POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS

Vol. 27, No. 4, April 2018, pp. 734–755
ISSN 1059-1478 | EISSN 1937-5956 | 18 | 274 | 000734
DOI 10.1111/poms.12829


# Investigating Operational Predictors of Future Financial Distress in the US Airline Industry

Yasin Alan, Michael A. Lapré*
Owen Graduate School of Management, Vanderbilt University, 401 21st Avenue South, Nashville, Tennessee 37203, USA,
yasin.alan@owen.vanderbilt.edu, michael.lapre@owen.vanderbilt.edu

We investigate the predictive power of operational performance on future financial distress in the context of the US airline industry. We focus on four areas of operational performance: revenue management, operational efficiency, service quality, and operational complexity. Using quarterly data from 1988 through 2013, we find that airlines that have inferior revenue management, lower aircraft utilization, and higher operational complexity face higher future financial distress. Interestingly, average service quality, measured by on-time performance and mishandled baggage rate, is not associated with future financial distress, but extreme service failures, measured by long delays (over two hours) and passenger complaints with the government regarding mishandled bags, have a positive association with future financial distress. Using the association between current operational performance and future financial distress, we build a model to predict financial distress. Out-of-sample analyses show that our forecasting model outperforms a financial ratio-based benchmark model up-to eight quarters before the measurement of financial distress. Our findings inform firms, regulators, and investors by demonstrating that operational performance metrics contain useful information to predict future financial distress.

*Key words:* industry studies; airline operations; financial distress; bankruptcy
*History*: Received: January 2017; Accepted: November 2017 by Nitin Joglekar, after 2 revisions.

## 1. Introduction

Financial distress predictions provide early warning signals regarding a firm's future financial health. Early warning signals are useful for firms as well as their outside stakeholders (e.g., regulators and investors), especially in turbulent industries characterized by occasional bankruptcies and large financial losses. The US airline industry is one such industry. US airlines lost nearly $60 billion since deregulation of the industry in 1978–2009 (Borenstein 2011). Since 1978, more than 100 carriers have filed for bankruptcy, and many airlines have ceased to exist including well-known airlines such as Pan Am and TWA (Gritta et al. 2008). Given the importance of air transportation and the data availability, scholars have studied airline bankruptcies and airlines' restructuring efforts while operating under bankruptcy protection. However, scholars have largely overlooked whether operational performance can predict future financial distress.

Airline financial distress studies build on the broader bankruptcy literature. In the bankruptcy literature, Altman's (1968) classic paper introduced the use of financial ratios to predict corporate bankruptcy. Other scholars have advanced the field of bankruptcy prediction by (i) developing duration models instead of static models (Shumway 2001), (ii) examining industry effects (Chava and Jarrow 2004) and estimating industry-specific models, e.g., Becchetti and Sierra (2003) for manufacturers and Lu et al. (2015) for airlines, and (iii) expanding the phenomenon of binary bankruptcy classification to a more nuanced phenomenon of financial distress (Bharath and Shumway 2008). The bankruptcy prediction literature has often found that firms are less likely to go bankrupt if firms have (i) high sales to total assets, (ii) high retained earnings to total assets as well as high earnings before interest and taxes to total assets, (iii) high working capital to total assets, and (iv) high book value (or market value) of equity to total liabilities (Altman and Hotchkiss 2010).

Airline financial distress studies have proposed other financial ratios and industry-specific metrics to better capture operational and financial dynamics in the airline industry. For instance, Chow et al. (1991) demonstrate that high interest expense to total liabilities and low operating revenues to total miles flown are associated with higher bankruptcy risk. Subsequent research has established associations between some operational factors, such as fleet and labor productivity, and distress (e.g., Gudmundsson 2002, 2004). See Gritta et al. (2008) for a review of airline financial distress studies. These studies identify contemporaneous associations between an airline's operational performance and binary distress status (e.g., bankrupt vs. non-bankrupt), yet they do not perform out-of-sample forecasts to test the external validity of their findings. Hence, it is unclear whether such associations can be used to predict future financial distress. Moreover, binary classification models may not distinguish between healthy airlines and airlines on the verge of bankruptcy. Another stream of literature focuses on how airline bankruptcies affect pricing, capacity, and service quality. Airlines under bankruptcy protection offer lower prices (Borenstein and Rose 1995), reduce capacity, and downsize flight networks (Ciliberto and Schenone 2012). During restructuring efforts, airlines improve on-time performance and reduce mishandled baggage rates (Phillips and Sertsios 2013). Studies of bankruptcy episodes generate insights regarding firms' operations under bankruptcy protection, but do not reveal whether *current* operational performance can predict *future* financial distress.

More accurate financial predictions enable investors to make better investment decisions (Altman and Hotchkiss 2010). Similarly, such predictions allow regulators to better monitor the overall health of an industry. For instance, the Department of Transportation (DOT) monitors the financial strength of US airlines, using the Pilarski score model (Pilarski and Dinh 1999), which is a variant of the Altman's model (Gritta et al. 2008). However, the Pilarski score model relies solely on financial ratios. Hence, it is of interest to academics, practitioners, and regulators to test whether operational variables can be used to predict future financial distress. Accordingly, our main objective in this study is to examine whether current operational performance can predict future financial distress in the US airline industry. We identify four areas of operational performance: revenue management, operational efficiency, service quality, and operational complexity. We study the predictive power of all four areas on future financial distress using quarterly data from 1988 through 2013.

In a preliminary analysis, we focus on airline bankruptcy as an extreme form of financial distress. Extending existing binary bankruptcy classification models, we show that augmenting financial ratios with operational variables improves model fit. Subsequently, we focus on a more nuanced phenomenon of financial distress and use stock price data to measure financial distress with Bharath and Shumway's (2008) naïve distance to default (NDD). Linking operational performance to future NDD, we find that airlines that have inferior revenue management, lower aircraft utilization, and higher operational complexity face higher future financial distress. Interestingly, service quality is only associated with future financial distress if we consider extreme service failures. Average service quality, measured by on-time performance and mishandled baggage rate, is not associated with future financial distress. However, extreme service failures, measured by long delays (over two hours) and passenger complaints with the government regarding mishandled bags, are associated with future financial distress. Robustness tests show that these findings are not driven by the differences between legacy carriers (e.g., American, Delta) and low-cost carriers (e.g., JetBlue, Southwest) or the superior operational and financial performance of Southwest. Lastly, we use the above associations to perform out-of-sample forecasts of future financial distress. We show that our operational performance-based forecasting model outperforms a financial ratio-based benchmark model up-to eight quarters before the measurement of financial distress.

Joglekar et al. (2016, p. 1980) state that "the focus on context-rich industry data offers new and more nuanced theories of operations, and also strengthens the linkages between operations and sister disciplines like marketing and finance." As such, our paper contributes to the growing body of empirical literature that links operational performance to financial performance in different industries. For instance, in the pharmaceutical industry, failure of new drug development projects in the clinical trial stage decreases firm value (Girotra et al. 2007). In the airline industry, the stock market punishes long delays and flight cancellations (Ramdas et al. 2013). In the retail industry, inventory productivity predicts future stock returns (Alan et al. 2014). Despite the empirical evidence linking operational performance to stock performance, empirical association between operational performance and future financial distress is lacking.

To delineate our contribution, Table 1 compares our study with two research streams. Financial distress prediction studies in the broader finance literature (Panel A in Table 1) typically perform analyses in three steps. First, they propose potential predictors of future financial distress. Second, they perform regression analyses to test associations between the lagged values of those predictors and financial distress. Third, they compare the forecast accuracy of their models with a benchmark model (e.g., Altman 1968) through out-of-sample forecasting. These studies typically analyze large samples with firms from different industries. They do not test whether operational performance can predict future financial distress, because the determinants of operational success vary considerably across industries (e.g., inventory performance in the retail industry vs. new product development capabilities in the pharmaceutical industry).

Some of the airline financial distress studies listed in Panel B of Table 1 do pay attention to operational performance. However, these studies typically test associations between a small number of operational

**Table 1 Comparison of Our Study with the Financial Distress Literature**

| | Financial distress metric (binary or continuous?) | Lagged operational predictors | | | | Benchmarking via out-of-sample forecasting? |
|---|---|---|---|---|---|---|
| | | Revenue management? | Operational efficiency? | Service quality? | Operational complexity? | |
| *Panel A: General studies* | | | | | | |
| Altman (1968) | Binary | No | No | No | No | No** |
| Shumway (2001) | Binary | No | No | No | No | Yes |
| Chava and Jarrow (2004) | Binary | No | No | No | No | Yes |
| Bharath and Shumway (2008) | Both | No | No | No | No | Yes |
| *Panel B: Airline industry studies* | | | | | | |
| Chow et al. (1991) | Binary | No* | No | No | No | No** |
| Borenstein and Rose (1995) | Binary | No* | No | No | No | No |
| Pilarski and Dinh (1999) | Binary | No | No | No | No | No** |
| Gudmundsson (2002) | Binary | No* | No* | No | No* | No |
| Gudmundsson (2004) | Binary | No* | No* | No | No* | No |
| Ciliberto and Schenone (2012) | Binary | No* | No | No | No* | No |
| Phillips and Sertsios (2013) | Both | No* | No* | No* | No* | No |
| Lu et al. (2015) | Binary | No | No | No | No | No |
| Our study | Both | Yes | Yes | Yes | Yes | Yes |

*Notes.* *denotes that the paper uses an operational metric or multiple operational metrics to report a contemporaneous relation between this operational dimension and financial distress but does not explore whether current operational performance is associated with future financial distress. **denotes that the paper performs out-of-sample forecasting but does does not compare the forecast accuracy of its model with a benchmark model.

predictors and binary distress status (e.g., bankrupt vs. non-bankrupt). More importantly, they deviate from the three-step approach used in the broader finance literature in two important ways. First, they test contemporaneous associations between their predictors and financial distress. Hence, it is unclear whether such associations can be used to predict future financial distress. Second, as we discuss in section 4.3, applying a binary classification model to a small sample prevents these studies from comparing the forecast accuracy of their models with a benchmark model through out-of-sample forecasting. Thus, it is unclear whether these models can outperform a benchmark model that does not use operational predictors.

We contribute to the literature by combining the three-step empirical approach used in the broader financial distress literature with context rich industry data and thereby demonstrating that all four areas of operational performance contain information useful to predict future financial distress in the US airline industry. More specifically, we identify four areas of operational performance, yielding a comprehensive set of 11 potential operational predictors of future financial distress. In addition, we overcome the limitations of previous airline financial distress studies by (i) using a continuous financial distress metric, (ii) establishing associations between lagged values of operational predictors and financial distress, and (iii) showing the superior predictive power of our model compared to a financial ratio-based benchmark model with out-of-sample forecasting.

The study is organized as follows. In section 2, we develop hypotheses associating operational performance with future financial distress. Section 3 describes our data collection and variable construction. In section 4, we present a preliminary analysis of airline bankruptcies as well as limitations of using bankruptcy as the dependent variable. In section 5, we overcome these limitations in our analysis of financial distress. In section 6, we perform out-of-sample forecasts of future financial distress. In section 7, we generate managerial insights by comparing and contrasting two legacy carriers (i.e., American and Delta) and two low-cost carriers (i.e., JetBlue and Southwest). Lastly, in section 8, we close with a discussion of our contribution and questions for future research.

## 2. Identifying Operational Predictors of Financial Distress

In this section, we hypothesize associations between four areas of operational performance –revenue management, operational efficiency, service quality, and operational complexity– and financial distress. These associations enable us to identify possible operational predictors of financial distress in the context of the US airline industry.

### 2.1. Revenue Management

The basic idea behind yield management is "selling the right capacity to the right customers at the right prices" (Smith et al. 1992, p. 8). Airlines have extensively used yield management, also frequently referred to as revenue management, to maximize revenue given the fixed seat capacity of a plane (Bodea and Ferguson 2014, Talluri and Van Ryzin 2005).

Yield management systems typically comprise several elements such as overbooking, capacity allocation, and differential pricing (Metters et al. 2008). Airlines engage in overbooking –i.e., selling more reservations than seats available– to counter customer cancellations and no-shows. American Airlines estimated that 15% of seats on sold-out flights would be empty without overbooking (Smith et al. 1992). Low revenue passengers make reservations far in advance, whereas high revenue passengers make reservations much closer to flight departure. Hence, airlines engage in capacity allocation –reserving capacity for high revenue passengers– by denying reservations to low revenue passengers. Airlines charge different fares to different customer segments using restrictions such as number of days of advance purchase, ease of cancellation and refund policies, Saturday overnight stays, etc.

Optimally determining overbooking, capacity allocation, and differential pricing is very difficult. Former CEO of American Airlines, Bob Crandall (Smith et al. 1992, p. 31): "[Y]ield management is the single most important technical development in transportation management since we entered the era of airline deregulation in 1979. ... We estimate that yield management has generated $1.4 billion in incremental revenue in the last three years alone. ... We expect yield management to generate at least $500 million annually for the foreseeable future." Crandall's estimates imply that successful yield management has become a necessary condition to turn an operating profit. Indeed, by 1997 American estimated that yield management generated almost $1 billion in incremental annual revenue, whereas 1997 was the only year in American's history with operating profit approaching $1 billion (Cook 1998).

Airline industry practitioners use yield (unit revenue) and load factor (the degree to which an airline fills up planes with revenue paying passengers controlling for both seats and miles flown) as proxies to assess an airline's revenue management capabilities. For instance, Garvett and Hilton (1999, p. 181) state that "good revenue maximizers seek to obtain the best possible yields and load factors given the circumstance. Yet it is easy to get high yields by excessively raising fares and limiting the availability of discounts —and in the process ensuring nearly empty flights. Similarly, it is easy to maximize load factors by giving the product away. Neither strategy will lead to profitability. There are plenty of examples of both profitable and unprofitable low-load-factor airlines and vice versa. The same is true for yields. The art is to obtain both high load factors and high yields at the same time." Airlines with low yields and load factors could incur operating losses, which in turn could increase financial distress. Hence, using yield and load factor as proxies for revenue management, we test:

HYPOTHESIS 1. *Successful revenue management is negatively associated with future financial distress.*

### 2.2. Operational Efficiency

Collier and Evans (2014, p. 55) define operational efficiency as "the ability to provide goods and services with minimum waste and maximum utilization of resources." Companies from different industries such as Toyota Motor Corporation, Wal-Mart Stores Inc., and Southwest Airlines are considered best-in-class because their successful business practices drive operational efficiency, which in turn positively affects financial performance (Alan et al. 2014 and references therein).

We use fleet utilization and fuel efficiency as proxies for an airline's operational efficiency. Fleet utilization assesses effective capacity of planes –the percent of the time planes are used transporting passengers between departure at the gate of origin and arrival at the gate of destination. Higher fleet utilization allows an airline to operate more flights with the same number of aircraft. So, higher fleet utilization reduces investments in aircraft and at the same time increases revenue generating opportunities per aircraft. Both effects could reduce financial distress. Effective capacity is less than 100% because of (i) time spent turning around planes at the gate in between flights, (ii) idle time between the last flight of the day and the first flight of the next day, (iii) maintenance, and (iv) flying planes without passengers (to re-position planes). Southwest became known in the industry for its fast turnaround times keeping planes in the air where they make money (Gittell 2003). In contrast, legacy airlines –airlines serving inter-state routes at the time of deregulation, such as American and Delta– operate hub-and-spoke systems where planes arrive in banks of flights. Planes wait to allow for all possible connections before departure, leading to complex flight schedules (Atkinson et al. 2016). Thus, fleet utilization is lower for legacy airlines.

Our second measure for operational efficiency is fuel efficiency – the number of miles flown per gallon of fuel. The International Council on Clean Transportation (ICCT) performs benchmarking studies to compare fuel consumption across major US airlines. ICCT ranks American as the least efficient US airline, but also acknowledges that American has improved its fuel efficiency in recent years by phasing out fuel inefficient aircraft (e.g., Boeing 767-200 and MD-80) from its fleet. ICCT's findings indicate that Alaska Airlines consistently outperforms the rest of the industry in fuel efficiency and that there is a positive link between an airline's fuel efficiency and profitability (Li et al. 2015). Hence, increasing fuel efficiency could boost operating profitability, which in turn could reduce financial distress. Using fleet utilization and fuel efficiency as proxies for operational efficiency, we test:

HYPOTHESIS 2. *Operational efficiency is negatively associated with future financial distress.*

### 2.3. Service Quality

Poor service quality can lead to immediate extra costs. Whenever a service failure occurs, organizations engage in service recovery efforts, which require resources such as time spent by frontline employees and compensation expenses. In addition to immediate extra costs, there are also long-term, reputational consequences of poor service quality. Heskett et al. (1997) established a link between quality and profitability in their service-profit-chain framework. High service quality leads to customer satisfaction. Customer satisfaction, in turn, leads to repeat purchase and profitability. Conversely, poor service quality hurts repeat purchase and profitability, which in turn could increase financial distress.

In the airline industry, service quality is often measured by on-time performance and mishandled baggage (Phillips and Sertsios 2013). Late arrivals might misconnect, which could require costly re-booking of passengers on other flights of the same airline or even on flights of another airline. Tsikriktsis (2007) found that on-time performance improved profitability. Ramdas et al. (2013) estimated that the average cost of a flight delay for a US airline is $220 per minute. Mishandled bags need to be recovered. Airlines incur costs locating mishandled bags and delivering recovered bags to passengers. According to an industry report, between 2007 and 2016, the total cost of mishandled baggage for the airline industry was $27 billion (SITA 2017). Hence, poor service quality has a negative impact on financial performance.

Service failure severity affects satisfaction, trust, commitment, and negative word-of-mouth (Weun et al. 2004). Put differently, not all service failures are equal. Service failures that are more severe will be more harmful to the firm. Ramdas et al. (2013) investigated the impact of service quality on stock returns. The authors found that long delays (over two hours delayed) reduce contemporaneous stock returns, whereas late flights (over 15 minutes late) do not affect stock returns; that is, financial markets show a strong reaction to extreme service failures but not to average service quality.

After the occurrence of a service failure, an airline has an opportunity to engage in service recovery (Heskett et al. 1997, Lapré 2011). For example, if a passenger reports to the airline that she lost a bag, the airline has a chance to apologize, locate the missing bag, and deliver the bag to the owner. If customers are extremely dissatisfied after service recovery, they can engage in negative word-of-mouth. In the airline industry, passengers who are extremely dissatisfied have the option to file a complaint with DOT. DOT publishes the rate of consumer complaints (i.e., complaints per 100,000 passengers) against each major airline. Forbes (2008) found that consumers are more likely to complain to DOT if their expectations are not met. Behn and Riley (1999) found that consumer complaints increase operating expenses and decrease operating revenues and operating income. Luo (2007) found that consumer complaints to DOT about poor service by airlines harm future stock returns. Lapré (2011) decomposed the rate of baggage complaints (i.e., baggage complaints divided by the number of passengers) into two factors: the rate of mishandled bags (i.e., mishandled bags divided by the number of passengers) and the propensity to complain (i.e., baggage complaints divided by mishandled bags). Comparing mishandled baggage rates with the propensity to complain, the author found that airlines showed much more variation in the propensity to complain. That is, despite having similar mishandled baggage rates, some airlines are significantly better than others in terms of their service recovery capabilities. Based on negative consequences of poor service quality observed by Tsikriktsis (2007) and Heskett et al. (1997), we use two proxies of service quality – on-time performance and mishandled baggage rate – to test:

HYPOTHESIS 3A. *Higher service quality is negatively associated with future financial distress.*

Because extreme service failures can have a stronger association with an airline's financial performance (e.g., Ramdas et al. 2013), we test an alternative to Hypothesis 3A using two proxies of extreme service failures – long delays and propensity to complain to DOT after a mishandled baggage experience:

HYPOTHESIS 3B. *Extreme service failures are positively associated with future financial distress.*

### 2.4. Operational Complexity

Skinner's (1974) classic article introduced the notion of a focused factory. A factory that focuses on a narrow product mix for a particular market niche will outperform a plant that attempts a broader mission. Simplicity, repetition, experience, and task homogeneity build competence resulting in better performance. The concept of focus has also been extended to services. According to Heskett et al. (1997), service firms achieve higher profitability by having either market focus (focusing on specific target market segments) or operating focus (focusing the operations on servicing a narrow product mix). Firms that achieve both market and operating focus are nearly unbeatable. Lack of focus, on the other hand, leads to operational complexity, which in turn leads to additional operating expenses. Next, we discuss three elements that contribute to operational complexity in the airline industry.

One element that can contribute to significant complexity for an airline is operating a heterogenous fleet. Southwest grew from a local airline operating three routes in Texas to the largest US airline with just a single plane type – Boeing 737. Operating a single plane type reduces costs related to maintenance, gate configurations, training, servicing equipment, etc. At its inception, JetBlue copied the single-plane focus. However, in 2005, JetBlue added a second plane type to its fleet. Even just one additional plane type caused operational complexity and confusion related to new training programs, increased scheduling complexity due to reduced interchangeability for planes and pilots, greater task variety in servicing aircraft, more complex human resources issues, and the need to manage relations with multiple suppliers (Huckman and Pisano 2011). Costs associated with operational complexity are exacerbated for legacy carriers, which operate much more heterogeneous fleets compared to low-cost carriers, such as JetBlue (Atkinson et al. 2016).

A second element that contributes to operational complexity is flying to primary airports as opposed to secondary airports, which tend to be less congested. Southwest, for example, has traditionally chosen to operate out of secondary airports such as Chicago Midway and Dallas Love Field instead of Chicago O'Hare and Dallas Fort Worth (Gittell 2003). Operating out of less congested airports makes airlines less susceptible to costly service failures (Atkinson et al. 2016). Bonnefoy and Hansman (2004) provide a case study of the Boston regional airport system to illustrate how Southwest avoided the complexities associated with flying to the region's primary airport, Logan. In 1997, Logan reached 85% of its passenger capacity, leading to severe scheduling challenges and flight delays. Until entering Logan in 2009, Southwest avoided those problems by serving the Boston region from Providence and Manchester airports.

A third element that increases operational complexity is the extent to which a network of routes is sparse. A sparse network looks like a hub-and-spoke system with hubs that are connected to most spokes, whereas spokes have very few connections – typically to hubs. Legacy carriers use hub-and-spoke systems to allow for many connections. Connections require more customer contact points and extra baggage handling. Connections also increase dependencies. Hence, connections can lead to extra costs. Indeed, legacy carriers place a higher emphasis on the cost of flight delays (Deshpande and Arikan 2012). Even if an airline does not rely on connecting passengers, a sparse network can still suffer from dependencies. In 2007, JetBlue's network relied heavily on five key cities such as New York and Ft. Lauderdale. All destinations could only be reached from these key cities. A highly publicized Valentine's Day Crisis showed the adverse effects of dependencies. On February 14, 2007 a bad winter storm in New York crippled JetBlue's operations leading the airline to cancel 40% of its flights, affecting more than 131,000 customers (Huckman and Pisano 2011). In contrast, point-to-point systems –such as Southwest's route network– tend to connect many city pairs to offer more direct flights. Relying less on hubs or key cities reduces operational complexity and subsequently reduces costs. Since operational complexity carries additional costs, we use three proxies of operational complexity – fleet heterogeneity, use of primary airports, and network sparsity – to test:

HYPOTHESIS 4. *Operational complexity is positively associated with future financial distress.*

## 3. Data

Our main data sources are DOT, Compustat, and the Center for Research in Security Prices (CRSP). All US airlines are required to report financial and operating data on Form 41 to DOT. Form 41 is our source for quarterly data on revenue, cost, and transportation statistics. We obtain quality performance statistics from Air Travel Consumer Reports (ATCRs). Since October 1987, DOT has required major US airlines (airlines with at least 1% of US domestic passenger revenues) to file data on quality performance. DOT started publishing these statistics in ATCRs. Initially, there was some variation in reporting systems across airlines. The first quarter of 1988 was the first quarter for which quality performance could be compared across airlines (Lapré 2011). Consequently, we use quarterly data from 1988 to 2013. We collect flight level data from the DOT Bureau of Transportation Statistics website. We collect the financial data necessary to compute our financial distress metric from Compustat and CRSP. These data are accessed through Wharton Research Data Services.

### 3.1. Airlines

Our study requires operational, quality performance, and stock market data. Accordingly, our dataset is restricted to publicly traded major US airlines. Table 2 provides a summary of the airlines in our study. We have 25 airlines (10 legacy and 15 low-cost) and 1395 firm quarters. The number of quarters per airline varies from 7 (Piedmont Aviation) to 104 for the six airlines present during the whole time frame (Alaska, American, Delta, Southwest, United, and US Airways). Nineteen airlines entered the study after the first quarter of 1988 and/or left the study before the last quarter of 2013. Eight airlines in our study never declared bankruptcy. Fifteen airlines declared one bankruptcy, whereas two airlines, Trans World Airlines and US Airways, had multiple bankruptcy episodes. In total, our dataset includes 20 bankruptcies and 127 firm-quarters in which airlines stayed in bankruptcy. We obtained airline bankruptcy filing dates from the Airlines for America web site and cross checked the dates using the UCLA-LoPucki Bankruptcy Research Database, other airline studies such as Ciliberto and Schenone (2012), as well as the airlines' official web sites and Wikipedia pages.[1]

### 3.2. Dependent Variables

We study two dependent variables: *Bankruptcy Indicator*. $Bankrupt_{it}$ is an indicator variable that equals 1 if airline $i$ is bankrupt in quarter $t$ and 0 otherwise. We use this dependent variable in our preliminary bankruptcy analysis presented in section 4.

*Distance to Default*. We measure financial distress using the *naïve distance to default* metric constructed by Bharath and Shumway (2008). This metric is a simplified version of the Merton distance to default (DD) model, which applies Merton's (1974) bond pricing framework to model the evolution of a firm's debt and equity over time. Formally, the probability of bankruptcy implied by the Merton DD model can be written as

$$\mathcal{N}\left(-\frac{\ln(V/F) + \left(\mu - 0.5\sigma_V^2\right)T}{\sigma_V\sqrt{T}}\right), \tag{1}$$

where $\mathcal{N}(\cdot)$ is the cumulative standard normal distribution function, $V$ is market value of the firm (i.e., the sum of the market values of the firm's debt and equity), $F$ is the book value of the firm's debt, $T$ is the time horizon, $\mu$ is the firm's expected return, and $\sigma_V$ is the volatility of firm value.

Computing a firm's implied probability of bankruptcy via equation (1) is non-trivial because the market value of the firm's debt and its volatility are difficult to estimate. See Bharath and Shumway (2008)

**Table 2 Airlines in the Dataset**

| No. | Airline | First qtr in the sample | Last qtr in the sample | Qtrs in the sample | No. of bankruptices | No. of qtrs in bankruptcy |
|---|---|---|---|---|---|---|
| 1 | AirTran Airways | 2003 Q1 | 2012 Q1 | 37 | 0 | 0 |
| 2 | Alaska Airlines | 1988 Q1 | 2013 Q4 | 104 | 0 | 0 |
| 3 | America West Airlines | 1988 Q1 | 2005 Q4 | 72 | 1 | 14 |
| 4 | American Airlines* | 1988 Q1 | 2013 Q4 | 104 | 1 | 9 |
| 5 | American Eagle Airlines* | 2001 Q1 | 2013 Q4 | 52 | 1 | 9 |
| 6 | American Trans Air | 2003 Q1 | 2006 Q4 | 16 | 1 | 6 |
| 7 | Atlantic Southeast Airlines | 2003 Q1 | 2013 Q4 | 43 | 0 | 0 |
| 8 | Comair* | 2004 Q1 | 2010 Q3 | 28 | 1 | 7 |
| 9 | Continental Airlines* | 1988 Q1 | 2011 Q3 | 96 | 1 | 11 |
| 10 | Delta Airlines* | 1988 Q1 | 2013 Q4 | 104 | 1 | 7 |
| 11 | ExpressJet Airlines | 2003 Q1 | 2011 Q3 | 35 | 0 | 0 |
| 12 | Frontier Airlines | 2005 Q3 | 2013 Q4 | 34 | 1 | 6 |
| 13 | Hawaiian Airlines | 2004 Q1 | 2013 Q4 | 40 | 1 | 6 |
| 14 | Independence Air | 2003 Q1 | 2005 Q4 | 12 | 1 | 1 |
| 15 | JetBlue | 2003 Q1 | 2013 Q4 | 44 | 0 | 0 |
| 16 | Mesa Airlines | 2006 Q1 | 2013 Q4 | 32 | 1 | 5 |
| 17 | Northwest Airlines* | 1988 Q1 | 2009 Q4 | 88 | 1 | 7 |
| 18 | Pan Am. World Airways* | 1988 Q1 | 1991 Q3 | 15 | 1 | 3 |
| 19 | Piedmont Aviation | 1988 Q1 | 1989 Q3 | 7 | 0 | 0 |
| 20 | Pinnacle Airlines | 2007 Q1 | 2013 Q4 | 20 | 1 | 2 |
| 21 | Skywest | 2003 Q1 | 2013 Q4 | 44 | 0 | 0 |
| 22 | Southwest Airlines | 1988 Q1 | 2013 Q4 | 104 | 0 | 0 |
| 23 | Trans World Airlines* | 1988 Q1 | 2001 Q4 | 56 | 3 | 11 |
| 24 | United Airlines* | 1988 Q1 | 2013 Q4 | 104 | 1 | 15 |
| 25 | US Airways* | 1988 Q1 | 2013 Q4 | 104 | 2 | 8 |
| | | | Total | 1395 | 20 | 127 |

*Notes*. Our dataset consists of publicly traded major US airlines. We report each airline's name as well as the first and the last quarter (qtr) in which it appears in our sample. We also report the number of bankruptcy filings and the number of quarters each airline stayed in bankruptcy. For instance, US Airways filed for bankruptcy twice and spent a total of eight quarters in bankruptcy. * indicates a legacy carrier.

for the derivation of equation (1) and a discussion of the Merton DD model's underlying assumptions and computational hurdles. Bharath and Shumway (2008) propose an alternative, easy to compute distance to default metric called naïve distance to default (NDD), and show that NDD performs better in out-of-sample forecasts than the Merton DD model. Because of its simplicity and superiority to the Merton DD model, the recent finance literature uses NDD to measure financial distress (e.g., Campello and Gao 2017, Chava and Purnanandam 2011, Phillips and Sertsios 2013). We use NDD as our dependent variable in our financial distress analysis.

Revisiting equation (1), Bharath and Shumway (2008) approximate the market value of debt with its face value $F$. Consequently, the authors set $V = E + F$, where $E$ is the market value of the firm's equity. The authors approximate the volatility of the firm's debt as $\sigma_D = 0.05 + 0.25\sigma_E$, where $\sigma_E$ is the volatility of the firm's equity. Thus, the volatility of the firm's market value can be approximated by $\sigma_V = \frac{E}{E+F}\sigma_E + \frac{F}{E+F}\sigma_D = \frac{E}{E+F}\sigma_E + \frac{F}{E+F}(0.05 + 0.25\sigma_E)$. Lastly, the expected return, $\mu$, is approximated as the firm's stock return in the last period. Replacing each term in equation (1) with these approximations yields the bankruptcy probability, $\mathcal{N}(-NDD)$, where $NDD$ is:

$$NDD = \frac{\ln\left(\frac{E+F}{F}\right) + \left(\mu - 0.5\left[\frac{E}{E+F}\sigma_E + \frac{F}{E+F}(0.05+0.25\sigma_E)\right]^2\right)T}{\left[\frac{E}{E+F}\sigma_E + \frac{F}{E+F}(0.05+0.25\sigma_E)\right]\sqrt{T}}. \quad (2)$$

We compute $NDD_{it}$ for airline $i$ in quarter $t$ as follows. Following Vassalou and Xing (2004) and Bharath and Shumway (2008), we compute the face value of debt, $F_{it}$, as debt in current liabilities (Compustat item DLCQ) plus one-half of long-term debt (Compustat item DLTTQ). We compute the mean and the standard deviation of an airline's quarterly return using its daily stock price information from CRSP. Following Phillips and Sertsios (2013), we require at least 25 daily stock price observations for these computations. We compute the market value of equity as the number of shares outstanding times the closing stock price in quarter $t$. Because our time horizon is one quarter (i.e., $T = 1$), $\mathcal{N}(-NDD)$ is the probability of declaring bankruptcy within a quarter.

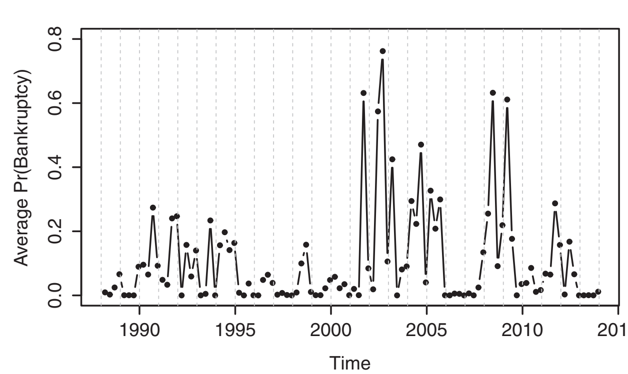
A complication arises when a parent corporation owns more than one airline in the sample because the association between an airline's operational performance and the parent company's financial distress may be weak. For instance, prior to its bankruptcy, AMR Corporation owned both American and American Eagle. It is unclear whether the smaller airline (i.e., American Eagle) had a significant impact on the parent company's financial distress. In our main model, we follow Phillips and Sertsios's (2013) approach. That is, when a corporation owns more than one airline in the sample, we use the parent company's *NDD* for all airlines owned by that company. In robustness tests, we define a smaller sample in which we only keep the largest airline the parent company owns and link the parent company's financial distress to its largest airline's operational performance. For instance, in our main model, we compute *NDD* for AMR and use it for both American and American Eagle. In robustness tests, we drop American Eagle from our dataset, and link AMR's *NDD* to American's operational performance. We adopt the same approach when two airlines merge or one airline buys another airline.

Figure 1 shows the average probability of bankruptcy implied by the naïve distance to default metric in the US airline industry over time. In order to generate this figure, we first compute the probability of bankruptcy for airline $i$ in quarter $t$ as $\mathcal{N}(-NDD_{it})$. Next, we take a simple average over all publicly traded airlines in quarter $t$. Despite only using stock market and debt data, this metric captures the negative impact of important events that have affected the airline industry. For example, spikes in the average probability of bankruptcy occur during the Gulf War, the terrorist attacks of September 11, 2001 (9/11), and the financial crisis in 2008.

### 3.3. Operational Performance Metrics

*Revenue Management*. We use yield to assess pricing. Yield is defined as passenger revenues divided by revenue passenger miles (RPMs). One RPM is flying one passenger over one mile in revenue service. We adjust yield for inflation, using the 2000 consumer price index, as we do for all other variables measured

**Figure 1 The Average Probability of Bankruptcy in the US Airline Industry**




*Notes.* We compute the probability of bankruptcy for airline $i$ in quarter $t$ using Bharath and Shumway's (2008) naïve distance to default metric. We compute the average probability of bankruptcy in quarter $t$ by taking a simple average over all publicly traded airlines in our dataset.

in dollars in our dataset. We use load factor, the standard measure in the industry, for utilization of effective capacity. Load factor captures the degree to which an airline fills up planes with revenue paying passengers controlling for both seats and miles flown. Load factor is defined as revenue passenger miles divided by available seat miles (ASMs). One ASM is flying one seat over one mile available for revenue service.

*Operational Efficiency*. We use fleet utilization to assess effective capacity of planes. Fleet utilization is the percent of time planes are used transporting passengers between departure at the gate of origin and arrival at the gate of destination. We use fuel efficiency to assess raw materials usage. Fuel efficiency is defined as ASMs divided by gallons of fuel used.

*Service Quality*. We use on-time performance and mishandled bags to measure the service quality of an airline. Following DOT's definitions, we measure on-time performance as the proportion of flights arriving less than 15 minutes after the scheduled arrival time (Ramdas et al. 2013) and mishandled bags as the number of mishandled baggage reports per 1000 passengers (Lapré 2011). Airlines receive these reports from passengers concerning lost, damaged, delayed, or pilfered baggage.

*Extreme Service Failures*. We measure extreme service failures using long delays and propensity to complain. Following Ramdas et al. (2013), we measure long delays as the percentage of flights that are 120 or more minutes late. Following Lapré (2011), we define propensity to complain as the number of baggage complaints filed with DOT per 1000 mishandled baggage reports.

*Operational Complexity*. To capture complexity associated with managing a diverse fleet, we use fleet heterogeneity, measured with the Blau index (Staats and Gino 2012) as one minus the sum of squared proportions of different aircraft types within an airline's fleet. For instance, if an airline in a quarter has only two types of aircraft with a 40–60% split, then the airline's fleet heterogeneity is computed as $1 - (0.4^2 + 0.6^2) = 0.48$. To capture complexity associated with flying to primary airports, we use landing fees, measured as the average landing fee per landing. Landing fees are higher for primary airports (e.g., Chicago O'Hare) compared to secondary airports (e.g., Chicago Midway). Lastly, to measure complexity associated with concentrating flights in a few major hubs or key cities rather than operating a point-to-point network, we use network sparsity measured as the sum of squared proportions of flights originating from each airport in an airline's network. Airlines with sparse networks have higher network sparsity scores. For instance, if an airline in a quarter serves ten airports with an equal proportion of flights (i.e., 10%) originating from each airport, then the network sparsity variable equals $10 \times 0.1^2 = 0.1$. Conversely, if 50% of flights originate from a focus city, whereas the remaining 50% is equally distributed among the remaining nine airports, then the network sparsity variable equals $0.5^2 + 9 \times (0.5/9)^2 = 0.28$.

### 3.4. Summary Statistics

Table 3 provides an overview of the variables we use in our financial distress analysis, while Panel A in Table 4 reports summary statistics. Although we report 1395 firm-quarter observations in Table 2, we report summary statistics for the independent variables based on the 1267 firm-quarters we use in our regression analysis. In our preliminary analysis of airline bankruptcies presented in section 4 and financial distress analysis presented in section 5, we lose 100 firm-quarters (i.e., the first four quarters for each airline) because we lag independent variables by four quarters. We lag independent variables by four quarters for two reasons. First, some operational metrics (e.g., on-time performance) have quarterly

**Table 3 Variables Used in Financial Distress Analysis**

| Variable | Formula | Data source |
|---|---|---|
| NDD | Equation (2) in section 3.2 | Compustat and CRSP |
| Yield | Passenger revenues/revenue passenger miles | Form 41 |
| Load factor | Revenue passenger miles/available seat miles | Form 41 |
| Fleet utilization | Block hours/(24 × Aircraft days) | Form 41 |
| Fuel efficiency | Available seat miles/Gallons of fuel used | Form 41 |
| On-time performance | Proportion of flights arriving less than 15 minutes after the scheduled arrival time | Air Travel Consumer Reports |
| Mishandled bags | The number of mishandled baggage reports/Total number of passengers (1000s) | Air Travel Consumer Reports |
| Long delays | Proportion of flights arriving more than 120 minutes after the scheduled arrival time | DOT On-Time Performance Data |
| Propensity to complain | Number of baggage complaints filed with DOT/Number of mishandled baggage reports (1000s) | Air Travel Consumer Reports |
| Fleet heterogeneity | Blau index computed from the proportion of different aircraft types within an airline's fleet (raw data: aircraft days by equipment type) | Form 41 |
| Landing fees | Total landing fees/Total number of flights | Form 41 |
| Network sparsity | Sum of squared proportions of flights originating from each airport in an airline's network | DOT On-time Performance Data |

**Table 4 Summary Statistics and Correlation Matrix**

| | | Panel A | | | Panel B | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *N* | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | NDD | 1086 | 5.59 | 5.10 | 1 | | | | | | | | | | | |
| 2 | Yield | 1267 | 13.16 | 3.86 | 0.00 | 1 | | | | | | | | | | |
| 3 | Load factor | 1267 | 0.73 | 0.08 | −0.15 | −0.58 | 1 | | | | | | | | | |
| 4 | Fleet utilization | 1267 | 0.42 | 0.05 | 0.23 | −0.42 | 0.13 | 1 | | | | | | | | |
| 5 | Fuel Efficiency | 1267 | 18.38 | 3.52 | −0.14 | 0.70 | −0.40 | −0.60 | 1 | | | | | | | |
| 6 | On-time perf. | 1267 | 0.79 | 0.06 | 0.20 | −0.13 | 0.06 | −0.08 | −0.21 | 1 | | | | | | |
| 7 | Mishandled bags | 1267 | 5.50 | 2.56 | −0.15 | 0.59 | −0.30 | −0.32 | 0.63 | −0.42 | 1 | | | | | |
| 8 | Long delays | 1267 | 0.01 | 0.01 | −0.31 | −0.15 | 0.40 | 0.09 | 0.06 | −0.68 | 0.22 | 1 | | | | |
| 9 | Prop. to complain | 1267 | 0.43 | 0.48 | −0.17 | −0.16 | 0.02 | 0.00 | −0.15 | −0.04 | −0.16 | 0.01 | 1 | | | |
| 10 | Fleet heterogeneity | 1267 | 0.53 | 0.24 | −0.35 | 0.22 | −0.06 | −0.33 | 0.17 | −0.14 | 0.16 | 0.02 | 0.18 | 1 | | |
| 11 | Landing fees | 1267 | 270 | 151 | −0.17 | −0.39 | 0.24 | 0.20 | −0.39 | 0.02 | −0.35 | 0.04 | 0.48 | 0.25 | 1 | |
| 12 | Network sparsity | 1267 | 0.10 | 0.05 | −0.23 | −0.33 | 0.33 | 0.13 | −0.28 | 0.11 | −0.13 | 0.02 | 0.09 | −0.07 | 0.10 | 1 |

*Notes.* In Panel A, we present the number of observations, mean, and standard deviation (SD) for the naïve distance to default (*NDD*) and the independent variables we use in our financial distress analyses presented in section 5 and 6. Naïve distance to default, fleet heterogeneity, and network sparsity are scalars. Yield and *OR*/*ASM* are expressed in 2000 US cents per seat mile. Load factor, fleet utilization, on-time performance, and long delays are fractions. Fuel efficiency is expressed as miles per gallon. Mishandled bags are expressed as the number of mishandled baggage per 1000 customers. Propensity to complain is expressed as the number of baggage complaints with DOT per 1000 mishandled bags. Landing fees are expressed in 2000 US dollars per landing. In Panel B, we report the correlation between naïve distance to default and the independent variables lagged by four quarters used in our financial distress analyses presented in section 5 and 6. Because the independent variables shown in rows and columns 2–12 are lagged by four quarters, we report their contemporaneous pairwise correlations.

seasonality. Second, and more importantly, lagging operational variables by four quarters eliminates a potential look-ahead bias and provides sufficient time to airlines and regulators to take actions based on the early warning signals the operational variables generate. Moreover, section 6 shows that the operational variables convey useful information up-to eight quarters before the measurement of financial distress and thereby demonstrates the robustness of our findings with respect to the time lag between operational performance and future financial distress.

We lose 28 firm-quarters due to missing data. We use all 1395 − (100 + 28) = 1267 firm-quarter observations in our preliminary analysis presented in the next section. However, we can calculate *NDD* and the corresponding probability of bankruptcy for only 1047 firm-quarters because stock price information is unavailable in firm-quarters during which an airline was either privately owned or operating under bankruptcy protection. Panel B in Table 4 shows the correlations among our variables. Because we lag independent variables by four quarters in sections 4 and 5, we report the correlations between naïve distance to default and the four-quarter lagged values of the independent variables.

# 4. Preliminary Analysis of Airline Bankruptcies

In this section, we provide preliminary evidence regarding the role of operational performance in financial distress models. We test whether operational characteristics can improve model fit in bankruptcy classification analyses.

## 4.1. Methodology

Modeling corporate bankruptcies has been an active research domain in accounting, economics, and finance since the early 1960s. A large body of literature uses binary classification models to identify financial ratios that explain corporate bankruptcies. In his seminal work, Altman (1968) applied multiple discriminant analysis to a matched sample of bankrupt and non-bankrupt firms to identify financial ratios that capture a firm's financial status. Because there are two distinct groups (bankrupt vs. non-bankrupt), the classification analysis can be transformed into a discriminant function of the form $z_n = \beta x_n$, where $z_n$ is the $z$-score for the $n^{th}$ observation in the sample and $x_n$ is the corresponding vector of firm specific variables. The goal is to estimate the coefficient vector $\beta$ such that every firm is correctly identified as bankrupt or non-bankrupt based on its $z$-score.

Altman (1968) identified five variables that perform well in firm classification: working capital to total assets (WC/TA), retained earnings to total assets (RE/TA), earnings before interest and taxes to total assets (EBIT/TA), market value of equity to total liabilities (ME/TL), and sales to total assets (S/TA). In a subsequent study, Altman (1993) replaced ME/TL with book value of equity to total liabilities (BE/TL), which facilitates the computation of a $z$-score for firms without stock price data. We use the independent variables of this alternative model, known as the

$z'$-score private firm model, in our first base model because stock price data are unavailable for firm quarters during which the firm is in bankruptcy. Both $z$ and $z'$ score models are still widely used by academics and practitioners to assess a firm's financial distress (Altman and Hotchkiss 2010).

While Altman's $z$ and $z'$ scores are useful in terms of identifying the financial ratios that explain corporate bankruptcies, their model coefficients are typically specified using a sample of firms from different industries. Thus, they may not be able to capture dynamics that are specific to the airline industry. In order to overcome this hurdle, researchers have proposed variables that are specific to the airline industry. One airline industry specific model is the Airscore model (Chow et al. 1991). The independent variables of the Airscore model are market value of equity to total liabilities (ME/TL), interest expense to total liabilities (IE/TL), and operating revenues to total miles flown (OR/TMF). Although the first two variables of this model are generic, the last variable is industry specific, and serves as a proxy for an airline's revenue generation capabilities. We use the Airscore model's independent variables in our second base model after replacing $ME/TL$ with $BE/TL$ so that we can calculate an Airscore in the absence of stock price information.

We test four versions of the Altman's $z'$ score model in our preliminary analysis. Model 1 includes the Altman's $z'$ score model's independent variables, all computed from Compustat data. Model 2 expands the set of independent variables of the $z'$ score model by adding $OR/ASM$ (which is yield $\times$ load factor), mishandled bags, fleet utilization, and fleet heterogeneity. Model 3 expands Model 2 by adding a legacy dummy that takes a value of 1 if airline $i$ is a legacy carrier and 0 otherwise. Given that 14 out of 20 bankruptcy episodes presented in Table 2 belong to legacy airlines, Model 3 allows us to test whether operational variables are associated with bankruptcy classification rather than the higher financial distress of legacy carriers compared to low-cost carriers. Lastly, Model 4 has the same independent variables as Model 3, but excludes Southwest. Comparing Model 4 with Model 3 allows us to check whether our findings are driven by the superior operational and financial performance of Southwest. We follow a similar approach with four versions of the Airscore model: Models 5–8. We compute the Airscore model's independent variables $BE/TL$ and $IE/TL$ from Compustat data and $OR/TMF$ from Form 41 data.

The discriminant analysis approach can be considered as a single period model that uses only one observation from each firm (typically the last observation in the data). Shumway (2001) demonstrates that this approach leads to biased and inconsistent coefficient estimates due to its static nature, overlooking data on healthy firms that eventually go bankrupt. A more appropriate approach is a duration model (also known as a hazard model), which captures the dynamic nature of bankruptcy risk. Shumway (2001) and Chava and Jarrow (2004) show that the likelihood function of a discrete-time hazard model is identical to that of a multi-period logit model and that a discrete-time hazard model can be estimated via logistic regression. We adopt the same approach and seek to classify firm-quarter observations via logistic regression. In the logit model, $Pr(Bankrupt_{it} = 1) = \frac{1}{1+\exp(-\beta X_{i,t-4})}$, where $X_{i,t-4}$ denotes the vector of independent variables lagged by four quarters.

### 4.2. Results

Table 5 shows the results of our preliminary bankruptcy analysis. All independent variables except $WC/TA$ are significant at $p = 0.05$ in Model 1. Consistent with Altman's findings, we find that lower $RE/TA$, $EBIT/TA$, $BE/TL$, and $S/TA$ are associated with higher future financial distress. The insignificance of $WC/TA$ is consistent with Shumway (2001), who documents the insignificance of this variable using a broader sample of firms.

A likelihood ratio test between Models 1 and 2 has a $p$-value of 0.002. So, adding revenue management, operational efficiency, service quality, and operational complexity to the Altman's $z'$ score model leads to a statistically significant improvement in model fit. Model 2 indicates that lower operating revenue per ASM, lower fleet utilization, and higher fleet heterogeneity are associated with higher future financial distress. Despite having the correct sign, mishandled bags is statistically insignificant. Replacing mishandled bags with propensity to complain does not change our results. A likelihood ratio test between Models 2 and 3 has a $p$-value of 0.247, indicating that controlling for the low-cost–legacy difference does not improve model fit. A comparison of the estimates of Models 2 and 3 reveals that the addition of the legacy fixed effect does not change the statistical significance (or insignificance) of Model 2's independent variables. Moreover, the legacy fixed effect is statistically insignificant in Model 3. The estimates of Model 4, which excludes Southwest, are qualitatively similar to Model 3. Estimating the Airscore model leads to similar insights (Models 5–8).

Our preliminary analysis shows that operational performance metrics can improve goodness of fit in bankruptcy classification models and that this improvement is not driven by the low-cost–legacy difference or a Southwest effect. In particular, we find support for Hypotheses 1, 2, and 4, whereas we do

**Table 5 Logistic Regression Results**

| | Altman's $z'$ model | | | | Airscore model | | | |
|---|---|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
| (Intercept) | −1.726*** | 1.910 | 2.003 | 2.071 | −0.718 | 1.601 | 1.786 | 1.789 |
| | (0.292) | (1.704) | (1.704) | (1.703) | (0.541) | (1.362) | (1.366) | (1.364) |
| *WC/TA* | 0.696 | 0.965 | 0.899 | 0.827 | | | | |
| | (0.833) | (0.938) | (0.935) | (0.942) | | | | |
| *RE/TA* | −3.289*** | −3.596*** | −3.692*** | −3.703*** | | | | |
| | (0.589) | (0.609) | (0.624) | (0.623) | | | | |
| *EBIT/TA* | −17.940*** | −12.905** | −13.384** | −13.321** | | | | |
| | (4.133) | (4.526) | (4.532) | (4.523) | | | | |
| *BE/TL* | −1.396* | −0.922 | −0.974 | −0.910 | −3.676*** | −3.508*** | −3.397*** | −3.378*** |
| | (0.578) | (0.571) | (0.575) | (0.584) | (0.390) | (0.402) | (0.412) | (0.415) |
| *S/TA* | −2.837** | −3.231** | −3.482** | −3.523** | | | | |
| | (0.995) | (1.189) | (1.221) | (1.222) | | | | |
| *OR/ASM* | | −0.237*** | −0.230** | −0.230** | | | | |
| | | (0.071) | (0.071) | (0.071) | | | | |
| Fleet utilization | | −6.417* | −6.727* | −6.716* | | −6.080* | −5.712* | −5.680* |
| | | (3.019) | (3.017) | (3.002) | | (2.788) | (2.806) | (2.802) |
| Mishandled bags | | 0.065 | 0.068 | 0.066 | | 0.063 | 0.065 | 0.065 |
| | | (0.054) | (0.053) | (0.053) | | (0.050) | (0.051) | (0.051) |
| Fleet heterogeneity | | 1.353* | 1.766* | 1.662* | | 3.175*** | 2.713*** | 2.680*** |
| | | (0.680) | (0.776) | (0.791) | | (0.688) | (0.779) | (0.784) |
| *IE/TL* | | | | | 45.343* | 37.248 | 33.891 | 33.700 |
| | | | | | (21.742) | (22.728) | (23.035) | (23.029) |
| *OR/TMF* | | | | | −1.032** | −2.463*** | −2.698*** | −2.693*** |
| | | | | | (0.356) | (0.455) | (0.499) | (0.499) |
| Legacy | | | −0.359 | −0.348 | | | 0.405 | 0.410 |
| | | | (0.308) | (0.308) | | | (0.339) | (0.339) |
| # of airlines | 25 | 25 | 25 | 24 | 25 | 25 | 25 | 24 |
| # of observations | 1267 | 1267 | 1267 | 1167 | 1267 | 1267 | 1267 | 1167 |
| # of bankrupt qtrs. | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 |
| # of non-bankrupt qtrs. | 1147 | 1147 | 1147 | 947 | 1147 | 1147 | 1147 | 947 |
| AIC | 620.128 | 611.310 | 611.969 | 611.340 | 657.502 | 634.738 | 635.281 | 635.070 |
| BIC | 743.594 | 817.086 | 838.323 | 834.077 | 739.813 | 778.782 | 799.902 | 797.061 |
| Pseudo $R^2$ | 0.293 | 0.317 | 0.319 | 0.301 | 0.231 | 0.264 | 0.277 | 0.255 |
| Log likelihood | −304.064 | −295.655 | −294.984 | −294.670 | −324.751 | −310.369 | −309.641 | −309.535 |
| Likelihood ratio test | | 16.819 ($p$ = 0.002) | | | | 28.764 ($p < 10^{-3}$) | | |
| Likelihood ratio test | | | 1.341 ($p$ = 0.247) | | | | 1.457 ($p$ = 0.227) | |

*Notes.* Standard errors in parentheses. *, **, and *** indicate statistical significance at 0.05, 0.01, and $p$ = 0.001, respectively. The dependent variable is $Bankrupt_{it}$. All explanatory variables except the legacy dummy are lagged by four quarters. We perform parameter estimation via logistic regression. In Models 4 and 8, we drop the Southwest observations.

not find a link between service quality (Hypothesis 3A) or extreme service failures (Hypothesis 3B) and bankruptcy in subsequent quarters.

### 4.3. Limitations

Despite their widespread use in the airline bankruptcy literature (e.g., Chow et al. 1991, Gritta et al. 2008, Lu et al. 2015), binary classification models have some significant limitations. First, bankruptcy is an extreme form of financial distress. By labeling airlines as bankrupt vs. non-bankrupt, binary classification models may not distinguish between healthy airlines and airlines that are on the verge of bankruptcy. Moreover, out-of-sample evidence is necessary to establish the empirical reliability of a forecasting model (Rapach et al. 2010). However, the binary classification models in the airline industry rely on a small number of bankruptcy filings, which prevents researchers from testing the out-of-sample forecast accuracy. Consequently, inferences from binary classification models regarding the association between operational performance and future financial distress may be inaccurate.

Second, a general rule of thumb in logistic regression is to have at least 10 events per parameter in order to have reliable parameter estimates (Hosmer and Lemeshow 2004). Because we only have 120 bankrupt quarters in our logistic regression sample, we should have fewer than 12 independent variables. Therefore, we attempt to capture the association between each area of operational performance and future financial distress with only variable (e.g., *OR/ASM* to proxy revenue management) rather than using multiple variables (e.g., yield and load factor).

Furthermore, we cannot introduce parameters to capture airline or time specific effects that may influence financial distress. Not being able to incorporate fixed effects impedes the reliability of inferences in industry-specific studies (Joglekar et al. 2016). For instance, we do not find support for Hypothesis 3B (extreme service failures) in our preliminary analysis of airline bankruptcies, but we do find support for this hypothesis in the next section, where we control for airline and time specific effects.

Lastly, as discussed in Phillips and Sertsios (2013), the weights assigned to each financial ratio (i.e., the estimated coefficients of each financial ratio) can be quite unstable in a binary classification model because of industry trends that systematically change those ratios. As a result, the estimated $z$–scores can be unreliable. In the next section, we overcome these hurdles by replacing the binary classification with a continuous financial distress metric. Consequently, we can (i) distinguish between healthy airlines and financially distressed but non-bankrupt airlines, (ii) jointly test the significance of a relatively large number of independent variables, and (iii) perform out-of-sample forecasts.

## 5. Financial Distress Analysis

In this section, we examine the association between operational performance and future financial distress. Section 5.1 explains our methodology, and section 5.2 discusses our results and robustness of our findings.

### 5.1. Methodology

Our model specification has the following form:

$$-NDD_{it} = \alpha_i + \gamma_t + \sum_k \beta_k x_{k,i,t-4} + \theta\left(-NDD_{i,t-4}\right) + u_{it}, \tag{3}$$

where $\alpha_i$ is the airline dummy, $\gamma_t$ is the time dummy, $x_{k,i,t-4}$ are the lagged values of operational variables, and $u_{it}$ is the error term. We use $-1 \times NDD$ (rather than $NDD$) as our dependent variable to ensure that the estimated coefficient signs are consistent with the ones we obtained in section 4.

Many economic and financial variables dynamically evolve over time. In empirical settings, dynamic relationships are captured by the presence of a lagged dependent variable among the independent variables. (See Baltagi (2008, ch. 8) for an overview of dynamic panel data models and their applications in economics.) Accordingly, we use $-NDD_{i,t-4}$ as an independent variable to capture the dynamic nature of financial distress. Another reason to use $-NDD_{i,t-4}$ as an independent variable is a potential link between an airline's financial distress state and operational performance. A financially distressed airline could also anticipate financial distress in future periods and thereby take actions that could influence our operational predictors. For instance, a financially distressed airline could lower its ticket prices, reduce its capacity, downsize its flight networks, and/or lower its service quality to improve its short-term liquidity. Including $-NDD_{i,t-4}$ allows us to test whether operational performance is associated with future financial distress after controlling for an airline's financial distress state at the time we measure operational performance.

We test seven models. Model 1 is the base model with airline and time dummies. Model 2 expands Model 1 by adding yield, load factor, on-time performance, mishandled bags, fleet utilization, fuel efficiency, fleet heterogeneity, landing fees, and network sparsity. Model 3 expands Model 2 by adding $NDD_{i,t-4}$. To test whether our findings are driven by the superior operational and financial performance of Southwest, Model 4 has the same independent variables as Model 3, but excludes Southwest. In Models 5–7, we replace the service quality variables in Models 2–4 (on-time performance and mishandled bags) with the extreme service failure variables (long delays and propensity to complain).

Models 1, 2, 3, 5, and 6 have 25 airlines, whereas Models 4 and 7 have 24 airlines due to the exclusion of Southwest. The number of quarters per airline varies from 3 to 100 in all seven models. Thus, we have unbalanced time-series cross-section (TSCS) data, which is different from panel data (Beck 2001). In panel data, the asymptotics are in the number of units. In TSCS data, the units are fixed, there is no sampling, and we are interested in specific units. The asymptotics are in time periods. Consequently, the estimation method needs to deal with panel heteroskedasticity ($E[u_{it}^2] = \sigma_i^2 \neq E[u_{jt}^2] = \sigma_j^2$ for $i \neq j$), contemporaneous correlation ($E[u_{it}u_{jt}] = \sigma_{ij}$), and autocorrelation ($u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}$).[2] Following Beck and Katz (1995), we use Prais–Winstein regression with a single first-order autoregressive process common to all airlines and estimate panel-corrected standard errors, which account for panel heteroskedasticity and contemporaneous correlation.[3] Other airline industry studies that use our estimation methodology include Lapré and Tsikriktsis (2006) and Lapré (2011).

### 5.2. Results

Table 6 shows the regression results. Models 1 explains 57.6% of the variation in financial distress, whereas Model 2 explains 61.3%. Both revenue management variables, yield and load factor, are negatively associated with future financial distress.

**Table 6 Financial Distress Regression Results**

| | | Average service quality | | | Extreme service failures | | |
|---|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
| Yield | | −0.314*** | −0.277*** | −0.240*** | −0.311*** | −0.275*** | −0.241*** |
| | | (0.072) | (0.067) | (0.065) | (0.071) | (0.066) | (0.065) |
| Load factor | | −10.569* | −9.354* | −11.884** | −11.559** | −9.967* | −12.320** |
| | | (4.484) | (4.387) | (4.256) | (4.408) | (4.314) | (4.158) |
| Fleet utilization | | −12.071* | −12.766** | −10.897* | −12.954* | −13.373** | −11.024* |
| | | (5.169) | (4.895) | (4.791) | (5.127) | (4.859) | (4.756) |
| Fuel efficiency | | 0.084 | 0.070 | 0.039 | 0.050 | 0.043 | 0.015 |
| | | (0.081) | (0.077) | (0.078) | (0.081) | (0.078) | (0.079) |
| On-time perf. | | −0.029 | −0.028 | −0.026 | | | |
| | | (0.026) | (0.026) | (0.025) | | | |
| Mishandled bags | | −0.161 | −0.155 | −0.137 | | | |
| | | (0.097) | (0.091) | (0.092) | | | |
| Long delays | | | | | 48.528** | 41.955* | 40.837* |
| | | | | | (17.851) | (17.820) | (17.821) |
| Prop. to complain | | | | | 1.287** | 1.208** | 1.253** |
| | | | | | (0.407) | (0.402) | (0.392) |
| Fleet heterogeneity | | 4.294** | 3.473** | 3.574** | 3.679* | 2.937* | 2.989* |
| | | (1.450) | (1.309) | (1.331) | (1.432) | (1.307) | (1.330) |
| Landing fees | | 0.011*** | 0.010** | 0.007* | 0.011*** | 0.009** | 0.007* |
| | | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Network sparsity | | 27.450** | 24.706** | 16.233* | 26.495** | 23.649** | 15.673* |
| | | (8.441) | (7.683) | (7.750) | (8.284) | (7.622) | (7.766) |
| Lagged NDD | | | 0.190*** | 0.177*** | | 0.180*** | 0.166*** |
| | | | (0.038) | (0.038) | | (0.038) | (0.038) |
| # of airlines | 25 | 25 | 25 | 24 | 25 | 25 | 24 |
| # of observations | 1047 | 1047 | 1047 | 947 | 1047 | 1047 | 947 |
| $R^2$ | 0.576 | 0.613 | 0.648 | 0.600 | 0.624 | 0.656 | 0.608 |
| Autocor. coeff., $\rho$ | 0.512 | 0.437 | 0.360 | 0.369 | 0.419 | 0.351 | 0.361 |

*Notes.* The dependent variable is $-NDD_{it}$. All independent variables are lagged by four quarters. Because $-NDD_{i,t-4}$ is unavailable in 39 firm-quarters, Models 1, 2, 3, 5, and 6 have 1086−39 = 1047 observations. Because Models 4 and 7 exclude Southwest, they have 24 airlines and 947 observations. All models have airline and time dummies. We do not report the coefficient estimates for airline and quarter dummies due to space limitations. We perform parameter estimation via Prais-Winsten regression with panel-corrected standard errors and a single first-order autoregressive process common to all airlines, i.e., $u_{it} = \rho u_{i,t-1} + \varepsilon_{it}$. Panel-corrected standard errors, which correct for panel heteroskedasticity and contemporaneous correlation, are in parentheses. *, **, and *** indicate statistical significance at 0.05, 0.01, and 0.001, respectively.

Among the operational efficiency variables, fleet utilization is negative and significant, whereas fuel efficiency is statistically insignificant. Both service quality variables, on-time performance and mishandled bags, are statistically insignificant at $p = 0.05$. All three operational complexity variables (fleet heterogeneity, landing fees, and network sparsity) are statistically significant with the expected signs. Model 3 explains 64.8% of variation in financial distress. The lagged value of *NDD* is significant indicating that the current financial distress state is an important determinant of the future distress state. The addition of the lagged *NDD* in Model 3 does not change the significance findings for the operational variables in Model 2. The coefficient estimates of Model 4, which excludes Southwest, are qualitatively similar to those of Model 3 indicating that our findings are not driven by the superior operational and financial performance of Southwest.

Model 5 explains 62.4% of the variation in financial distress. Unlike service quality variables, extreme service failure variables are significant. Both long delays and propensity to complain are positively associated with future financial distress. Replacing service quality with extreme service failures does not change the significance findings for the other operational variables in Model 2. Model 6 explains 65.6% of the variation in financial distress. Lagged *NDD* is significant. The addition of the lagged *NDD* does not change the significance findings in Model 5 for the operational variables. Lastly, the estimates of Model 7, which excludes Southwest, are qualitatively similar to Model 6.

In sum, our findings from Models 1–7 are consistent. We find support for Hypothesis 1, i.e., lower values of yield and load factor are associated with higher future financial distress. We find partial support for Hypothesis 2, as low fleet utilization is associated with higher future financial distress, whereas fuel efficiency is not significant. The insignificance of fuel efficiency is primarily driven by the lack of contemporaneous heterogeneity among airlines. In a given quarter, there is very little variation among airlines in terms of their fuel efficiencies. We do not find support for Hypothesis 3A. Service quality, measured by on-time performance and mishandled bags, does not have a significant association with future financial distress. However, we do find support for Hypothesis 3B. Extreme service failures are associated with higher future financial distress. Lastly, we find support for Hypothesis 4 as fleet heterogeneity, landing fees, and network sparsity are associated with higher future financial distress.

We conducted several tests to assess the robustness of our findings. Recall from section 3 that when a corporation owns more than one airline in the sample, we use the parent company's *NDD* for all airlines owned by that company. We test the robustness of our findings by replicating Models 1–7 on a smaller sample in which we only keep the largest airline owned by the parent company. Our findings regarding the associations between four operational dimensions and future financial distress remained unchanged in the smaller sample. In addition, we tested different versions of our models with only one variable for each operational dimension. For instance, we used either on-time performance or mishandled bags to measure service quality. The impact of revenue management, operational efficiency, service quality (or extreme service failures), and operational complexity remained qualitatively unchanged under those alternative specifications. Furthermore, we introduced additional variables (fuel cost, labor cost, market share) to control for an airline's cost structure, size, and pricing power. These variables were insignificant. More importantly, they did not change the associations between operational variables and future financial distress. Several studies (e.g., Hofer et al. 2009, Phillips and Sertsios 2013) document contemporaneous endogeneity between revenue management variables (yield and load factor) and financial distress. We checked the robustness of our coefficient estimates by estimating two-stage regressions. We used five-quarter lagged values of yield as an instrument for four-quarter lagged values of yield and verified through Durbin–Wu–Hausman tests that the coefficient estimates reported in Table 6 do not suffer from endogeneity. We reached the same conclusion using this approach for load factor. See Table 7.

Unobservable shocks may have a differential impact on the low-cost and legacy carriers' financial distress. Although we cannot introduce a legacy fixed effect due to the presence of airline dummies, we can control for the potential differences between the low-cost and legacy carriers by fitting separate time fixed effects. Fitting separate time fixed effects for the low-cost and legacy carriers did not change the associations between operational variables and future financial distress. We also tested whether the inclusion of airlines with short time series affected our results. Dropping American Trans Air, which had 12 firm-quarters in

**Table 7 Endogeneity Tests**

| | Revenue management variables | | | | | |
|---|---|---|---|---|---|---|
| | Yield | | | Load factor | | |
| Model in Table 6 | First stage *F* test | Durbin–Wu–Hausman $\chi^2$ test | Wu–Hausman *F* test | First stage *F* test | Durbin–Wu–Hausman $\chi^2$ test | Wu–Hausman *F* test |
| Model 2 | 2040.770 | 1.747 | 1.065 | 366.130 | 2.220 | 1.968 |
| | | (0.186) | (0.302) | | (0.136) | (0.161) |
| Model 3 | 2006.270 | 1.370 | 0.821 | 359.684 | 1.653 | 1.440 |
| | | (0.242) | (0.365) | | (0.198) | (0.230) |
| Model 4 | 1928.090 | 1.347 | 0.778 | 355.040 | 1.824 | 1.577 |
| | | (0.246) | (0.378) | | (0.177) | (0.210) |
| Model 5 | 2095.692 | 1.432 | 0.950 | 373.685 | 1.862 | 1.704 |
| | | (0.231) | (0.330) | | (0.172) | (0.192) |
| Model 6 | 2060.897 | 1.050 | 0.684 | 368.056 | 1.586 | 1.425 |
| | | (0.305) | (0.408) | | (0.208) | (0.233) |
| Model 7 | 1977.870 | 1.003 | 0.644 | 362.067 | 1.715 | 1.550 |
| | | (0.317) | (0.423) | | (0.190) | (0.213) |

*Notes.* We check the endogeneity of yield in Model 2 in Table 6 as follows. First, we estimate a two-stage regression model, where we used $Yield_{i,t-5}$ as an instrument for $Yield_{i,t-4}$ in the first stage. Second, we check the relevance of the instrument using an *F* test. The *F* statistic of 2040.770 is much higher than the rule of thumb threshold of 10 indicating that we have a valid instrument. Third, we use the Durbin–Wu–Hausman and Wu–Hausman tests to compare our original coefficient estimates with the two-stage estimates. The results show that we do not reject the null hypotheses. Thus, we conclude that the coefficient estimates we report in Table 6 do not suffer from endogeneity. We repeat the same steps for all model (Models 2–7 in Table 6) – revenue management variable (yield and load factor) combinations. We report the test statistics for the first stage *F*-test, Durbin-Wu-Hausman test, and Wu–Hausman test for each model–revenue management variable combination. *p* values for the Durbin-Wu-Hausman and Wu–Hausman tests are in parentheses.

our models after dropping the first four quarters due to lagging, Independence Air (8 firm-quarters), and Piedmont Aviation (3 firm-quarters) did not change our findings. Lastly, in case an airline declared bankruptcy, exited bankruptcy, and started trading later, we fitted separate firm dummies for pre- and post-bankruptcy quarters, which allows us to treat the pre- and post-bankruptcy episodes as two separate firms. Once again, our findings remained unchanged.

# 6. Out-of-Sample Forecasts of Future Financial Distress

In this section, we switch our focus from testing associations to generating out-of-sample forecasts to investigate whether operational performance can be used to predict future financial distress. Section 6.1 explains our methodology, and section 6.2 presents our findings.

## 6.1. Methodology

We generate out-of-sample forecasts of *NDD* using a recursive estimation window for different lags. Recursive forecasting models are commonly used in economics and finance to test the predictive ability of a variable of interest (e.g., Goyal and Welch 2003, Rapach et al. 2010, and references therein). To test of the robustness of our findings, we perform sensitivity analyses with different lag periods.

In order to generate out-of-sample forecasts, we first divide our study period of 104 quarters into an in-sample portion consisting of the first 40 quarters (from the first quarter of 1988 until the last quarter of 1997) and an out-of-sample portion consisting of the last 64 quarters (from the first quarter of 1998 until the last quarter of 2013). We generate out-of-sample forecasts for the first quarter of 1998 (i.e., $t = 41$) as follows. Let $l$ denote the time lag between the operational variables and future financial distress. For $l = 1, \cdots, 8$, we use data from the first 40 quarters and the estimation methodology described in section 5 to estimate the coefficients of the model specification described as Model 6 in section 5 as

$$-NDD_{is} = \alpha_i^{[OM,40,l]} + \gamma_s^{[OM,40,l]} + \sum_k \beta_k^{[OM,40,l]} x_{k,i,s-l} - \theta^{[OM,40,l]} NDD_{i,s-l} + u_{is}^{[OM,40,l]}, \tag{4}$$

where $s \in \{5, \cdots, 40\}$ and the superscript $[OM, 40, l]$ indicates that we estimate the coefficients of an operational performance model (*OM*) using data from the first 40 quarters. Then, for $l = 1, \cdots, 8$, we predict naïve distance to default for airline $i$ at $t = 41$, $\widehat{NDD}_{i,41}^{[OM,l]}$, as

$$\widehat{NDD}_{i,41}^{[OM,l]} = -\left(\widehat{\alpha}_i^{[OM,40,l]} + \widehat{\gamma}_{41-l}^{[OM,40,l]} + \sum_k \widehat{\beta}_k^{[OM,40,l]} x_{k,i,41-l} - \widehat{\theta}^{[OM,40,l]} NDD_{i,41-l}\right), \tag{5}$$

where $\widehat{\alpha}_i^{[OM,40,l]}$, $\widehat{\gamma}_{41-l}^{[OM,40,l]}$, $\widehat{\beta}_k^{[OM,40,l]}$, and $\widehat{\theta}^{[OM,40,l]}$ denote the coefficient estimates. Because we estimate model coefficients of equation (4) using data from the first 40 quarters, and because the right-hand side of equation (5) has the lagged values of independent variables, $\widehat{NDD}_{i,41}^{[OM,l]}$ is an out-of-sample forecast.

We generate out-of-sample forecasts for the remaining quarters of the out-of-sample portion of our data in a recursive manner. Specifically, in order to generate forecasts for quarter $t = 41, \cdots, 104$, we first estimate the coefficients of the model specified in equation (4) using data up to quarter $t - 1$. We then use the estimated model coefficients, $\widehat{\alpha}_i^{[OM,t-1,l]}$, $\widehat{\gamma}_{t-l}^{[OM,t-1,l]}$, $\widehat{\beta}_k^{[OM,t-1,l]}$, and $\widehat{\theta}^{[OM,t-1,l]}$, and calculate a one-step ahead forecast for airline $i$ at lag $l$ and time $t$ as

$$\widehat{NDD}_{i,t}^{[OM,l]} = -\left(\widehat{\alpha}_i^{[OM,t-1,l]} + \widehat{\gamma}_{t-l}^{[OM,t-1,l]} + \sum_k \widehat{\beta}_k^{[OM,t-1,l]} x_{k,i,t-l} - \widehat{\theta}^{[OM,t-1,l]} NDD_{i,t-l}\right). \tag{6}$$

Given the observed distance to default, $NDD_{it}$, and the corresponding forecast, $\widehat{NDD}_{it}^{[OM,l]}$, we calculate the forecast error of our model for airline $i$ at lag $l$ and time $t$ as

$$e_{it}^{[OM,l]} = \mathcal{N}(-NDD_{it}) - \mathcal{N}\left(-\widehat{NDD}_{it}^{[OM,l]}\right). \tag{7}$$

We calculate forecast errors based on the observed and predicted bankruptcy probabilities (rather than the observed and predicted distance to default values) because forecast errors are easier to interpret in the probability domain as they lie in the interval $[-1, 1]$. For each lag ($l = 1, \cdots, 8$), we compare the performance of our forecasting model with a benchmark model.

Our benchmark model combines the Altman variables with the lagged value of *NDD*. Using data up to period $t - 1$, we first fit the following benchmark model (BM):

$$-NDD_{is} = \alpha_i^{[BM,t-1,l]} + \gamma_s^{[BM,t-1,l]} + \sum_{k=1}^{5} \beta_k^{[BM,t-1,l]} x_{k,i,s-l} - \theta^{[BM,t-1,l]} NDD_{i,s-l} + u_{is}^{[BM,t-l]}, \quad (8)$$

where $x_{1,i,s-l}, \cdots, x_{5,i,s-l}$ denote the $l$-quarter lagged values of $WC/TA$, $RE/TA$, $EBIT/TA$, $BE/TL$, and $S/TA$ for airline $i$, respectively. Then, we calculate our forecast as

$$\widehat{NDD}_{i,t}^{[BM,l]} = -\left(\widehat{\alpha}_i^{[BM,t-1,l]} + \widehat{\gamma}_{t-l}^{[BM,t-1,l]} + \sum_{k=1}^{5} \widehat{\beta}_k^{[BM,t-1,l]} x_{k,i,t-l} - \widehat{\theta}^{[BM,t-1,l]} NDD_{i,t-l}\right), \quad (9)$$

where $\widehat{\alpha}_i^{[BM,t-1,l]}$, $\widehat{\gamma}_{t-l}^{[BM,t-1,l]}$, $\widehat{\beta}_k^{[BM,t-1,l]}$, and $\widehat{\theta}^{[BM,t-1,l]}$ are the coefficients estimated using data up to $t-1$. For lag $l = 1, \cdots, 8$, the forecast error of the benchmark model for airline $i$ at time $t$ is $e_{it}^{[BM,l]} = \mathcal{N}(-NDD_{it}) - \mathcal{N}\left(-\widehat{NDD}_{it}^{[BM,l]}\right)$.

### 6.2. Results

We use absolute forecast errors (i.e., absolute deviations) and squared errors, which take values between 0 and 1, to assess forecast accuracy. Panel A in Table 8 reports results based on the absolute deviations of our model as well as the benchmark model for each lag. Due to data availability, the number of observations decreases as the time lag increases. For instance, we generate forecasts for 818 firm-quarters when $l = 1$ and 703 firm-quarters when $l = 8$.

To formally compare our model with the benchmark model, we follow Kesavan's et al. (2010) approach of assessing forecast accuracies with three statistical tests. First, we test whether our model gives a lower mean absolute deviation (MAD) than the benchmark model using one tailed $t$-tests. The differences between our model's MAD and the benchmark model's MAD for lags 1, ⋯, 8 are all statistically lower than zero at $p = 0.001$. Second, we use Johnson's skewness-adjusted $t$-test (Johnson 1978) to ensure that a potential skewness in the difference between absolute deviations obtained from two different forecasting models does not influence our findings. The skewness-adjusted $t$-statistics indicate that our model's MAD is statistically lower than that of the benchmark model for $l = 1, \cdots, 8$. Third, we test if our model performs better than the benchmark model

**Table 8 Comparison of Our Model's Forecast Accuracy Against a Benchmark Model**

| | Lags | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *Panel A: Absolute Deviations* | | | | | | | | |
| No. of observations | 818 | 798 | 779 | 761 | 743 | 727 | 714 | 703 |
| Our model's MAD | 0.131 | 0.129 | 0.131 | 0.128 | 0.125 | 0.141 | 0.154 | 0.118 |
| Benchmark model's MAD | 0.173 | 0.162 | 0.164 | 0.162 | 0.184 | 0.174 | 0.213 | 0.181 |
| $\Delta MAD$ | −0.042 | −0.033 | −0.034 | −0.034 | −0.059 | −0.033 | −0.059 | −0.062 |
| $t$-stat | −4.614*** | −3.876*** | −3.821*** | −3.820*** | −6.887*** | −4.498*** | −7.635*** | −6.698*** |
| Skewness adjusted $t$-stat | −4.929*** | −4.087*** | −4.065*** | −4.030*** | −8.230*** | −4.842*** | −9.376*** | −7.632*** |
| % of observations where our model gives lower absolute deviations | 70.05 | 65.79 | 65.47 | 68.20 | 67.97 | 67.81 | 73.67 | 74.96 |
| $p$-value for binom. sign test | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |
| *Panel B: Squared Errors* | | | | | | | | |
| No. of observations | 818 | 798 | 779 | 761 | 743 | 727 | 714 | 703 |
| Our model's MSE | 0.093 | 0.094 | 0.096 | 0.094 | 0.092 | 0.095 | 0.108 | 0.084 |
| Benchmark model's MSE | 0.119 | 0.111 | 0.114 | 0.112 | 0.133 | 0.121 | 0.160 | 0.131 |
| $\Delta MSE$ | −0.026 | −0.017 | −0.018 | −0.018 | −0.041 | −0.027 | −0.052 | −0.047 |
| $t$-stat | −2.932** | −2.144* | −2.251* | −2.231* | −5.189*** | −3.712*** | −6.972*** | −5.548*** |
| Skewness adjusted $t$-stat | −3.037** | −2.195* | −2.329* | −2.297* | −5.963*** | −3.971*** | −8.624*** | −6.272*** |
| % of observations where our model gives lower squared errors | 73.84 | 72.68 | 72.27 | 73.19 | 75.64 | 76.20 | 78.01 | 79.52 |
| $p$-value for binom. sign test | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

*Notes.* In Panel A, we use absolute deviations to compare our model's out-of-sample forecast accuracy with that of the benchmark model. The $\Delta MAD$ row reports the difference between the mean absolute deviations (MADs) obtained from our forecasting model and the benchmark model. In Panel B, we use squared errors to compare our model's out-of-sample forecast accuracy with that of the benchmark model. The $\Delta MSE$ row reports the difference between the mean squared errors (MSEs) obtained from our forecasting model and the benchmark model. In both panels, we report the $t$ statistic of a one tailed $t$-test. We also report the $t$ statistic obtained from a skewness adjusted $t$-test to ensure that a potential skewness does not affect our findings. Lastly, we use the binomial sign test to test if the benchmark model is more likely to produce absolute deviations that are lower than the ones our model generates. *, **, and *** indicate statistical significance at 0.05, 0.01, and 0.001, respectively.

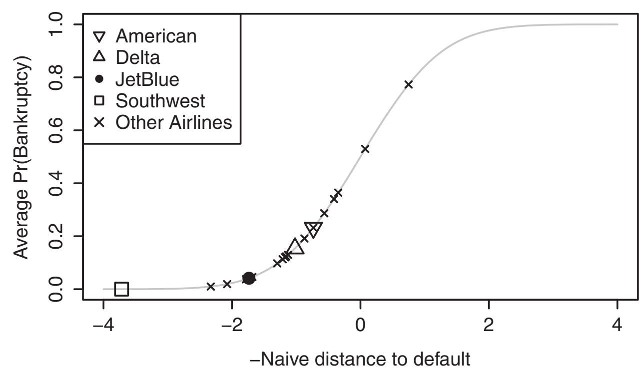
for more than half of the forecasts (e.g., if it gives a lower absolute deviation for more than 818/2 = 409 data points when $l = 1$). We perform this analysis using a non-parametric binomial sign test. Our model generates lower absolute deviations than the benchmark model. The difference from 50% is statistically significant at $p = 0.001$ for all $l = 1, \cdots, 8$. Panel B in Table 8 reports results based on the squared errors of our model as well as the benchmark model for each lag, $l = 1, \ldots, 8$. Performing forecast comparisons between our model and the benchmark model based on squared errors does not change our findings.

Our analyses indicate that our model is more likely to generate smaller forecast errors than the benchmark model, on average. Similarly, our non-parametric analyses show that for a randomly picked firm-quarter observation, our model is more likely to generate a more accurate forecast than the benchmark model. These results are robust to the time lag between predictive variables and future financial distress. Thus, we conclude that operational performance can contribute to predicting future financial distress.

## 7. Managerial Insights

Our findings suggest that airlines and outside stakeholders (e.g., regulators) can use operational performance to generate early warning signals regarding future financial distress. Additionally, our empirical model can be used for comparative purposes. To further illustrate these ideas, we compare and contrast four major players in the US airline industry, American, Delta, JetBlue, and Southwest, in terms of their financial distress and ability to manage key operational variables post 9/11. The US airline industry experienced a major shock from the terrorist attacks of September 11, 2001. Figure 1 shows that the average bankruptcy probability exceeded 0.50 for several quarters after 9/11, primarily due to high volatility and steep declines in stock prices. Both American and Delta felt the negative impact immediately with their stock prices declining more than 35% within a week of the attacks. While most carriers were downsizing their workforce and flight schedules after 9/11, Southwest and JetBlue did the opposite by moving into markets deserted by their competitors. Shortly after 9/11, Southwest became the third airline after American and Delta in terms of the number of enplaned passengers, and JetBlue gained major airline status. In light of these observations, we examine American, Delta, JetBlue, and Southwest during the period 2005–2013. Starting in 2005 enables us to exclude the industry-wide effect of 9/11 observed for a few years after the attacks and examine JetBlue after it obtained a relatively large market share in the industry.

**Figure 2 The Average Probability of Bankruptcy and the Corresponding Naïve Distance to Default by Airline Between 2005 Q1 and 2013 Q4**

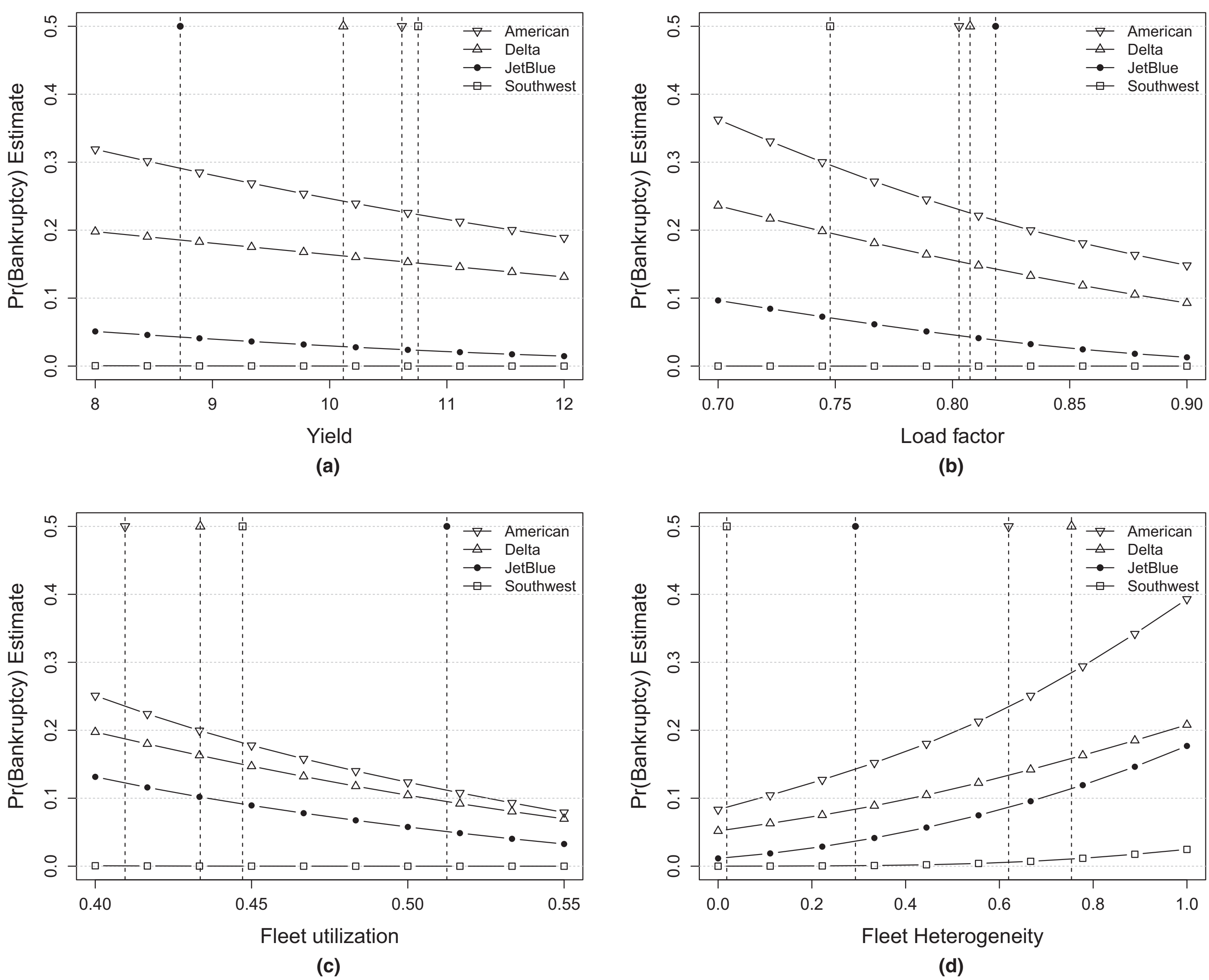



*Notes.* We compute the probability of bankruptcy for airline $i$ in quarter $t$ using Bharath and Shumway's (2008) naïve distance to default, *NDD* metric. We compute the average probability of bankruptcy for airline $i$ by taking a simple average over all quarters in which the airline was publicly traded between the first quarter of 2005 and the last quarter of 2013. We plot $-1 \times NDD$ on the $x$-axis to be consistent with our regression model specifications.

Figure 2 shows the average probability of bankruptcy for airlines between 2005 Q1 and 2013 Q4. Southwest is clearly superior to its competitors in terms of financial stability. Its average probability of bankruptcy is less than 0.01. JetBlue also has a relatively low average bankruptcy probability of 0.04, whereas the average bankruptcy probabilities equal 0.23 and 0.15 for American and Delta, respectively. Our empirical model specification (equation 3) implies a non-linear relationship between operational variables and the probability of bankruptcy, because we transform $NDD$ to $\mathcal{N}(-NDD)$. Consequently, our bankruptcy probability estimates are more sensitive to operational performance for airlines with distance to default values between $-2$ and 2. As illustrated in Figure 2, the slope of the bankruptcy probability curve is steepest in this range.

Figure 3 shows how four operational variables yield, load factor, fleet utilization, and fleet heterogeneity affect bankruptcy probability estimates. Figure 3a captures the sensitivity of our estimates to yield. We generate Figure 3a by using the coefficient estimates of Model 6 reported in Table 6. Estimating airline fixed effects enables to us derive a separate bankruptcy probability curve for each airline. We estimate bankruptcy probabilities by replacing the actual value of yield with numerous yield values varying between 7 and 14 ¢/ASM while keeping the original values of the remaining explanatory variables. (See the notes to Figure 3 for the details of our methodology.) Although high yield is associated with lower

**Figure 3 Sensitivity Analysis**




*Notes.* Figures (a)–(d) demonstrate the sensitivity of our bankruptcy probability estimate to yield, load factor, fleet utilization, and fleet heterogeneity, respectively. We use the coefficient estimates of Model 6 in Table 6 to generate these figures. The vertical dashed lines show the average value of the operational variable (e.g., yield) for each airline for the period between the first quarter of 2005 and the last quarter of 2013. We generate the sensitivity curves for yield as follows. For each airline, we first replace the realized value of yield with a fixed value (e.g., 7) in each quarter. Then we compute the fitted naïve distance to default, using the new value of yield while setting all other variables equal to their realized values. Lastly, we convert the fitted naïve distance to default values to bankruptcy probabilities, and compute the average probability of bankruptcy between the first quarter of 2005 and the last quarter of 2013. Repeating these steps for a range of yield values between 7 and 14 leads to the curves depicted in Figure (a). We generate the sensitivity curves in Figures (b)–(d) using the same approach.

financial distress, it is worthwhile to note that JetBlue has the lowest average yield among the four carriers we examine. Nonetheless, its financial distress is lower compared to American and Delta in part because it compensates in other dimensions (high load factor and high fleet utilization). Bankruptcy probability estimates for American and Delta are sensitive to yield. For instance, if American with an average yield of 10.62 ¢/ASM had Delta's average yield of 10.12 ¢/ASM, then American's bankruptcy probability estimate would change from 0.23 to 0.25. Southwest's bankruptcy probability estimate is insensitive to yield, because Southwest is on the very flat left tail of the curve in Figure 2.

Figure 3b illustrates the sensitivity of bankruptcy probability estimates to load factor. Among the four carriers, JetBlue has the highest load factor. If JetBlue's average load factor declined from 0.82 to Southwest's average load factor of 0.75, then JetBlue's bankruptcy probability estimate would go up from 0.04 to 0.07. Similarly, the bankruptcy probability estimate would increase by more than 0.05 for both

American and Delta if they had Southwest's average load factor. Despite having a low load factor, Southwest has no financial distress in part because it compensates in other dimensions (e.g., low operational complexity).

Figure 3c shows the sensitivity of bankruptcy probability estimates to fleet utilization. Among the four carriers, JetBlue has a clear fleet utilization advantage due to its high proportion of red-eye flights (Huckman and Pisano 2011). In contrast, American and Delta have relatively low fleet utilization in part due to their hub-and-spoke systems. It is unlikely for legacy carriers to achieve JetBlue's fleet utilization. However, if American increased its average fleet utilization from 0.41 to Delta's average utilization, 0.43, American would lower its bankruptcy probability estimate from 0.23 to 0.20. Lastly, Figure 3d shows the sensitivity of bankruptcy probability estimates to fleet heterogeneity. American and Delta have relatively high fleet heterogeneity scores in part due to serving both small and large markets with different aircraft types. Nonetheless, American has lower fleet heterogeneity than Delta. If American's average fleet heterogeneity increased from 0.62 to Delta's average, 0.75, American's bankruptcy probability estimate would increase from 0.23 to 0.28. Among the low-cost carriers, JetBlue's average fleet heterogeneity score is 0.29 due to the introduction of a second plane type in 2005. If JetBlue had not introduced the second plane type, its average fleet heterogeneity would have been zero, and its bankruptcy probability estimate would have been 0.01 instead of 0.04.

In summary, our analyses demonstrate how analysts and regulators can use operational variables for comparative and predictive purposes. When examining an association between an operational variable (e.g., yield) and future financial distress, one should keep airline fixed effects and other operational dimensions (e.g., operational complexity) in mind. Underperforming in one dimension (e.g., yield for JetBlue) does not necessarily imply that an airline will face financial distress, especially if that airline is overperforming in other operational dimensions.

## 8. Conclusion

Financial distress models are useful for firms and their outside stakeholders, such as regulators and investors, because such models provide early warning signals regarding a firm's future financial health. Despite the abundance of financial distress studies, the literature has largely overlooked whether operational performance can predict future financial distress. We identify four areas of operational performance, yielding a comprehensive set of 11 potential operational predictors of future financial distress. In addition, we overcome the limitations of previous airline financial distress studies by (i) using a continuous financial distress metric, (ii) testing associations between lagged values of operational predictors and financial distress, and (iii) showing the superior predictive power of our model compared to a financial ratio-based benchmark model with out-of-sample forecasting. Thus, we contribute to the literature by combining the three-step empirical approach used in the broader financial distress literature with context-rich industry data and thereby demonstrating that all four areas of operational performance contain information useful to predict future financial distress in the US airline industry.

The superior predictive ability of our model implies that financial metrics do not fully explain the future financial performance of a firm and that operational performance can be a leading indicator of financial performance. Consequently, firms and their outside stakeholders should pay close attention to operations as operational performance can help better predict future financial distress. For instance, our model can help a regulatory agency, such as DOT, to monitor the health of airlines by not only tracking operational performance but also deciphering time trends and airline fixed effects, which play important roles in industry studies (Joglekar et al. 2016).

Future industry studies can use our dependent variable to investigate the predictive ability of operational characteristics on future financial distress in other contexts. Another avenue for future research is to identify operational variables that predict the timing of corporate bankruptcies. Even though our logistic regressions show an association between operational performance and bankruptcy status, our dataset with only 20 bankruptcy filings is not large enough to develop a forecasting model that predicts the timing of bankruptcy filings and conduct formal statistical tests to measure its out-of-sample forecast accuracy. Studying bankruptcy prediction in an industry with a larger number of bankruptcy filings (e.g., manufacturing) should lead to insights regarding operational variables that predict the timing of bankruptcy. Hopefully, studies along these lines can advance our understanding of the relationship between operational performance and financial performance.

## Acknowledgments

We gratefully acknowledge the thoughtful and constructive comments provided by Department Editor Nitin Joglekar, the Senior Editor, and three anonymous reviewers. We also thank the participants of the 2016 Supply Chain Finance & Risk Management workshop, especially the discussant Tunay Tunca, for their helpful suggestions. Support for this research was provided by the Dean's Fund for Faculty

Research, Owen Graduate School of Management, Vanderbilt University.

## Notes

[1]The list of bankruptcies in the US airline industry is available at the Airlines for America web site, http://airlines.org/dataset/u-s-bankruptcies-and-services-cessations/. The UCLA-LoPucki bankruptcy database is available at http://lopucki.law.ucla.edu/.

[2]Breusch-Pagan Lagrange multipler tests verified the presence of cross-sectional heteroskedasticity, and Breusch–Godfrey tests verified the presence of serial correlation in Models 1–7.

[3]Panel data models that include lagged dependent variables in their regressors are estimated via generalized method of moments estimators when the number of units is large and time horizon is short (typically less 15 time periods). In contrast, TSCS data estimators perform better in settings like ours where there is a small number of units observed over a long time horizon (Beck and Katz 2011).

## References

Alan, Y., G. P. Gao, V. Gaur. 2014. Does inventory productivity predict future stock returns? A retailing industry perspective. *Management Sci.* **60**(10): 2416–2434.

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Finance* **23**(4): 589–609.

Altman, E. I. 1993. *Corporate Financial Distress and Bankruptcy*, 2nd edn. John Wiley & Sons, Hoboken, NJ.

Altman, E. I., E. Hotchkiss. 2010. *Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt*, 3rd edn. John Wiley & Sons, Hoboken, NJ.

Atkinson, S. E., K. Ramdas, J. W. Williams. 2016. Robust scheduling practices in the US airline industry: Costs, returns, and inefficiencies. *Management Sci.* **62**(11): 3372–3391.

Baltagi, B. 2008. *Econometric Analysis of Panel Data*, 5th edn. John Wiley & Sons, West Sussex, UK.

Becchetti, L., J. Sierra. 2003. Bankruptcy risk and productive efficiency in manufacturing firms. *J. Bank. Finance* **27**(11): 2099–2120.

Beck, N. 2001. Time-series cross-section data: What have we learned in the past few years? *Ann. Rev. Polit. Sci.* **4**(1): 271–293.

Beck, N., J. N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *Am. Polit. Sci. Rev.* **89**(3): 634–647.

Beck, N., J. N. Katz. 2011. Modeling dynamics in time-series-cross-section political economy data. *Ann. Rev. Polit. Sci.* **14**: 331–352.

Behn, B. K., R. A. Riley, Jr. 1999. Using nonfinancial information to predict financial performance: The case of the US airline industry. *J. Accou. Auditing Finance* **14**(1): 29–56.

Bharath, S. T., T. Shumway. 2008. Forecasting default with the Merton distance to default model. *Rev. Financ. Stud.* **21**(3): 1339–1369.

Bodea, T., M. Ferguson. 2014. *Segmentation, Revenue Management and Pricing Analytics*. Routledge, New York, NY.

Bonnefoy, P., R. J. Hansman. 2004. Emergence and impact of secondary airports in the United States. 4th AIAA ATIO Forum, Chicago, IL. Available at http://hdl.handle.net/1721.1/34958

Borenstein, S. 2011. On the persistent financial losses of US airlines: A preliminary exploration. National Bureau of Economic Research, Working paper #16744.

Borenstein, S., N. L. Rose. 1995. Do airlines in chapter 11 harm their rivals? Bankruptcy and pricing behavior in US airline markets. National Bureau of Economic Research, Working paper #5047.

Campello, M., J. Gao. 2017. Customer concentration and loan contract terms. *J. Financ. Econ.* **123**(1): 108–136.

Chava, S., R. A. Jarrow. 2004. Bankruptcy prediction with industry effects. *Rev. Finance* **8**(4): 537–569.

Chava, S., A. Purnanandam. 2011. The effect of banking crisis on bank-dependent borrowers. *J. Financ. Econ.* **99**(1): 116–135.

Chow, G., R. D. Gritta, E. Leung. 1991. A multiple discriminant analysis approach to gauging air carrier bankruptcy propensities: The AIRSCORE model. *J. Transport. Res. Forum* **31**(2): 371–377.

Ciliberto, F., C. Schenone. 2012. Bankruptcy and product-market competition: Evidence from the airline industry. *Int. J. Ind. Organ.* **30**(6): 564–577.

Collier, D. A., J. R. Evans. 2014. *Operations Management*, 5th edn. Cengage Learning, Boston, MA.

Cook, B. M. 1998. SABRE soars. *OR/MS Today* **25**(3): 26–31.

Deshpande, V., M. Arikan. 2012. The impact of airline flight schedules on flight delays. *Manuf. Serv. Oper. Manag.* **14**(3): 423–440.

Forbes, S. J. 2008. The effect of service quality and expectations on customer complaints. *J. Ind. Econ.* **56**(1): 190–213.

Garvett, D. S., K. J. Hilton. 1999. What drives airline profits? A first look. G. F. Butler, M. R. Keller, eds. *Handbook of Airline Finance*. Aviation Week, McGraw-Hill, 173–197.

Girotra, K., C. Terwiesch, K. T. Ulrich. 2007. Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. *Management Sci.* **53**(9): 1452–1466.

Gittell, J. H. 2003. *The Southwest Airlines Way: Using the Power of Relationships to Achieve High Performance*. McGraw-Hill, New York, NY.

Goyal, A., I. Welch. 2003. Predicting the equity premium with dividend ratios. *Management Sci.* **49**(5): 639–654.

Gritta, R. D., B. Adrangi, S. Davalos, D. Bright. 2008. A review of the history of air carrier bankruptcy forecasting and the application of various models to the US airline industry, 1980–2005. K. Gratzer, D. Stiefel, eds. *History of Insolvency and Bankruptcy from an International Perspective*. Sodertorn Academic Studies, Huddinge, 193–214.

Gudmundsson, S. V. 2002. Airline distress prediction using non-financial indicators. *J. Air Transport.* **7**(2): 3–24.

Gudmundsson, S. V. 2004. Management emphasis and performance in the airline industry: An exploratory multilevel analysis. *Transport. Res. Part E Logistics Transport. Rev.* **40**(6): 443–463.

Heskett, J. L., W. E. Sasser, L. A. Schlesinger. 1997. *The Service Profit Chain*. The Free Press, New York, NY.

Hofer, C., M. E. Dresner, R. J. Windle. 2009. The impact of airline financial distress on US air fares: A contingency approach. *Transport. Res. Part E Logistics Transport. Rev.* **45**(1): 238–249.

Hosmer, D. W., S. Lemeshow. 2004. *Applied Logistic Regression*, 2nd edn. John Wiley & Sons, Hoboken, NJ.

Huckman, R. S., G. P. Pisano. 2011. *JetBlue Airways: Managing Growth*. Harvard Business School Case No.9-609-046, Harvard Business School Publishing, Boston, MA.

Joglekar, N. R., J. Davies, E. G. Anderson. 2016. The role of industry studies and public policies in Production and Operations Management. *Prod. Oper. Manag.* **25**(12): 1977–2001.

Johnson, N. J. 1978. Modified t tests and confidence intervals for asymmetrical populations. *J. Am. Statist. Assoc.* **73**(363): 536–544.

Kesavan, S., V. Gaur, A. Raman. 2010. Do inventory and gross margin data improve sales forecasts for US public retailers? *Management Sci.* **56**(9): 1519–1533.

Lapré, M. A. 2011. Reducing customer dissatisfaction: How important is learning to reduce service failure? *Prod. Oper. Manag.* **20**(4): 491–507.

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

Li, G., I. Kwan, D. Rutherford. 2015. US domestic airline fuel efficiency ranking, 2014. International Council on Clean Transportation, Washington, DC. Available at http://www.theicct.org/us-domestic-airline-fuel-efficiency-ranking-2014.

Lu, C., A. S. Yang, J. Huang. 2015. Bankruptcy predictions for US air carrier operations: A study of financial data. *J. Econ. Finance* **39**(3): 574–589.

Luo, X. 2007. Consumer negative voice and firm-idiosyncratic stock returns. *J. Market.* **71**(3): 75–88.

Merton, R. C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *J. Finance* **29**(2): 449–470.

Metters, R. D., K. H. King-Metters, M. Pullman, S. Walton. 2008. *Successful Service Operations Management*, 2nd edn. South-Western College Publishing, Mason, OH.

Phillips, G., G. Sertsios. 2013. How do firm financial conditions affect product quality and pricing? *Management Sci.* **59**(8): 1764–1782.

Pilarski, A., T. Dinh. 1999. Numerical scoring approach to credit risk analysis. G. F. Butler, M. R. Keller, eds. *Handbook of Airline Finance*. McGraw-Hill, 329–342.

Ramdas, K., J. Williams, M. Lipson. 2013. Can financial markets inform operational improvement efforts? Evidence from the airline industry. *Manuf. Serv. Oper. Manag.* **15**(3): 405–422.

Rapach, D. E., J. K. Strauss, G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Rev. Financ. Stud.* **23**(2): 821–862.

Shumway, T. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *J. Bus.* **74**(1): 101–124.

SITA. 2017. The baggage report. Societe Internationale de Telecommunications Aeronautiques, Geneva, Switzerland. Available at https://www.sita.aero/resources/type/surveys-reports/baggage-report-2017

Skinner, W. 1974. The focused factory. *Harv. Bus. Rev.* **52**(3): 113–121.

Smith, B. C., J. F. Leimkuhler, R. M. Darrow. 1992. Yield management at American Airlines. *Interfaces* **22**(1): 8–31.

Staats, B. R., F. Gino. 2012. Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* **58**(6): 1141–1159.

Talluri, K. T., G. J. Van Ryzin. 2005. *The Theory and Practice of Revenue Management*. Springer Science & Business Media, New York, NY.

Tsikriktsis, N. 2007. The effect of operational performance and focus on profitability: A longitudinal study of the U.S. airline industry. *Manuf. Serv. Oper. Manag.* **9**(4): 506–517.

Vassalou, M., Y. Xing. 2004. Default risk in equity returns. *J. Finance* **59**(2): 831–868.

Weun, S., S. E. Beatty, M. A. Jones. 2004. The impact of service failure severity on service recovery evaluations and post-recovery relationships. *J. Serv. Market.* **18**(2): 133–146.