance calculation when evaluating policy impact. When operational performance was narrowed down to carrier-induced delays that may most appropriately capture policy impact, we find that operational performance does not improve in a linear trend but instead demonstrates an inverted U-shaped performance improvement curve (the reverse to Figure 3 as Figure 3 is measured in delays), in contrast to existing findings showing a straightforward improvement path following baggage fee implementation (Scotti et al. 2016, Nicolae et al. 2017).

Third, despite the relatively constant number of passengers before and after the policy implementation, our findings suggest that consumer complaints have steadily increased in the six post-policy quarters, corroborating Tuzovic et al.'s (2014) survey results, which show that consumers tend to complain more after feeling betrayed by airlines. In addition, there seems to be a halo effect in consumers' response to the baggage fee policy as consumer complaints started to increase three quarters before the policy implementation. Without another full-length test of the halo effect, we leave this as a potential future research avenue.

Fourth, our study explores the moderating effect in the early stage of baggage fee implementation. To the best of our knowledge, our study is the first to analyze the moderating effect and accordingly find interesting results. To this end, we contribute to knowledge accumulation in the related literature. We find that in the early stage (six post-policy quarters) of implementing baggage fees, for carriers who charged baggage fees, increasing fares will compensate for the loss in financial performance; increasing load factors does not create more delays; and increased number of mishandled bags does not lead to more consumer complaints. We call for future research to continue investigating other potential moderators to continue knowledge accumulation in this stream of literature.

Lastly, with the advancement of econometrics, traditional TWFE, when used on the original dataset, has been proven to generate biased estimates in situations where there are multiple time periods with variation in treatment timing (Callaway and Sant'Anna 2021, Baker et al. 2022). We accordingly adopt the latest approaches to avoid the biases associated with TWFE by constructing a clean dataset to estimate the treatment effect. We, therefore, call for future operations researchers to adopt the latest econometrics advancement to increase research rigor.

7.2 Managerial Contribution

In 2008, airlines introduced baggage fees as a means to improve their financial situation (McCartney 2010b). Did this policy effectively improve the financial situation for airlines? When financial performance was measured by OPOR, unfortunately, we see OPOR actually deteriorated in the six post-policy quarters: OPOR first dropped an instantaneous 0.17 unit in the quarter when baggage fee policy was implemented; OPOR then continued to deteriorate before eventually started to improve but never exceeded the pre-policy level. The implication from our analysis is that despite the claimed revenue increase garnered from collecting ancillary fees, the bottom-line performance of airlines actually deteriorated. A potential explanation for this may be that when airlines de-bundled the ancillary fees, airlines also reduced their airfares to avoid angering customers (Jenkins et al. 2011). Henrickson and Scott (2012) and Scotti and Dresner (2015) confirmed a small but significant negative impact of baggage fees on airfares. As result of this, the revenue loss incurred by the reduction in airfares may have exceeded the extra revenue generated from charging baggage fees. On the other hand, our moderation analysis also shows that increasing airfares along with charging baggage fees can help improve OPOR for carriers who charged baggage fees. Combined, there are two implications here. First, despite carriers have touted collecting ancillary fees as a means to increase revenues, collecting ancillary fees to improve profit, however, does not materialize, at least in the six post-policy quarters in our data. Policymakers should anticipate the potential decrease in the bottom-line performance associated with collecting ancillary fees. Second, increasing airfare along with collecting ancillary fees can help improve profit but policymakers should carefully implement this strategy to avoid potential negative impact such as customer deflection.

Another decision-making process relates to service quality in the airline industry. The airline industry is a consumer-centric industry, where competition for customers is tight, and excellent service is key for sustainable business growth (Ostrowski et al. 1993, Park et al. 2006, Hussain et al. 2015). However, our results show that consumer complaints have increased even before the policy implementation despite the relatively constant number of enplaned passengers. It has been long established that dissatisfied customers may eventually walk away (Park et al. 2006, Hussain et al. 2015), hurting company's financial performance. Although our research does not focus on studying consumer deflection behaivor, the increased number of complaints clearly indicates dissatisfied customers, who may eventually walk away. Therefore, how to prevent consumers from walking away while implementing similar ancillary fees in the future is another task for airline policymakers to solve.

Our results also provide implications to airline operations managers. Our hypotheses testing results regarding carrier-induced delays indicate that the below cabin effect is impacted by the amount of baggage. Although DOT does not report the exact number of checked bags, our theoretical arguments and related airline literature (McCartney 2008b, Nicolae et al. 2017, Cho and Dresner 2018) suggest that, following the baggage fee implementation, consumers checked fewer bags. As a result, carrier-induced delays were reduced in the six quarters in our event window. The challenge here for operations manager is that as the number of bags were reduced, airlines may not need the same number of ground operations staff to handle the ground operations work. How to re-schedule the ground operations staff on other jobs to achieve operational efficiency is a task operations managers should accordingly take into consideration when facing future similar policy implementations that may result in a reduction or even an increase of checked bags in the early stage of policy implementation.

8. LIMITATION AND FUTURE RESEARCH

Like all research, our research demonstrates a few limitations. First, many carriers implemented multiple baggage fee policies from 2008 to 2012, i.e., carriers have increased baggage fees several times (Barone et al. 2012, Yazdi et al. 2017). While our research uses six post-policy quarters to only investigate the impact of the initial implementation, it would be interesting to examine the effects of multiple implementations by modeling multiple baggage fee policies. We observe this as a limitation, but, at the same time, it also serves as a potential topic for future research.

Our next limitation is that our unit of analysis is at the carrier level, the financial data of which is only reported quarterly. Accordingly, we aggregated all relevant variables at quarterly level. However, if researchers could have access to internal corporate financial data that can be analyzed at more minute levels, such as daily or weekly, researchers can then conduct more minute level analysis and potentially produce more interesting and specific results and accordingly provide more pertinent and detailed managerial recommendations. We treat this as another limitation, but it also serves as another fruitful future research avenue to use minute level financial data to further explore baggage fee policy impacts.

Our last limitation lies in our theoretical reasoning regarding the amount of baggage. As with the extant literature (Nicolae et al. 2017), we argue that the amount of baggage impacts operational performance to a great degree. However, DOT does not report the exact number of bags that were checked in and carried on. If the exact data were available, the number of bags could be linked both to revenue generated from charging baggage fees and to operating revenue and operating profit, allowing for a more precise investigation of the impact of baggage fees on carrier financial performance.

In conclusion, we leverage appraisal theory and the latest advancement in difference-in-difference to re-investigate the impact of baggage fees on carriers' financial performance, operational performance, and consumer complaints. In addition, we conduct moderation analysis and provide new findings to the literature of baggage fee implementation. We call for operations researchers to adopt the latest development in econometrics to assess treatment effect when there are multiple time periods with variation in treatment timing. We also call for operations researchers to continue investigate other moderation effects in policy implementation to provide more academic and managerial implications.

REFERENCES

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

Anderson, S.W., Baggett, L.S. and Widener, S.K. 2009. The impact of service operations failures on customer satisfaction: evidence on how failures and their source affect what matters to customers. *Manufacturing & Service Operations Management*, 11(1), 52-69.

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

Barone, G.J., Henrickson, K.E. and Voy, A. 2012. September. Baggage fees and airline stock performance: A case study of initial investor misperception. *Journal of the Transportation Research Forum*, 51 (1), 5-18.

Bureau of Economic Analysis. 2024. Gross Domestic Product. https://www.bea.gov/data/gdp/gross-domestic-product.

Ben-Akiva, M.E. and Lerman, S.R. 1985. Discrete choice analysis: theory and application to travel demand (Vol. 9). MIT press.

Boyd, A. 1998. Airline Alliance Revenue Management: Global alliances within the airline industry add complexity to the yield management problem. *Or Ms Today*, 25, 28-31.

Buzzell, R.D., Gale, B.T. and Sultan, R.G. 1975. Market share-a key to profitability. Harvard business review, 53(1), 97-106.

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

Cho, W. and Dresner, M.E. 2018. The impact of a baggage fee waiver on airline choice: Evidence from the Washington-Baltimore region. *Transportation Research Part A: Policy and Practice*, *112*, 4-17.

Chowdhury, S., Schulz, E., Milner, M. and Van De Voort, D. 2014. Core employee based human capital and revenue productivity in small firms: An empirical investigation. *Journal of Business Research*, 67(11), 2473-2479.

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

Department of Transportation (DOT). 2024. https://www.transtats.bts.gov/tables.asp?DB_ID=135.

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

Ernst, A.T., Jiang, H., Krishnamoorthy, M. and Sier, D. 2004. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153(1), 3-27.

Forbes, S.J., 2008. The effect of service quality and expectations on customer complaints. *The Journal of Industrial Economics*, 56(1), 190-213.

Garcia, M. 2024. Airlines Expect Record $996 Billion Revenue But Modest Profit. https://www.forbes.com/sites/marisagarcia/2024/06/03/airlines-expect-305-billion-net-profit-from-record-passengers/.

Garrow, L.A., Hotle, S. and Mumbower, S. 2012. Assessment of product debundling trends in the US airline industry: Customer service and public policy implications. *Transportation Research Part A: Policy and Practice*, 46(2), 255-268.

Halstead, D., Morash, E.A. and Ozment, J. 1996. Comparing objective service failures and subjective complaints: An investigation of domino and halo effects. *Journal of Business Research*, 36(2), 107-115.

Henrickson, K.E. and Scott, J. 2012. Chapter 8 Baggage Fees and Changes in Airline Ticket Prices. In Pricing behavior and non-price characteristics in the airline industry (pp. 177-192). Emerald Group Publishing Limited.

Higgens, M. 2010. Dodging Those Pesky Airline Fees. Retrieved from https://www.nytimes.com/2010/12/05/travel/05pracfees.html.

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.

Jiang, C. and Zheng, S. 2020. Airline baggage fees and airport congestion. *Transportation Research Part C: Emerging Technologies*, *117*, 102686.

Kalemba, N. and Campa-Planas, F. 2017. The Quality–Profitability Link in the US Airline Business: A Study Based on the Airline Quality Rating Index. Procedia Engineering, 187, 308-316.

Landon, E. L. 1977. A Model of Consumer Complaining Behavior. In Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, ed. R.L. Day, Bloomington, Indiana: Bureau of Business Research, 31-35.

Lapré, M.A. and Tsikriktsis, N. 2006. Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Science*, 52(3), 352-366.

Lazarus, R.S. and Folkman, S. 1984. Stress, appraisal, and coping. Springer publishing company.

Lazarus, R.S. 1991. Emotion and adaptation. Oxford University Press on Demand.

Manchiraju, C., Sohoni, M.G. and Deshpande, V. 2023. It's not simply luck: The impact of network strategy, schedule padding, and operational improvements on domestic on-time performance in the US airline industry. *Production and Operations Management*, *32*(11), 3559-3574.

McCartney, S. 2008a. Space race: A battle looms for the overhead bins. *Wall Street Journal* (June 17) D1–D2.

McCartney, S. 2008b. Why your bags aren't better off on a big airline. *Wall Street Journal* (September 2) D1–D3.

McCartney, S. 2008c. What it costs an airline to fly your luggage. *Wall Street Journal* (November 25) D1–D8.

McCartney, S. 2010a. What's behind new baggage fees. *Wall Street Journal* (April 29) D1–D4.

McCartney, S. 2010b. An airline report card: Fewer delays, hassles last year, but bumpy times may be ahead. *Wall Street Journal* (January 7) D1–D3.

McCartney, S. 2012a. The tough tactics to avoid luggage check-in fees. *Wall Street Journal* (February 2) D1–D3.

McCartney, S. 2012b. Reality check: Why airlines are shrinking flight times. *Wall Street Journal* (June 14) D1–D2.

Mellat-Parast, M., Golmohammadi, D., McFadden, K.L. and 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.

Morrison, S.A. and Winston, C., 2005. What's wrong with the airline industry? Diagnosis and possible cures. Hearing before the Subcommittee on Aviation, Committee on Transportation and Infrastructure, United States House of Representatives.

Nicolae, M., Arikan, M., Deshpande, V. and Ferguson, M. 2016. Do bags fly free? An empirical analysis of the operational implications of airline baggage fees. *Management Science*, 63(10), 3187-3206.

Nicolae, M.L., Ferguson, M.E. and Garrow, L.A. 2016. Measuring the benefit of offering auxiliary services: Do bag-checkers differ in their sensitivities to airline itinerary attributes? *Production and Operations Management*, *25*(10), 1689-1708.

Peterson, E.B., Neels, K., Barczi, N. and Graham, T. 2013. The economic cost of airline flight delay. *Journal of Transport Economics and Policy*, 47(1), 107-121.

Prince, J.T. and Simon, D.H. 2009. Multimarket contact and service quality: Evidence from on-time performance in the US airline industry. *Academy of Management Journal*, 52(2), 336-354.

Rambachan, A. and Roth, J. 2023. A more credible approach to parallel trends. *Review of Economic Studies*, *90*(5), 2555-2591.

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

Rupp, N., Owens, D. and Plumly, L. 2006. Does competition influence airline on-time performance? *Advances in Airline Economics*, 1, 251-272.

Scherer, K.R., Schorr, A. and Johnstone, T. eds. 2001. Appraisal processes in emotion: Theory, methods, research. Oxford University Press.

Schumann, H.O. and Singh, H. 2014. The market structure dynamics created by de-bundling of airline bag fees. *Journal of Economics and Economic Education Research*, 15(3), 187- 228.

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

Scotti, D., Dresner, M. and Martini, G. 2016. Baggage fees, operational performance and customer satisfaction in the US air transport industry. *Journal of Air Transport Management*, 55, 139-146.

Shaffer, B., Quasney, T.J. and Grimm, C.M. 2000. Firm level performance implications of nonmarket actions. *Business and Society*, 39(2), 126-143.

Steven, A.B., Dong, Y. and Dresner, M. 2012. Linkages between customer service, customer satisfaction and performance in the airline industry: Investigation of non-linearities and moderating effects. *Transportation Research Part E: Logistics and Transportation Review*, 48(4), 743-754.

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

Suzuki, Y. 2000. The relationship between on-time performance and airline market share: a new approach. *Transportation Research Part E: Logistics and Transportation Review*, 36(2), 139-154.

Suzuki, Y. 2004. The impact of airline service failures on travelers' carrier choice: A case study of central Iowa. *Transportation Journal,* 26-36.

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.

Tuzovic, S., Simpson, M.C., Kuppelwieser, V.G. and Finsterwalder, J. 2014. From "free" to fee: Acceptability of airline ancillary fees and the effects on customer behavior. *Journal of Retailing and Consumer Services*, 21(2), 98-107.

Warnock-Smith, D., O'Connell, J.F. and Maleki, M. 2017. An analysis of ongoing trends in airline ancillary revenues. *Journal of Air Transport Management*, *64*, 42-54.

Wooldridge, J.M. 2015. Introductory Econometrics: A modern approach. Nelson Education.

Xia, L., Monroe, K.B. and Cox, J.L. 2004. The price is unfair! A conceptual framework of price fairness perceptions. *Journal of Marketing*, *68*(4), 1-15.

Yazdi, A.A., Dutta, P. and Steven, A.B. 2017. Airline baggage fees and flight delays: A floor wax and dessert topping?. *Transportation Research Part E: Logistics and Transportation Review*, 104, 83-96.

Yimga, J. 2017. Airline on-time performance and its effects on consumer choice behavior. *Research in Transportation Economics*, 66, 12-25.

Zou, L. and Chen, X. 2017. The effect of code-sharing alliances on airline profitability. *Journal of Air Transport Management*, 58, 50-57.

APPENDICES

Appendix A Correlation Matrix (Log Transformed Variables)

	Panel A			Panel B													
	N	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 OPOR	1050	0.04	0.15	1.00													
2 Carrier Delay	1061	8.22	0.96	0.05	1.00												
3 Complaint	1010	4.12	1.26	0.07	0.48	1.00											
4 Load Factor	1058	(0.22)	0.07	0.34	(0.15)	0.39	1.00										
5 Yield	1050	(2.02)	0.28	0.11	0.32	(0.21)	(0.33)	1.00									
6 Fleet Utilization	1050	(1.10)	0.28	0.04	(0.16)	0.26	0.21	(0.28)	1.00								
7 Fleet Heterogeneity	1035	(0.59)	0.49	(0.00)	0.27	0.38	0.19	0.04	(0.00)	1.00							
8 Network Sparsity	1058	(2.42)	0.55	(0.13)	(0.63)	(0.28)	0.13	0.04	0.08	(0.12)	1.00						
9 # of Passengers	1058	15.61	0.90	0.09	0.73	0.76	0.23	(0.03)	0.20	0.35	(0.52)	1.00					
10 # of Employees	1050	10.51	1.12	(0.02)	0.66	0.77	0.20	0.06	0.18	0.42	(0.31)	0.93	1.00				
11 Market Share	1058	(3.34)	1.07	0.05	0.51	0.78	0.35	(0.15)	0.39	0.35	(0.30)	0.93	0.93	1.00			
12 Mishandled Bags	1009	2.47	0.57	(0.12)	0.44	(0.13)	(0.53)	0.41	(0.33)	0.05	(0.14)	(0.06)	(0.03)	(0.26)	1.00		
13 Fare	1061	5.84	0.08	0.24	(0.06)	0.19	0.50	(0.07)	(0.04)	0.05	(0.12)	0.08	0.02	0.12	(0.42)	1.00	
14 % GDP Change	1061	0.28	0.58	0.03	(0.00)	0.06	0.07	(0.10)	0.04	0.01	(0.02)	0.05	0.03	0.04	(0.07)	0.02	1.00