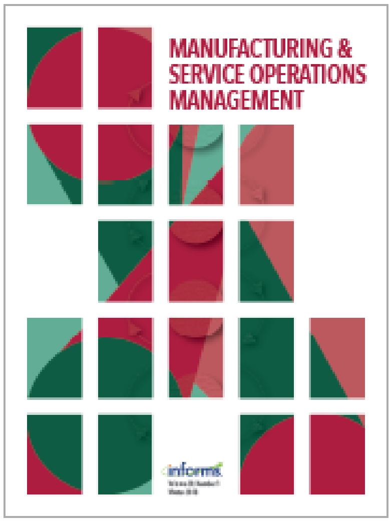


Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Can Financial Markets Inform Operational Improvement Efforts? Evidence from the Airline Industry

Kamalini Ramdas, Jonathan Williams, Marc Lipson,

M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 15, No. 3, Summer 2013, pp. 405–422
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

http://dx.doi.org/10.1287/msom.2013.0448

Can Financial Markets Inform Operational Improvement Efforts? Evidence from the Airline Industry

Kamalini Ramdas
Management Science and Operations, London Business School, London NW1 4SA, United Kingdom, kramdas@london.edu

Jonathan Williams
Department of Economics, Terry College of Business, University of Georgia, Athens, Georgia 30602, jonwms@uga.edu

Marc Lipson
Darden School of Business, University of Virginia, Charlottesville, Virginia 22903, mlipson@virginia.edu

We investigate whether stock price movements can inform operations managers as to where they should focus improvement efforts. We examine how unexpected performance along several dimensions of service quality—on-time performance, long delays and cancellations, lost bags, and denied boardings—impacts contemporaneous stock returns. Prior research suggests that airlines buffer their flight schedules and engage in expensive employee incentive programs to increase the likelihood of on-time arrival. We find that only long delays are penalized by the market, and we identify a number of carrier-specific factors that alter the financial impact of long delays. We find that the penalty a carrier faces for long delays is significantly higher if it operates a high percentage of short-haul or connecting flights, or if its competitors incur fewer long delays in the same time period. Our findings suggest that developing ways to curtail long delays is a useful future research area.

Key words: econometric analysis; empirical research; service operations; quality management; OM–finance interface
History: Received: February 12, 2012; accepted: March 8, 2013. Published online in *Articles in Advance* May 23, 2013.

1. Introduction

In any industry, operations managers can improve performance along a number of different dimensions that impact the quality and the cost of their offerings. Thus, operations managers need to know where to focus their efforts. Although managers often consider the profit impact of potential improvements to prioritize their efforts, this practice is consistent with maximizing shareholder value only if the estimated profit impact fully reflects the present value of the changes in all current and future relevant revenues and costs associated with the operational improvements.

Recent research has used event studies to examine how individual operational levers impact market value (e.g., Hendricks and Singhal 2003, 2005). Researchers in accounting have analyzed the relationship between analysts' estimates of a firm's future earnings and its stock returns (e.g., Abarbanell and Bushee 1997). Cannon et al. (2008) include operational variables to predict future earnings and find that the incremental predictive information from operational variables is not fully reflected in the earnings predictions of financial analysts.

In this paper, we directly examine how investors react to changes in service quality in the airline industry. Airlines compete along several easily measurable dimensions of service quality, and timeliness is a core element of airline service quality (Anderson et al. 2005). Although the Department of Transportation (DOT) classifies any flight that arrives 15 minutes later than its scheduled arrival time as a late arrival, delays can range from a few minutes to several hours. Cancelled flights can result in long delays plus the inconvenience of rebooking for travelers. Passengers can also be inconvenienced by poor in-flight service, damaged or lost baggage, or being denied boarding onto an overbooked flight. Given the multiple facets of airline service quality, airline operations managers must know which ones to target for improvement. We use information from stock price movements to shed light on this issue. In this regard, our contribution is similar to that of Girotra et al. (2007), who use stock market data to guide portfolio decisions in the pharmaceutical industry.

The Federal Aviation Administration requires the major U.S. airlines to report detailed data on their flight operations to the DOT within 15 days of month-end, in addition to disclosing quarterly and annual financial statements (see Shumsky 1993 for a historical perspective on the origins of on-time disclosures). The DOT makes these data available to the public via downloadable databases with flight-by-flight data and publications that contain summary statistics on each carrier's flight delays, mishandled bags, and denied boardings. Each of these items is released by the DOT regularly, typically within a month or two of the end of any particular month. In the first week of each month, U.S. airlines also voluntarily release summary statistics on their flight operations in the previous month. Information regarding each airline's operations is therefore released more frequently than accounting information and often well in advance of the release of financial statements.

Besides the periodic information releases described above, information about airline service quality also becomes publicly available on a daily basis. Important dimensions of service quality on each scheduled flight are public information, and passengers themselves possess this information in real time. In recent years, airlines have also started to display flight delay information on their websites in real time. Several websites and airline blogs also provide real-time information on flight delays.

Using standard event study methodology, we found no movement in abnormal returns surrounding the dates of official release of airline service quality bulletins by the DOT. This led us to investigate the possibility that information about airline service quality is rapidly incorporated into stock prices in real time. Maloney and Mulherin (2003) document the remarkable speed and accuracy of price discovery by financial markets when information is particularly diffuse.

Using monthly data for 11 major U.S. carriers over the period 1990–2009, we model investors' expectations of operational performance along different dimensions of service quality, and then examine how unexpected operational performance impacts contemporaneous airline stock price movements. If financial markets efficiently process and aggregate information, an airline's stock price at the end of each month should incorporate all available information, including real-time information about the service quality that its passengers experienced in the month. Therefore, invoking the efficient markets hypothesis (Fama 1970), any unpredicted realization of service quality in a month should be correlated contemporaneously with stock price movements in that month.

In an airline industry study, Forbes (2008) examines how the gap between actual and expected quality impacts customer complaints, modeling expectations of the frequencies of delays and mishandled bags using month and carrier dummies. By relying on the financial market as an efficient processor of all available information, our forecasting models in principle incorporate more information on service quality than existing studies. Unlike Luo (2007), who relates consumer complaints to next period returns in the airline industry, we explicitly control for expectations that would otherwise result in correlated omitted variable bias.

We use a panel data approach that relates unexpected performance to contemporaneous stock returns, as has been done by several widely cited papers including Aaker and Jacobson (1994), Bartov et al. (2002), Kasznik and McNichols (2002), and Roll (1984). Bartov et al. (2002, p. 177) describe this approach as the "information content/events study paradigm."

We find that an unexpected increase in the proportion of flights incurring long delays or cancellations has a significant negative impact on stock returns, whereas an unexpected increase in the proportion of on-time flights has no significant impact. This is an important finding because there is very clear evidence that airlines try hard to get their planes to arrive "on time." Deshpande and Arıkan (2012) and Arıkan et al. (2013) find evidence that this is the case, and Forbes et al. (2011) find that airlines actually implement costly employee incentive programs to ensure this. Knowing how the market penalizes different lengths of delay is also very useful because what managers need to do operationally to reduce short delays can differ substantially from what they must do to reduce long delays.

We highlight several successful approaches used by airlines to specifically reduce long delays and cancellations. For example, delay propagation is an important source of long delays and cancellations. We learned from Southwest managers that they try to limit delay propagation from one flight to the next on an aircraft by isolating flights to or from delay-prone airports into pockets of their network, where possible, to avoid contamination of generally delay-free regions of their network. Delays can also be propagated from one aircraft to another, by customers who have connecting flights, crew who have been scheduled to be moved from one aircraft to another, or scarce facilities. Managers at Southwest Airlines shared with us how their point-to-point network structure and the capacity flexibility they gain from having only one type of aircraft enable them to limit this type of delay propagation. Atkinson et al. (2012) show empirically that increased capacity flexibility in an airline network reduces long delays and cancellations.

By providing a way to estimate the negative market impact of long delays, we provide a yardstick against which to compare the costs of investing in practices that reduce such delays. Using average cost estimates provided by the Air Transport Association (ATA), we estimate that the reduction in market value associated with larger-than-predicted long delays is far greater than the airlines' estimates of average out-of-pocket costs for these delays. This finding suggests that many of the costs associated with long delays are opportunity costs faced in the future—such as loss of future sales due to passengers switching away from an airline or a reduction in earnings due to passengers' unwillingness to pay higher fares.

To further unpack the mechanism though which unexpected long delays impact stock returns, we examine how a number of carrier-specific factors affect this relationship. We find that the negative effect of unexpected long delays on stock returns is greater for airlines that fly a greater percentage of short-haul flights, for which alternative modes of transport are more readily available. The stock market penalty for unexpected long delays is also greater for airlines with a higher frequency of connecting passengers, consistent with the notion that missed flight connections are a source of both higher out-of-pocket costs and greater customer dissatisfaction, which would impact an airline's future revenues. We also show that unexpected long delays hurt stock prices more if competitors are offering better-than-expected service in the same time period, suggesting that the market expects consumers' future buying behavior to be driven by current realizations of service quality.

2. Hypothesis Development

In any industry, operational variables impact profitability, and ultimately firm value. Flight delays and cancellations result in two broad categories of costs to the airlines: direct out-of-pocket costs and costs associated with lost revenue due to consumers' reduced willingness to pay and reduced inclination to fly on an airline after experiencing poor service (Forbes 2008). Additional fuel and crew costs are the largest component of out-of-pocket costs (Hauser 2002). The ATA estimated the direct cost of the 100 million minutes of delays for the U.S. carriers in 2009 as \$6.1 billion, or \$61 per minute of delay. The costs associated with lost revenues are much harder to estimate, because they are opportunity costs and therefore do not show up directly in accounting statements. Also, a large portion of these costs may be incurred only in future periods.

With a view to estimating the total financial impact of poor service, which includes both the direct short-term costs and the longer-term opportunity costs, we examine how service quality affects firm value. Market prices reflect the present value of future cash flows, less the present market value of debt obligations. Since a firm's operations impact its cash flows, variation in its operational performance should have a measurable impact on its stock price. An anticipated reduction in future cash flows due to poor service will result in a reduction in market value. The negative impact on market value will be larger if the reduction in future cash flows induced by current poor service is more long lasting. If enough passengers were to switch away from an airline after experiencing significant flight delays, the impact of these delays on stock prices could exceed the direct short-term costs of delays by a large margin.

Since stock prices react only to new information, the timing of information availability is important. The information available about a firm in any industry is partly public knowledge produced by firms in compliance with disclosure and reporting requirements or for voluntary disclosure, partly information that can be bought from entities that track the industry, and partly information diffused in the knowledge and experience of consumers. In this relatively transparent information environment, it is reasonable to expect that information about service quality will be incorporated into stock prices in real time, particularly since such stock price reactions occur even in far less transparent environments. For example, only a day after the crash of the space shuttle Challenger in 1986, stock price movements pointed toward Morton Thiokol, manufacturer of the defective O-ring in its rocket booster, as the source of the shuttle's quality problem (Maloney and Mulherin 2003). It took months of deliberation and debate by a presidential panel of experts to reach the same conclusion.

Assuming that the efficient markets hypothesis (Fama 1970) holds, financial markets should incorporate all publicly available forecasts about the service quality of U.S. airlines. Stock prices should, therefore, incorporate the predictable part of the important dimensions of each airline's service quality. Unpredicted changes in service quality should be contemporaneously correlated with stock price movements. Since we know of no existing forecasts of airline service quality, we form our own forecasts using historical data, as described in §4.

Hypothesis 1. *Unpredicted changes in airline service quality will have a contemporaneous impact on stock returns.*

We also consider various dimensions of service quality. The U.S. DOT tracks several indicators of on-time performance, including the percentage of late arrivals, the percentage of very late arrivals, and the percentage of cancelled flights. The percentage of late arrivals is frequently reported in the press, and airlines set their flight times so as to meet target service levels on this metric. Despite this focus on late arrivals, we expect that among all late flights, those incurring longer delays and cancellations will have a disproportionately larger negative impact on market returns than those that are simply late, for several reasons. First, flights that incur long delays or cancellations for reasons within the airlines' control result in substantial costs for refunds, rebooking of passengers, hotels, meals and incidental expenses, over and above additional fuel and staff costs incurred by all late flights (Ferguson et al. 2012). Second, in terms of passenger experience, passengers making connections at hub airports have a higher chance of missing their connections when the first leg of their flight has a longer delay. Third, subsequent flights on an aircraft that incurs a long delay are much more likely to incur long delays or cancellations due to the cascading nature of delays, increasing passenger discomfort. Fourth, with long delays, business passengers, who have less flexible schedules, are more likely to miss appointments, causing major inconvenience. Finally, research in psychology suggests that extreme events are much more likely to be remembered vividly than mundane events (Rubin and Talarico 2003).

So, extreme delays and cancellations are more likely to permanently hurt customer perceptions about an airline than short delays, and affect future buying behavior, which in turn will affect stock returns. Thus we expect the marginal cost of delay to be greater for long delays than for short delays.

Hypothesis 2. *The marginal cost of a minute of delay is greater for long delays than for short delays.*

We now delve deeper into the mechanisms by which delays impact returns. We expect the negative impact of unexpected delays on returns to be greater for short-haul flights than for long-haul flights for two reasons. First, for short-haul flights, road transport provides a viable alternative mode of transport. Passengers are much more likely to substitute driving for flying for a short-haul flight (Russon and Bowers 1990). Second, a delay of any length is greater relative to flight time for a short-haul flight than for a long-haul flight.

The impact of delays is greater when passengers miss connecting flights due to the delay. Bratu and Barnhart (2005) highlight the importance of considering passenger disruption due to missed flight connections in evaluating the impact of flight delays. Aside from the inconvenience to passengers, missed connections are more likely to result in higher cost to airlines due to potential rebooking at higher fares or accommodation provided at the airline's expense. We therefore expect that the negative impact of long delays on stock returns will be even greater when there is a higher percentage of passengers taking connecting flights.

Load factor, the percentage of seats filled by fare-paying passengers, is an industry metric of how full an airline's flights are. Prior research suggests a trade-off between load factor and delays (Atkinson et al. 2012, Rupp and Holmes 2006). If flights are less full, additional passengers can be more easily accommodated at the last minute, reducing the disruption due to long delays. We therefore expect that the negative impact of long delays on stock returns will be mitigated by lower levels of load factor.

Mazzeo (2003) finds that delays are significantly higher on monopoly routes. This finding suggests that the airlines impute a lower likelihood of lost future sales in markets where there are no alternatives for passengers who wish to travel by air. We therefore expect the negative effect of unexpected long delays on stock returns to be exacerbated when a carrier's competition is concurrently performing better than expected. On the other hand, the additional future sales an airline obtains by beating expectations regarding operational performance should be amplified if its close competitors are concurrently performing worse than expected.

Research in consumer psychology suggests that customers react differently to product failures or service inconvenience depending on whether the underlying cause for the problem was within the firm's control (Hamilton 1980). In the airline context, since 2003 the DOT has classified the "cause" of a delay into five categories—air carrier delay, aircraft arriving late, security delay, national aviation system delay, and extreme weather delay—the first two of which are within the carriers' control. Folkes et al. (1987) and Anderson et al. (2009) find that flight delays that can be attributed to the airlines result in greater dissatisfaction than those that can be attributed to causes out of the airlines' control. Consistent with this view, we expect that delays attributable to causes within the airlines' control will have a greater negative impact on stock returns than those attributable to causes outside the airlines' control.

Hypothesis 3. *The negative impact of unexpected long delays will be exacerbated by shorter haul lengths, increased frequency of connecting passengers, higher levels of load factor, competitors' performance, and a larger proportion of delays being attributable to the carrier.*

The transparency of information on key dimensions of service quality in the airline industry makes it an excellent candidate for a study of the impact of service quality on market value. The industry's service quality and operations are also sensitive to exogenous shocks such as weather, fuel prices, and national security threats. In fact, the exogeneity of weather and its substantial impact on delays makes it a very suitable choice for an instrument for delays. We test the robustness of our analysis using weather as an instrument. In the next section, we describe our data and the variables used to operationalize the above hypotheses.

3. Data

3.1. Operational Data

3.1.1. Delays. Data on flight delays and cancellations were collected from the Bureau of Transportation Statistics' (BTS) Airline On-Time Performance database, which records the details of delays and incidence of cancellations for all domestic flights for those 11 carriers required to report to the DOT from 1988 on. From these raw data, we constructed different moments of the distribution of flight delays for every carrier in each month.

As mentioned earlier, the DOT uses a 15-minute threshold to classify flights as being on time or late, and classifies cancelled flights as a separate class. For brevity, we will use the term "late" to denote "late or cancelled." Let $P(>15)_{it}$ denote the percentage of scheduled flights that are 15 or more minutes late for carrier i in month t. Similarly, we construct four measures of long delays: $P(>120)_{it}$, $P(>180)_{it}$, $P(Cancel)_{it}$, and $P(3\sigma)_{it}$ (percentage of delays that are three standard deviations above the mean).

Let $E(Delay)_{it}$ denote the average delay for carrier i in month t, where the average is taken over delays greater than zero. Including the average delay in our specifications allows us to examine whether the marginal cost of an additional minute of delay is greater for longer delays. For example, consider two different carriers i and j in time period t (or the same carrier in different time periods) that both operate a total of 100 flights and incur an average delay of 10 minutes (1,000 total minutes of delay). However, suppose that the first carrier had no delays of more than 120 minutes ($P(>120)_{it} = 0$), whereas the second carrier had five flights delayed over 120 minutes ($P(>120)_{jt} = 0.05$). If the marginal cost of delays is increasing in the length of delay, the second carrier will incur more costs for the same 1,000 minutes of delay.[1]

From 2003 onward, the DOT also classifies the "cause" of a delay into one of five categories. The first two categories are attributable to a carrier, *air carrier delay* and *aircraft arriving late*. The other categories outside the carriers' control are *security delay*, *national aviation system delay*, and *extreme weather*. This allows us to examine whether the source of the delay has a differential impact on an airline's profitability (e.g., on how consumers view the delay and make subsequent travel plans). We define $Attrib_{it}$ as the proportion of carrier i's total minutes of delays in month t that fall into the two categories attributable to the carrier. This proportion, across all carriers, was approximately 58% in 2003, but steadily rose to nearly 70% by 2009.

Table 1 presents overall summary statistics for the variables described above. About 21% of flights by the 11 major carriers studied are over 15 minutes late. In the tail of the delays distribution, about 3.0% of flights are over two hours late, about 2.1% are over three hours late, and about 1.7% are cancelled. Flights arriving after the scheduled arrival time have an average delay of just over 24 minutes. Figures 1(a) and 1(b) show the significant seasonality in each measure of delays, spiking during the winter and summer months. The noticeable spike in long delays in September demonstrates the importance of singular events (e.g., September 11th). We discuss how we control for these trends and events in our empirical analysis in §4. Figure 1(c) presents the distribution of delay length across all carriers and months. The large standard deviation for $Attrib_{it}$ is partly due to strategic decisions by carriers (e.g., Southwest serving secondary airports and largely fair-weather routes). By modeling investor expectations of performance by carrier, we control for these differences.

3.1.2. Mishandled Baggage and Denied Boardings. Data on the ratio of number of passengers who experienced mishandled baggage or were denied boarding to total enplanements, denoted by $MissedBags_{it}$ and $Oversales_{it}$, respectively, were collected from the *Air Travel Consumer Report*, a DOT publication. From Table 1, we see that, on average, carriers mishandle about 4.5 bags per thousand handled, whereas 0.83 passengers out of every 10 thousand enplanements were denied boarding. There is no clear pattern across months or carriers for denied boardings, but mishandled bags do spike during the December holiday season. These two series are available on the DOT website from 1998 onward, whereas data for other operational variables are available from 1990 onward. Therefore, we test our hypotheses on both a long panel (1990–2009) and a short panel (1998–2009) to include these additional operational metrics.

3.1.3. Revenue Passenger Miles, Load Factor, Haul Length, Congestion, and Connectivity. Value-maximizing airlines should aim to spread their high fixed costs over as many passengers as possible on

[1] When calculating $E(Delay)_{it}$, we average over all flights arriving after the scheduled arrival time. If early flights were included in the average, we would be assuming that the marginal cost of an additional minute of delay was symmetric, but opposite in sign, on either side of the origin. Although there may be benefits associated with arriving early, we believe these are small relative to the costs of arriving late.

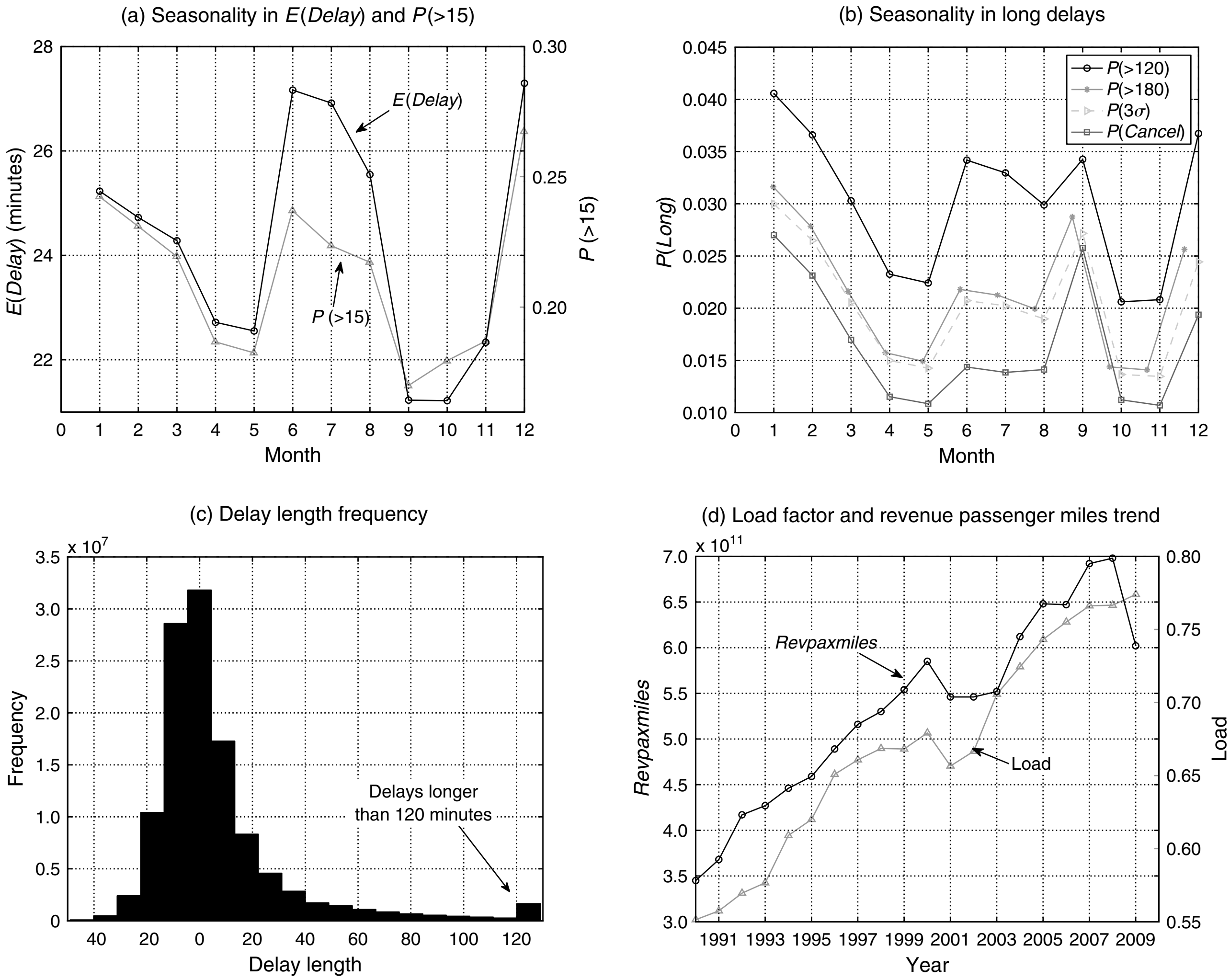
Table 1 Summary Statistics

	Description	Mean	Std. dev.	Min	Max	No. of obs.
Financial and operational						
Ret	Raw excess returns (%)	0.008	0.164	−0.634	1.133	1,708
AbnRet	Abnormal returns (%)	0.000	0.145	−0.586	1.039	1,708
MktVal	Market value of equity (unit = millions)	3,110.0	3,330.0	33.9	16,800.0	1,713
$P(>15)$	Prob. of delay greater than 15 minutes	0.212	0.074	0.026	0.639	2,162
$P(>120)$	Prob. of delay greater than 120 minutes	0.030	0.027	0.001	0.588	2,162
$P(>180)$	Prob. of delay greater than 180 minutes	0.021	0.024	0.000	0.586	2,162
$P(>3\sigma)$	Prob. of delay 3 std. dev. greater than mean	0.019	0.022	0.000	0.585	2,162
$P(Cancel)$	Prob. of cancel	0.017	0.022	0.000	0.584	2,162
$\Delta Revpaxmiles$	Change in revenue passenger miles (unit = billions)	0.004	0.676	−5.244	3.654	2,151
MissedBags	Mishandled bags per 1,000 passengers	4.516	1.792	1.360	17.130	1,392
Oversales	Denied boardings per 10,000 passengers	0.830	0.644	0.000	3.770	1,387
Strike	Formal work stoppage indicator	0.009	0.096	0.000	1.000	2,162
Slowdown	Informal slowdown or sick-out by employees indicator	0.024	0.153	0.000	1.000	2,162
Bankrupt	Bankruptcy indicator	0.060	0.238	0.000	1.000	2,162
$E(Delay)$	Average delay for flights arriving after scheduled time	24.267	7.403	9.304	54.489	2,162
$\Delta Fuel$	Change in price of jet fuel (unit = $ per gallon)	0.005	0.154	−0.996	0.401	2,130
Attrib	Proportion of total delays attributable to carrier	0.672	0.229	0.251	0.883	936
Load	Proportion of seats filled	0.687	0.090	0.415	0.910	2,162
Haul	Average haul length for flights (100 s of miles)	26.377	11.320	6.290	78.357	2,162
Connect	Ratio of connections to enplanements	0.361	0.337	0.188	0.452	2,162
Congest	Ratio of departures to gates	4.802	2.065	3.509	5.471	2,162
Weather						
Prec	Total precipitation (millimeters)	24.444	10.247	1.217	108.182	2,162
SnowFall	Snow fall (millimeters)	1.371	2.387	0.000	21.700	2,162
SnowDepth	Snow depth on ground (millimeters)	8.284	20.337	0.000	219.549	2,162
Sun	Sunshine (minutes)	351.453	236.636	211.636	2,125.426	2,162
TempMax	Maximum temperature (degrees Celsius)	18.047	17.769	11.836	27.707	2,162
AvgWind	Average wind speed (meters per second)	38.507	5.007	25.880	55.277	2,162
PkWind	Peak wind speed (meters per second)	43.236	66.056	32.815	602.935	2,162

each flight. A key operational variable for air carriers is revenue passenger miles. One revenue-paying passenger flying one mile generates one revenue passenger mile. An unpredicted improvement in an airline's revenue passenger miles, holding service quality constant, should raise the airline's stock price, as higher revenue passenger miles indicate better utilization of fixed resources (i.e., aircraft and airport facilities). We obtained data on revenue passenger miles, $Revpaxmiles_{it}$, from the BTS' publicly available T-100 Domestic Segment database, which contains both the number of passengers transported and total seats on every flight segment flown by a carrier in a month. By multiplying the number of passengers by flight distance and then aggregating across flight segments for each carrier, we get total revenue passenger miles for each carrier in each month. From the T-100 data, we calculate three other variables: $Load_{it}$ is defined as the fraction of total seats occupied by a revenue-generating passenger across all flights in carrier i's network in month t, whereas $Haul_{it}$ is the average haul length (in thousands of miles) of those flights; $Congest_{it}$ is calculated as a weighted average of the ratio of daily total departures to boarding gates at each airport in a carrier's network, where the weights are the share of the carrier's enplanements at each airport. See Atkinson et al. (2012) and Williams (2012) for a discussion of the data on gates. The variable $Congest_{it}$ serves as an important control in each of our regressions, because it may be correlated with both delays and returns and would result in omitted variable bias if left out. These data series are available from 1990 to 2009.

Figure 1(d) shows an upward trend in both total revenue passenger miles (in billions) and average load factor from 1990 to 2009. There are significant deviations from this trend, driven by recessions and recent consolidation in the industry (e.g., Delta's acquisition of Northwest). For this reason, in our forecasts of operational variables, we focus on monthly changes in revenue passenger miles and load factor, $\Delta Revpaxmiles_{it}$ and $\Delta Load_{it}$, respectively. These variables are summarized in Table 1. The large amount of variation in these variables captures the rapid expansion of operations by some carriers (e.g., Southwest), whereas others were actually contracting over time (e.g., United). There are also significant differences across carriers in the average haul length, with legacy carriers (i.e., United, US Air, Delta, American, and Northwest) operating longer flights on average.

Figure 1 Descriptive Statistics on Operations



We measure the degree of connectivity in a carrier's network from the DOT's Origin and Destination Survey, Data Bank 1B (DB1B), data. The DB1B is a 10% survey of all domestic itineraries, which allows us to observe the fraction of all enplanements that involve a connection. For example, a one-way itinerary from New York (JFK) to Los Angeles (LAX) with a connection in Atlanta (ATL) involves two enplanements (at JFK and ATL) and one connection (at ATL). If a third leg to Salt Lake City (SLC) were added to the one-way itinerary, it would involve three enplanements and two connections. For each carrier, i, and month, t, we calculate the ratio of connections to enplanements, $Connect_{it}$, as our measure connectivity. Thus, $Connect_{it}$ will be closer to one for carriers serving more passengers via connecting service. The large standard deviation reported for $Connect_{it}$ in Table 1 is due to the variation across legacy and low-cost carriers in the degree of connectivity in their networks.

3.1.4. Other Operational Variables. Specialized labor resources, from pilots to mechanics, play an important role in the airline industry. Their skills are not easily replaced, giving unions representing these labor interests a great deal of power. As a result, operational disruptions in the form of formal strikes or less formal sick-outs or slowdowns by employees are common in the industry. Slowdowns and sick-outs involve employees performing everyday tasks more slowly or collectively calling in sick to work to disrupt normal operations. Also, five of the 11 carriers we examine experienced at least one bankruptcy from 1990 to 2009. The U.S. General Accounting Office (2003) and Ciliberto and Schenone (2012) discuss the impact of labor disruptions and bankruptcies, respectively, on operational performance. For our purposes, it is necessary to control for these events, because they impact both operational performance and the market valuation of a carrier. To this end, we collected data on labor disruptions and bankruptcies from the National Mediation Board and Ciliberto and Schenone (2012). From these data, we construct three carrier-specific dummy variables, $Strike_{it}$, $Slowdown_{it}$, and $Bankrupt_{it}$, to indicate periods in which a carrier experiences a formal strike, slowdown or sick-out by employees, and bankruptcy, respectively. These variables are summarized in Table 1. From 1990 to 2009, American and Northwest experienced strikes, whereas American, Delta, Northwest, and United incurred slowdowns or sick-outs. The carriers filing for bankruptcy protection during our period of study were Continental, Delta, Northwest, United, and US Airways. Since the months leading up to bankruptcy and including the filing may provide the largest simultaneous service disruption and stock market effect, we estimate specifications including indicators for each of the four months prior to and following each type of event. The only significant coefficients were in the month of or preceding an event, and the remainder of the results are unchanged, so we present these more parsimonious specifications.

Jet fuel, along with labor, is a critical determinant of costs in the industry, and variation in prices can significantly alter managerial decisions and operational outcomes (e.g., scheduling additional flights when fuel costs are low may lead to additional delays). To ensure that the price of jet fuel is not a correlated omitted variable in our regressions relating operational performance to stock market valuations of carriers, we collected data on jet fuel prices from the U.S. Energy Information Administration. As with revenue passenger miles, there is an upward trend in fuel prices over time even after adjusting for inflation. For this reason, we focus on monthly changes in fuel prices, $\Delta Fuel_{it}$. This variable is summarized in Table 1. In the empirical analysis, we allow for unexpected changes in $\Delta Fuel_{it}$ to have a differential impact on the market value of each firm, as some firms (e.g., Southwest) aggressively employ fuel-hedging practices.

3.2. Airport Weather Data

To test the robustness of the exogeneity assumptions made in our regression analysis, it is useful to have an instrument for the measures of delays discussed above. Specifically, we want to increase the likelihood that the relationship between market returns and deviations from expected performance that we estimate is a causal one (ruling out unobserved causes such as new management at an airline, which might impact both delays and market value). One important and exogenous cause of delays, which can reasonably be assumed to be uncorrelated with unobserved factors impacting returns for airlines, is the weather. According to the DOT's classification, weather events were responsible for between 5% and 7% of the total minutes of delays annually since 2003. Since most air passengers buy their tickets at least a month before the travel date, weather in any month is also very unlikely to impact returns by affecting demand for air travel in that month, aside from predictable monthly patterns that we control for with month-year fixed effects.

We collected daily weather data from the National Oceanic and Atmospheric Administration's National Climatic Data Center for the top 200 airports in terms of enplanements in 2009. We observe various measures of wind, precipitation, and sun/clouds. To match the unit of observation for the remainder of our data, we aggregate each weather metric to get a carrier-month-specific weighted average. For example, suppose we are constructing a measure of monthly precipitation (measured in millimeters) for carrier i's network in month t. Furthermore, suppose carrier i serves N_{it} airports, and w_{ijt} is the proportion of the carrier's total enplanements in month t that occurred at airport j. Our measure of precipitation is then $Prec_{it} = \sum_j^{N_{it}} w_{ijt} Prec_{jt}$. We construct similar weighted averages for snow fall ($SnowFall_{it}$) and snow depth ($SnowDep_{it}$) in millimeters, sunshine (Sun_{it}) in hours, maximum temperature ($TempMax_{it}$) in degrees Celsius, and the average wind speed ($AvgWind_{it}$) and peak wind speed ($PkWind_{it}$) in meters per second.

Each of these variables is described in Table 1. There is a great deal of variation in weather patterns across carriers' networks, with Southwest and Hawaiian operating networks largely free of extreme weather. In Table 2, we present the correlations between the measures of delays and weather. The strong correlations suggest the weather variables will serve as appropriate instruments for delays in our analysis. We discuss the results of statistical tests measuring the strength of the instruments and the need to instrument in §5.

Table 2 Correlation, Delays and Weather

	Prec	*SnowFall*	*SnowDepth*	*Sun*	*TempMax*	*AvgWind*	*PkWind*
$P(>15)$	0.377	0.352	0.244	−0.170	−0.248	−0.065	−0.173
$P(>120)$	0.269	0.261	0.196	−0.193	−0.217	−0.036	−0.124
$P(>180)$	0.159	0.251	0.198	−0.192	−0.238	−0.026	−0.170
$P(>3\sigma)$	0.132	0.257	0.199	−0.241	−0.182	−0.022	−0.147
$P(Cancel)$	0.091	0.238	0.194	−0.230	−0.161	−0.020	−0.120
$E(Delay)$	0.274	0.218	0.128	−0.125	−0.126	−0.091	−0.379

Table 3 Correlation, Operational and Financial Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. *Ret*	1															
2. *AbnRet*	0.88	1														
3. *MktVal*	0.03	0.04	1													
4. $P(>15)$	−0.06	−0.03	0.03	1												
5. $P(>120)$	−0.11	−0.07	−0.01	0.74	1											
6. $P(>180)$	−0.13	−0.08	−0.01	0.64	0.97	1										
7. $P(>3\sigma)$	−0.12	−0.09	0.01	0.57	0.92	0.97	1									
8. $P(\mathit{Cancel})$	−0.14	−0.09	0.01	0.54	0.91	0.98	0.99	1								
9. Δ*Revpaxmiles*	0.13	0.09	0.03	0.07	−0.14	−0.20	−0.23	−0.24	1							
10. *MissedBags*	−0.05	−0.02	0.16	0.54	0.37	0.31	0.25	0.26	0.06	1						
11. *Oversales*	−0.02	−0.02	0.20	0.16	0.02	0.04	0.04	0.06	0.01	0.23	1					
12. *Strike*	−0.01	−0.01	0.01	0.02	0.07	0.09	0.12	0.10	−0.08	—	—	1				
13. *Slowdown*	−0.03	−0.02	0.24	0.03	0.15	0.17	0.21	0.18	−0.01	0.01	−0.12	−0.01	1			
14. *Bankrupt*	−0.04	−0.03	−0.04	−0.04	−0.01	−0.01	−0.02	−0.01	−0.03	0.01	−0.02	0.00	−0.01	1		
15. Δ*Fuel*	−0.24	−0.27	−0.01	0.02	0.01	0.01	0.01	0.01	−0.03	0.01	0.02	−0.01	−0.02	0.04	1	
16. $E(\mathit{Delay})$	−0.02	0.00	−0.12	0.73	0.61	0.43	0.22	0.25	0.08	0.36	−0.08	0.02	−0.01	−0.02	0.02	1

3.3. Financial Data

Monthly returns for the 11 major airlines for the 1990–2009 sample period were collected from the Center for Research in Security Prices (CRSP) database. We use the residual returns obtained after adjusting for market and other risk factors, which are commonly referred to as "abnormal returns," as the dependent variable in our analysis of stock returns. In our discussion, we focus on those abnormal returns recovered from the four-factor model of Carhart (1997) because our results and conclusions are nearly identical for three alternative models, a three-factor model, market model, and industry model. The three-factor model is that of Fama and French (1993). The market model recovers abnormal returns as the residual from a regression of returns on a constant and the CRSP U.S. Total Market Index. Abnormal returns from the industry model are calculated as the deviation of a carrier's return from a weighted (based on market value) industry average in a month. The details of each model, and a sample of the results, are provided in the online supplement (available at http://dx.doi.org/10.1287/msom.2013.0448).

Between 1990 and 2009, an investor who held a portfolio of the 11 airline stocks consisting of an equal number of shares of each airline would have received a mean return of 0.68% per month (Table 1 reports returns net of the risk-free rate), which is about 8.56% per annum. The investor would have done better if she invested instead in the market portfolio, which had an annual return of 10.10% per annum based on the S&P 500 return over this period. The underperformance of the U.S. airline industry over this period is not surprising, in view of traumatic shocks such as the 9/11 terrorist attacks and the SARS epidemic, and the impact of three recessions (in 1991, 2001, and 2008). These events all contributed to the high volatility of airline stocks, and to some of the extreme returns realized during the sample period.

By focusing our empirical analysis on the abnormal or risk-adjusted returns, $AbnRet_{it}$, we can control for the impact of market conditions and other risk factors on stock price movements. Table 1 describes $AbnRet_{it}$, calculated using the four-factor model. In our regressions relating abnormal returns to operational performance, we include year-month fixed effects to control for shocks to the industry that impact both returns and operational performance. Thus, only the relative operational performance of carriers in response to exogenous shocks (i.e., within year-month variation) is used to explain returns.

3.4. Descriptive Evidence

Table 3 shows the contemporaneous correlations among returns and the operational variables, foreshadowing some of our findings. It also provides direct support for Hypothesis 1, that metrics of operational performance are contemporaneously reflected in the market's valuation of firms, in advance of the release of DOT statistics. In general, the correlation between abnormal returns and the operational variables has the expected sign. Variables $E(Delay)_{it}$, $P(3\sigma)_{it}$, $P(>15)_{it}$, $P(>120)_{it}$, $P(>180)_{it}$, and $P(Cancel)_{it}$ are each negatively correlated with returns; $\Delta Revpaxmiles_{it}$ and $\Delta LoadFactor_{it}$ are positively correlated with returns; and $MissedBags_{it}$ and $Oversales_{it}$ are negatively correlated with returns. The controls, $Strike_{it}$, $Slowdown_{it}$, $Bankrupt_{it}$, and $\Delta Fuel_{it}$, are all negatively correlated with returns, as expected.[2] The goal

[2] The nearly perfect linear relationship among different measures of long delays makes it impossible to simultaneously consider more than one measure describing the tail of the delays distribution in any regression. To mitigate concerns regarding which is the appropriate measure of long delays, we perform our analysis with each of the four measures.

of our regression analysis is to examine whether these descriptive relationships between operational performance and returns remain after carefully controlling for a variety of factors, including the market's expectation of a carrier's operational performance.

4. Methodology

We use a monthly time window to analyze the effects of operating outcomes on stock returns. Unlike a typical event study, we do not examine the impact of specific announcements about operating results on stock prices. Nonetheless, our central empirical method is very much in the spirit of an event study in that we relate monthly stock returns to contemporaneous, unpredicted changes in service quality and operating performance using a panel data set. In this way, our paper is conceptually similar to Bartov et al. (2002) and Kasznik and McNichols (2002).

As discussed in §2, since stock prices respond to *unpredicted* news, for each variable of interest we develop a prediction of current month performance, and then model abnormal stock returns each month as a function of actual and predicted values for each variable.

4.1. Prediction of Operational Variables

Our predictions are intended to mimic investors' predictions. To account for heterogeneity in the way investors form their expectations regarding carriers' future operational performance, we predict each variable separately for each carrier. For example, for carrier i in time t, we model the frequency of delays of over 15 minutes, $P(>15)_{it}$, as

$$P(>15)_{it} = \sum_{k=2}^{5} \phi_{ik} P(>15)_{i(t-k)} + \boldsymbol{\alpha}_i \mathbf{x}_{it} + \xi_{it}^{P(>15)}, \qquad (1)$$

where $P(>15)_{i(t-k)}$ denotes the kth lag of $P(>15)_{it}$. We include lags because if the previous periods' delays are abnormally high or low, investors may revise their expectations accordingly.[3] The vector $\mathbf{x}_{it}$ includes a constant; month dummies; a post-9/11 dummy; linear and quadratic time trends; indicators for strikes, slowdowns, or sick-outs and bankruptcies; and indicators for the month prior to these events. Let $\widehat{P(>15)}_{it}$ denote the predicted or fitted value of $P(>15)_{it}$ from the above regression, and let $\hat{\xi}_{it}^{P(>15)}$ denote the regression residual, which we use as the *unpredicted* portion of the percentage of delays over 15 minutes. Predictions and residuals for the other delay variables, missed bags, and oversales are similarly obtained. Since both revenue passenger miles and load factor exhibit a clear trend, we model the change, rather than level, in these variables from month $t-1$ to t, employing the same functional form above. We model the monthly changes in fuel prices, $\Delta Fuel_t$, excluding the carrier-specific controls. An overview of the results from these regressions and alternative forecasting models is in the online supplement. We find our results to be robust to a number of alternative models, giving us confidence that subtle differences between the model above and the actual model used by investors will not affect our results meaningfully. We also test for serial correlation and find none.

[3] The immediate lag of the operational variable being predicted is omitted, since it is not yet released to the public (delay in collecting and publishing previous months' data) when investors form their expectations.

4.2. Abnormal Returns Regressions

After creating predicted values and residuals for all variables of interest, we regress abnormal returns in each month, $AbnRet_{it}$, against the residuals from the prediction exercise and a full set of controls, pooling across carriers and time. For example, to examine the effect of $P(>15)_{it}$ and $P(>180)_{it}$, recalling that we can only include one measure of long delays at a time, we estimate the following regression:

$$\begin{aligned} AbnRet_{it} = {} & \lambda_1 \hat{\xi}_{it}^{E(Delay)} + \lambda_2 \hat{\xi}_{it}^{P(>15)} + \lambda_3 \hat{\xi}_{it}^{P(>180)} \\ & + \lambda_4 \hat{\xi}_{it}^{\Delta Revpaxmiles} + \lambda_5 \hat{\xi}_{it}^{Bags} + \lambda_6 \hat{\xi}_{it}^{Oversales} \\ & + \sum_i \beta_i \hat{\xi}_{it}^{\Delta Fuel} * 1[Carrier_i] + \boldsymbol{\gamma} \mathbf{z}_{it} + \mu_{it}. \qquad (2) \end{aligned}$$

The $\mathbf{z}_{it}$ vector includes a constant, $Congest_{it}$; indicators for strikes, slowdowns or sick-outs, and bankruptcies; along with indicators for the month prior to these events and year-month dummies.[4] The year-month dummies account for variation in returns and operational performance driven by industry-wide shocks, such as the immediate effect of the 9/11 terrorist attacks and its lingering effects. The carrier-specific β_i coefficients are intended to capture any differential effect, across carriers, in the impact of unexpected changes in fuel prices. For example, Southwest is well known for its aggressive (and successful) fuel-hedging practices, suggesting that it would be less negatively impacted by an unexpected increase in fuel prices. To account for sampling error in the first stage of our analysis, i.e., error in the prediction of operational performance and abnormal returns, we estimate standard errors for the model's parameters using a parametric bootstrap procedure, sampling the first-stage estimates and repeatedly estimating the second stage. The resulting distribution of second-stage estimates is used to calculate standard errors for the estimates in Tables 4–7.

[4] Carrier dummies are not separately identified, because both abnormal returns and the surprises in operational variables have zero mean for each carrier by construction.

Table 4 Abnormal Returns Regressions—Deviations Form

	Long panel				Short panel			
	(1) $P(>120)$	(2) $P(>180)$	(3) $P(Cancel)$	(4) $P(3\sigma)$	(5) $P(>120)$	(6) $P(>180)$	(7) $P(Cancel)$	(8) $P(3\sigma)$
$\xi^{E(Delay)}$	0.002	0.001	0.000	0.001	0.002	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\xi^{P(>15)}$	0.039	0.033	0.039	0.036	0.057	0.042	0.045	0.047
	(0.105)	(0.105)	(0.106)	(0.125)	(0.17)	(0.151)	(0.155)	(0.17)
$\xi^{P(Long)}$	-0.975^{***}	-1.125^{***}	-1.165^{***}	-1.115^{***}	-0.911^{**}	-1.031^{**}	-1.086^{***}	-1.125^{***}
	(0.233)	(0.269)	(0.266)	(0.222)	(0.319)	(0.378)	(0.379)	(0.377)
$\xi^{\Delta Revpaxmiles}$	0.352	0.286	0.257	0.256	0.305	0.248	0.209	0.201
	(1.669)	(1.831)	(1.639)	(1.573)	(1.945)	(1.933)	(1.663)	(1.645)
$\xi^{MissedBags}$					−0.006	−0.006	−0.006	−0.006
					(0.005)	(0.004)	(0.005)	(0.005)
$\xi^{Oversales}$					−0.010	−0.010	−0.010	−0.01
					(0.011)	(0.01)	(0.01)	(0.011)
Strike	−0.054	−0.053	−0.052	−0.051				
	(0.044)	(0.048)	(0.044)	(0.051)				
Prestrike	-0.067^{***}	-0.065^{***}	-0.065^{***}	-0.065^{***}				
	(0.019)	(0.022)	(0.022)	(0.02)				
Slowdown	−0.002	−0.003	−0.003	−0.003	−0.001	−0.002	−0.002	−0.002
	(0.011)	(0.01)	(0.011)	(0.01)	(0.012)	(0.013)	(0.012)	(0.012)
Preslowdown	−0.010	−0.012	−0.012	−0.012	−0.009	−0.010	−0.010	−0.011
	(0.01)	(0.011)	(0.011)	(0.01)	(0.009)	(0.01)	(0.01)	(0.01)
Bankrupt	0.021	0.022	0.020	0.02	0.028	0.028	0.026	0.025
	(0.047)	(0.047)	(0.04)	(0.042)	(0.042)	(0.038)	(0.041)	(0.044)
Prebankrupt	-0.249^{***}	-0.249^{***}	-0.249^{***}	-0.257^{***}	-0.246^{***}	-0.245^{***}	-0.246^{***}	-0.246^{***}
	(0.067)	(0.072)	(0.076)	(0.072)	(0.068)	(0.067)	(0.068)	(0.079)
$\xi^{\Delta Fuel} * 1(AA)$	-0.316^{***}	-0.317^{***}	-0.317^{***}	-0.331^{***}	-0.360^{***}	-0.361^{***}	-0.362^{***}	-0.347^{***}
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)
$\xi^{\Delta Fuel} * 1(AS)$	-0.169^{***}	-0.170^{***}	-0.171^{***}	-0.166^{***}	-0.176^{***}	-0.177^{***}	-0.178^{***}	-0.169^{***}
	(0.004)	(0.003)	(0.003)	(0.004)	(0.007)	(0.006)	(0.006)	(0.006)
$\xi^{\Delta Fuel} * 1(B6)$	-0.114^{***}	-0.116^{***}	-0.114^{***}	-0.113^{***}	-0.151^{***}	-0.153^{***}	-0.151^{***}	-0.152^{***}
	(0.023)	(0.025)	(0.024)	(0.022)	(0.019)	(0.02)	(0.021)	(0.019)
$\xi^{\Delta Fuel} * 1(CO)$	−0.176	−0.175	−0.177	−0.185	−0.137	−0.138	−0.139	−0.145
	(0.124)	(0.137)	(0.12)	(0.132)	(0.141)	(0.137)	(0.146)	(0.139)
$\xi^{\Delta Fuel} * 1(DL)$	-0.366^{***}	-0.364^{***}	-0.364^{***}	-0.349^{***}	-0.422^{***}	-0.420^{***}	-0.420^{***}	-0.415^{***}
	(0.025)	(0.028)	(0.024)	(0.026)	(0.034)	(0.031)	(0.033)	(0.036)
$\xi^{\Delta Fuel} * 1(FL)$	-0.083^{***}	-0.083^{***}	-0.083^{***}	-0.081^{***}	-0.122^{***}	-0.122^{***}	-0.122^{***}	-0.12^{***}
	(0.022)	(0.022)	(0.022)	(0.021)	(0.022)	(0.02)	(0.021)	(0.021)
$\xi^{\Delta Fuel} * 1(HA)$	-0.057^{***}	-0.058^{***}	-0.058^{***}	-0.06^{***}	-0.089^{***}	-0.090^{***}	-0.091^{***}	-0.093^{***}
	(0.02)	(0.019)	(0.017)	(0.021)	(0.019)	(0.02)	(0.018)	(0.019)
$\xi^{\Delta Fuel} * 1(NW)$	0.071	0.068	0.067	0.066	0.040	0.038	0.037	0.039
	(0.083)	(0.085)	(0.084)	(0.072)	(0.08)	(0.083)	(0.071)	(0.071)
$\xi^{\Delta Fuel} * 1(UA)$	-0.520^{***}	-0.521^{***}	-0.522^{***}	-0.509^{***}	-0.581^{***}	-0.582^{***}	-0.583^{***}	-0.579^{***}
	(0.026)	(0.025)	(0.025)	(0.025)	(0.026)	(0.029)	(0.027)	(0.025)
$\xi^{\Delta Fuel} * 1(US)$	-0.538^{***}	-0.538^{***}	-0.536^{***}	-0.531^{***}	-0.609^{***}	-0.609^{***}	-0.607^{***}	-0.612^{***}
	(0.017)	(0.019)	(0.018)	(0.021)	(0.024)	(0.021)	(0.022)	(0.023)
R^2	0.473	0.474	0.474	0.483	0.491	0.491	0.491	0.506
Adjusted R^2	0.378	0.379	0.379	0.377	0.397	0.397	0.397	0.379
No. of obs.	1,636	1,636	1,636	1,636	1,017	1,017	1,017	1,017

Notes. Year-month-specific dummies; the interaction of $\xi^{\Delta Fuel}$ with carrier dummies, *Congest*, and dummies for the period before and period of strikes, slowdowns, and bankruptcies are included in all regressions. AA, American Airlines; AS, Alaska Airlines; B6, JetBlue; CO, Continental; DL, Delta; FL, AirTran; HA, Hawaiian; NW, Northwest; UA, United; US, US Airways; WN, Southwest.

**Significance level is 0.05; ***significance level is 0.01.

Table 5 Abnormal Returns Regressions Expanded Form

	(1) $P(>120)$	(2) $P(>180)$	(3) $P(Cancel)$	(4) $P(3\sigma)$
$E(Delay)$	0.002	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$E(Delay)^{Forecast}$	−0.001	−0.001	−0.000	−0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$P(>15)$	0.040	0.035	0.040	0.039
	(0.115)	(0.123)	(0.123)	(0.123)
$P(>15)^{Forecast}$	−0.040	−0.036	−0.044	−0.047
	(0.171)	(0.179)	(0.174)	(0.174)
$P(Long)$	−0.988***	−1.139***	−1.174***	−1.104***
	(0.254)	(0.24)	(0.249)	(0.191)
$P(Long)^{Forecast}$	0.899*	1.030*	1.108***	1.177***
	(0.461)	(0.495)	(0.507)	(0.494)
$\Delta Revpaxmiles$	0.359	0.295	0.263	0.263
	(1.575)	(1.679)	(1.609)	(1.704)
$\Delta Revpaxmiles^{Forecast}$	−0.416	−0.354	−0.315	−0.315
	(1.878)	(1.902)	(1.935)	(1.876)
F-test, p-value	0.973	0.968	0.976	0.975
R^2	0.483	0.484	0.487	0.485
Adjusted R^2	0.376	0.377	0.377	0.378
No. of obs.	1,636	1,636	1,636	1,636

Notes. Year-month-specific dummies, the interaction of $\xi^{\Delta Fuel}$ with carrier dummies, *Congest*, and dummies for the period before and period of strikes, slowdowns, and bankruptcies are included in all regressions.

*Significance level is 0.1; ***significance level is 0.01.

We also regress abnormal returns against each variable and its predicted value separately, rather than capturing the difference by the residual:

$$\begin{aligned} AbnRet_{it} = {} & \delta_1 E(Delay)_{it} + \delta_1^P \widehat{E(Delay)}_{it} + \delta_2 P(>15)_{it} \\ & + \delta_2^P \widehat{P(>15)}_{it} + \delta_3 P(>180)_{it} + \delta_3^P \widehat{P(>180)}_{it} \\ & + \delta_4 \Delta Revpaxmiles_{it} + \delta_4^P \Delta \widehat{Revpaxmiles}_{it} \\ & + \sum_i \beta_i \hat{\xi}_{it}^{\Delta Fuel} * 1[Carrier_i] + \boldsymbol{\gamma} \mathbf{z}_{it} + \mu_{it}. \qquad (3) \end{aligned}$$

The controls in Equation (3) are identical to those in Equation (2).[5] Estimates from Equation (3) provide a direct test of the assumptions imposed on Equation (2). In particular, if the market prices in its expectation of a carrier's performance for some operational metric, it would require a realization of this metric above or below the prediction to result in variation in the stock's price. In the case for which expectations are fully priced into the market, we would expect equal and opposite signs on realized and expected performance metrics. For example, suppose in a particular month that carriers A and B have 15% of their flights delayed by 180 minutes or more. Furthermore, suppose the market had different expectations going into the period regarding the carriers' performance, 5% and 15% late, for carriers A and B, respectively. In this case, if expectations regarding operational performance are priced into the market, only carrier A should see a drop in market value as a result of its performance. For this to be the case, there must be equal and opposite signs on realized and expected performance. In §5, we discuss the results of an F-test that shows that one cannot reject the assumptions imposed on Equation (2). We also discuss the robustness of our results to a relaxation of the different parametric and exogeneity assumptions made above.

[5] As we discuss in §5, we estimate this regression on the long panel since we consistently find no relationship between abnormal returns and both missed bags and oversales.

4.3. Factors That Alter the Impact of Delays on Returns

We have posited a number of factors that may alter the impact of delays: average haul length, degree of connectivity, aircraft load factors, the relative performance of competitors, and the proportion of delays attributable to the carrier. To test these hypotheses, we estimate regressions similar to Equation (2) that include interactions of delays with these factors. Because of multicollinearity issues associated with a large number of interaction terms (see Angrist and Pischke 2008), we cannot jointly test these five hypotheses and instead test each separately.

For example, we estimate the regression

$$\begin{aligned} AbnRet_{it} = {} & (\lambda_1 + \lambda_1^+ \widetilde{Haul}_{it}) \hat{\xi}_{it}^{E(Delay)} + (\lambda_2 + \lambda_2^+ \widetilde{Haul}_{it}) \hat{\xi}_{it}^{P(>15)} \\ & + (\lambda_3 + \lambda_3^+ \widetilde{Haul}_{it}) \hat{\xi}_{it}^{P(>180)} + \lambda_4 \hat{\xi}_{it}^{\Delta Revpaxmiles} \\ & + \sum_i \beta_i \hat{\xi}_{it}^{\Delta Fuel} * 1[Carrier_i] + \boldsymbol{\gamma} \mathbf{z}_{it} + \mu_{it} \qquad (4) \end{aligned}$$

to test whether operating a network with more short-haul routes exacerbates the financial impact of delays. The variable $\widetilde{Haul}_{it}$ is a normalization of $Haul_{it}$, between zero and one, derived by subtracting the minimum in the sample from each observation and dividing by the support of the variable in the data (difference between the min and max). This significantly simplifies the interpretation of the interaction terms. The λ_3 coefficient measures the impact of long delays for the carrier with the lowest average haul length, whereas $\lambda_3 + \lambda_3^+$ gives the impact of delays for the carrier with the highest average haul length during our sample period. To be consistent with Hypothesis 3, λ_3^+ must be greater than zero so that long-haul carriers are less impacted when performing worse than expected (i.e., $\hat{\xi}_{it}^{P(>180)} > 0$). We also estimate the Equation (4) regression replacing $\widetilde{Haul}_{it}$ with $\widetilde{Connect}_{it}$ and $\widetilde{Load}_{it}$. In these regressions, we expect the coefficient of the corresponding interaction term to be less than zero if carriers with a greater degree of connectivity in their network and those operating aircraft at higher load factors are penalized more when performing worse than expected.

Table 6 Sensitivity Analysis

	Nonlinearity		Weather-IV	Cancellations		Alternative abnormal returns measures		
	(1)		(2)	(3)	(4)	(5)	(6)	(7)
	Direct	Interaction		180 minutes	300 minutes	Industry	Market basket	Three-factor
$\xi^{E(Delay)}$	0.002	−0.002	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
$\xi^{P(>15)}$	−0.019	0.143	0.082	0.070	0.081	0.032	0.031	0.036
	(0.137)	(0.203)	(0.083)	(0.097)	(0.089)	(0.141)	(0.12)	(0.1)
$\xi^{P(>180)}$	−1.017***	−0.086	−1.213***	−1.259**	−1.275*	−1.258**	−1.175***	−1.125***
	(0.373)	(0.389)	(0.431)	(0.534)	(0.689)	(0.356)	(0.322)	(0.548)
$\xi^{\Delta Revpaxmiles}$	1.889**	−2.014	0.504	0.328	0.334	0.310	0.327	0.319
	(0.656)	(1.803)	(1.684)	(1.644)	(1.715)	(0.708)	(1.854)	(1.64)
R^2	0.485		0.478	0.474	0.475	0.463	0.460	0.468
Adjusted R^2	0.379		0.367	0.379	0.379	0.373	0.367	0.379
No. of obs.	1,636		1,636	1,636	1,636	1,636	1,636	1,636

Notes. Year-month-specific dummies, the interaction of $\xi^{\Delta Fuel}$ with carrier dummies, *Congest*, and dummies for the period before and period of strikes, slowdowns, and bankruptcies are included in all regressions.
*Significance level is 0.1; **significance level is 0.05; ***significance level is 0.01.

The performance of a carrier's competitors is another potential factor that could alter the impact of delays on returns. For each measure of delays, we construct an index to measure the performance of a carrier's competitors. For example, the index for long delays of carrier i's competitors in month t is

$$Comp_{it}^{P(>180)} = \sum_{k\neq i} w_{ikt} \hat{\xi}_{kt}^{P(>180)}.$$

The weights, $\sum_{k\neq i} w_{ikt} = 1$ for each (i, t), are calculated to reflect the degree to which carrier i and k compete, or the number of common nonstop routes the pair of carriers serve (in the DB1B data). For example, suppose that carrier i serves 100 total nonstop routes and faces two different competitors on these routes, which serve 50 and 75 of the 100 routes, respectively. The weights for the competitors are then 0.4 and 0.6, respectively. Similar to the other factors moderating the impact of delays, we utilize a normalized variant of the index, $\widetilde{Comp}_{it}^{P(>180)} \in [0, 1]$. However, there is an important difference. We expect that when a carrier performs worse than expected (i.e., $\hat{\xi}_{it}^{P(>180)} > 0$) and a competitor does poorly (i.e., $\widetilde{Comp}_{it}^{P(>180)} \approx 1$), the penalty will be lessened, and when a carrier performs better than expected (i.e., $\hat{\xi}_{it}^{P(>180)} < 0$) and a competitor does poorly, the reward will be increased. This requires accounting for whether the carrier performed better or worse than expected. For long delays,

Table 7 Differential Impact of Delays

	Haul length (1)		Connectivity (2)		Load factor (3)		Competitors' performance (4)		Source of delays (5)	
	Direct	Interaction	Direct	Interaction	Direct	Interaction	Direct	Interaction	Direct	Interaction
$\xi^{E(Delay)}$	0.002	−0.003	0.002	−0.002	0.002	0.001	0.002	−0.000	0.002	−0.001
	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)	(0.004)	(0.002)	(0.004)
$\xi^{P(>15)}$	0.015	0.019	−0.018	0.010	0.025	0.027	−0.012	0.000	−0.016	0.015
	(0.298)	(0.236)	(0.272)	(0.289)	(0.233)	(0.257)	(0.284)	(0.313)	(0.277)	(0.289)
$\xi^{P(>180)}$	−1.217***	−0.589***	−1.144***	−0.351***	−1.169***	−0.181	−1.134***	−0.301***	−1.281***	−0.275
	(0.428)	(0.24)	(0.49)	(0.17)	(0.428)	(0.137)	(0.546)	(0.12)	(0.531)	(0.568)
$\xi^{\Delta Revpaxmiles}$	0.403		0.403		0.403		0.403		0.403	
	(2.431)		(2.574)		(2.054)		(2.037)		(2.131)	
R^2	0.53		0.536		0.53		0.504		0.518	
Adjusted R^2	0.383		0.389		0.387		0.372		0.379	
No. of obs.	1,636		1,636		1,636		1,636		936	

Notes. Year-month-specific dummies, the interaction of $\xi^{\Delta Fuel}$ with carrier dummies, *Congest*, and dummies for the period before and period of strikes, slowdowns, and bankruptcies are included in all regressions.
***Significance level is 0.01.

we do this by interacting $\hat{\xi}_{it}^{P(>180)}$ with the product of $\widetilde{Comp}_{it}^{P(>180)}$ and $\text{sign}(\hat{\xi}_{it}^{P(>180)})$, where

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0, \end{cases}$$

rather than just $\widetilde{Comp}_{it}^{P(>180)}$. We expect the coefficients on these interactions to be negative for each measure of delays.

Similarly, to test whether the source of delay is an important moderator of the impact of delays, we interact each measure of delays with the product of the sign function and $\widetilde{Attrib}_{it}$. For example, in the case of long delays, the interaction term is the product of $\hat{\xi}_{it}^{P(>180)}$, $\text{sign}(\hat{\xi}_{it}^{P(>180)})$, and $\widetilde{Attrib}_{it}$. We expect that the coefficients on these interactions will be negative, such that when a carrier is responsible for a larger proportion of delays and performs worse (better) than expected, the financial penalty (reward) will be increased (decreased).

5. Empirical Results

5.1. Impact of Operational Variables on Stock Returns

Because our hypotheses are derived on the presumption that financial markets are efficient, our empirical tests investigate jointly the impact of operational variables on abnormal returns and the ability of financial markets to aggregate and process information. Table 4 reports our estimates of Equation (2). Since we have data on mishandled bags and denied boardings from only 1998 to 2009, specifications (5)–(8), which include these variables, are run on this shorter panel, whereas specifications (1)–(4) are run on the long panel, 1990–2009.

Columns (1)–(4) report estimates of Equation (2), varying the measure of long delays used. The one operational measure that has a statistically and economically significant impact on abnormal returns is long delays. The reported results are nearly identical for each measure of long delays, including others that are not reported (e.g., $P(2\sigma)_{it}$). The coefficient estimates on the measures of long delays can be interpreted as the marginal change in returns when the unpredicted portion of the percentage of long delays goes up, keeping the surprise in both the total minutes of delay and the percentage of flights delayed over 15 minutes constant. This finding provides support for Hypothesis 2, that the marginal cost of a minute of delay is greater for long delays. Thus, although the most prominently reported statistic in the DOT's *Air Travel Consumer Report* is on-time performance, using a 15-minute delay as the on-time threshold, our results suggest that the financial market assigns a reduction in market value only to long delays or cancelations. Using an F-test we could reject the null that the coefficients of long delays and on-time performance were equal at the 1% level, appropriately standardizing each variable to account for differences in scale. The high variation in on-time performance, lack of market response to release of DOT statistics, and robustness to different modeling assumptions gives us confidence in this finding.

The signs of the coefficients on the controls are largely as expected in columns (1)–(4). The coefficients on the indicators for the month immediately before a strike and bankruptcy are negative and statistically significant. The interactions of surprises in jet fuel prices and carrier dummies reveal that all but two carriers (Continental and Northwest) performed significantly worse than Southwest, the omitted carrier, in insulating themselves from variability in jet-fuel prices over the sample period. Our estimates in columns (5)–(8) of Table 4, using the short panel, provide no evidence that lost baggage and oversales are significant drivers of returns. This may be due to the market's inability (or lack of incentive) to infer the costs associated with missing expectations for operational metrics that impact so few consumers. The remainder of our discussion focuses on the results using the long panel, as we consistently find no relationship between abnormal returns and both missed bags and oversales.

In Table 5, we report our estimates of Equation (3). As we discuss in §2, if our assumption that expectations regarding operational performance are fully incorporated into a stock's price at the beginning of the month is correct, then we would expect equal and opposite signs on realized and expected performance for each operational dimension. The results (p-value) from an F-test of the linear restrictions imposed on the coefficients of the operational variables in Equation (2) are presented at the bottom of Table 5. We cannot reject the restrictions imposed by Equation (2), and our results are very similar to those in Table 4. This result provides strong support for Hypothesis 1, that the market contemporaneously incorporates information regarding delays. For each measure of long delays, we again find a statistically significant effect on returns.

Table 6 presents the results of our sensitivity analysis. We present only the results for $\hat{\xi}_{it}^{P(>180)}$ since the different measures of long delays have a near-perfect linear relationship. Column (1) examines the sensitivity of our results to outliers, to ensure that the results in Tables 4 and 5 are not driven by a couple of instances of extremely poor operational performance. To this end, we reestimate Equation (2), including interactions of surprises in operational performance with indicators for whether operational performance along that dimension fell two standard deviations below expectations. If our results are driven by these extreme events, only the interactions would have explanatory power and be statistically significant. We find no statistically significant coefficients on the interactions.

Column (2) of Table 6 presents the results when we instrument for each measure of delays using the weather instruments. Our results are comparable to the analogous results in column (2) of Table 4. Regarding the efficiency of the instrumental variable approach and need to instrument at all, we cannot reject (even at the 10% level) the null that our delay measures are exogenous using a Wu–Hausman test.[6] Thus, ordinary least squares will be the efficient estimator and our focus for the remainder of the paper. Columns (3) and (4) of Table 6 provide a test of the sensitivity of our results to different assumptions regarding the average delay imposed before a flight is cancelled. The DOT does not report minutes of delay for flights that are eventually cancelled. For this reason, we cannot include these minutes of delay in our calculation of $E(Delay)$, which would result in a downward bias in $E(Delay)$. To test whether this potential source of measurement error in $E(Delay)$ has any meaningful impact on our results, we make varying assumptions regarding the average delay before a cancellation, recalculate $E(Delay)$, and reestimate Equation (2). Columns (3) and (4) report the results when we assume 180 and 300 minutes of delay are incurred prior to a cancellation, respectively. Again, we find very similar results to those in Table 4, suggesting the impact of any bias in $E(Delay)$ is at most a minor concern.

Next, consider the magnitude, in terms of market value, associated with unpredicted changes in long delays. We focus on the coefficient estimate on $\hat{\xi}_{it}^{P(>180)}$ in column (2) of Table 4, because it lies in the middle of the range of our estimates. A surprise of 0.02 (one standard deviation) in the proportion of delays over 180 minutes, holding the surprise element in both total minutes of delay and the percentage of flights that are over 15 minutes late constant, results in a $0.02 * (-1.125) = -2.25\%$ reduction in market value. In January 2009, we estimate the total market value of equity for our 11 carriers as $16,707 million, so a surprise one standard deviation above expectations in long delays results in a reduction in the market value of equity of $375.91 million.

As mentioned earlier, the ATA estimated the direct out-of-pocket average costs as $61 per minute of delay in 2009. Barnett et al. (2001) also report several estimates of out-of-pocket costs of delay, of the same order of magnitude, and note that these costs appear to grow linearly with delay length. We are interested in how our estimate of the loss in market value associated with an additional minute of delay complements the ATA's estimate of out-of-pocket average costs. There are two important reasons why the two estimates might diverge. First, our estimate incorporates opportunity costs of delays (e.g., lost future revenues from disgruntled passengers) in addition to out-of-pocket expenses. Second, the ATA estimates average cost, whereas our estimates allow us to examine the relationship between the marginal cost of a minute of delay and the length of delay.

With this in mind, we proceed as follows. Assume one holds constant both the total minutes of delays and the percentage of flights over 15 minutes late, and increases the proportion of flights over 180 minutes late by one standard deviation (0.02). This is equivalent to a mean-preserving spread of the conditional distribution of delays greater than 15 minutes in length. One such mean-preserving spread would result if for every 32 flights with delays of 24 minutes (which is the average nonnegative delay in our data, a conservative starting point) for which one decreased delays by five minutes, one increased a single flight's 24-minute delay by 160 minutes (pushing it over the 180-minute threshold). If we take 2% of all the flights operated by our carriers in January 2009 (i.e., 10,647 flights) and increase each of their delays by 160 minutes to cross the 180-minute threshold while reducing the flying times of other flights as above to obtain a mean-preserving spread, delays that are lengthened by 160 minutes would incur $10{,}647 * 160 = 1{,}703{,}520$ additional minutes of delay, under our assumption about which flights were further delayed. Since we find no measurable benefit to carriers of reducing delays of mild length, such a mean-preserving spread would result in a $375.91 million dollar loss to the industry, as we calculated above.

An estimate of average cost-per-minute of delay, a conservative estimate of the marginal cost, would be approximately $220 per minute. This is similar to the $279 per minute estimate of Ball et al. (2010) for 2007, which we calculate based on their estimates as the total costs of delays ($31.2 billion) divided by the total minutes of delays (111,831,870). One might use the ATA's estimate of the per-minute average cost of delay to measure the benefit to the carrier of reducing the short delays, resulting in a more conservative estimate of $159 per minute (220 − 61) for long delays. This is still nearly 2.5 times the ATA's out-of-pocket average cost estimate. The difference between our estimate and the ATA's suggests that if an airline's management were to ignore the full costs associated with delays and the fact that the marginal cost

[6] The first-stage regressions each have an adjusted R^2 over 0.72, and the instruments pass the weak-instrument test of Cragg and Donald (1993).

of delay is higher for long delays, it would be misled as to the optimal distribution of delays.

5.2. Factors That Alter the Impact of Delays

We hypothesized how a number of factors would alter the impact of delays on market value. The estimates of Equation (4) for each factor are presented in Table 7. We focus on a single measure of long delays, $\hat{\xi}_{it}^{P(>180)}$, since the results are very similar for all.

Column (1) of Table 7 presents the results for whether haul length exacerbates the impact of delays. We find a statistically and economically significant coefficient on the interaction of long delays and haul length, $\widetilde{Haul}_{it}$. We estimate the impact of delays for the longest-haul carriers in our sample ($\widetilde{Haul}_{it} = 1$) to be approximately half that of the shortest-haul carrier ($\widetilde{Haul}_{it} = 0$), $-1.217 + 0.589 = -0.628$ and -1.217, respectively. This result is consistent with a higher elasticity of demand with respect to delays on short-haul routes where there are better alternatives to air travel; see Berry and Jia (2010). Column (2) of Table 7 presents the results when we interact $\widetilde{Connect}_{it}$ with the measures of delays to measure the degree to which connectivity in a carrier's network amplifies the impact of delays. We find that the impact of delays for the most connection-intensive network in our data is 30.7% higher than that for the least intensive, and this effect is statistically significant. This difference reflects the direct costs of accommodating passengers who have missed connections along with any reductions in future demand from disgruntled passengers. Column (3) of Table 7 presents the results when delays are interacted with load factor, $\widetilde{Load}_{it}$. We find no statistically significant evidence that carriers operating at lower load factors can mitigate the financial impact of any types of delays by more easily accommodating passengers on subsequent flights to the same destination.

Mazzeo (2003) and others have demonstrated that competitive pressures induce airlines to reduce delays. Column (4) of Table 7 presents our results measuring the variation in the impact of delays due to the performance of a carrier's competition. We find that when a carrier's competition performs worse, or $\widetilde{Comp}_{it}^{P(>180)}$ increases, the loss (reward) for performing worse (better) than expected is significantly decreased (increased). For example, when a carrier's competition performs very poorly (i.e., $\widetilde{Comp}_{it}^{P(>180)} = 1$) rather than very well (i.e., $\widetilde{Comp}_{it}^{P(>180)} = 0$), the reward for the carrier to performing better than expected (i.e., $\hat{\xi}_{it}^{P(>180)} < 0$) increases about 26.5% (0.301/1.134).

The results for whether the source of delays, attributable to the carrier or not, is an important determinant of the financial impact of delays are presented in column (5) of Table 7. This effect is measured through the interaction of each measure of delay with $\widetilde{Attrib}_{it}$ and $\operatorname{sign}(\cdot)$. Interestingly, we find that none of these interactions are statistically significant, although the point estimate of the interaction with long delays has the correct sign. One explanation for this null result is consumers' inability to infer the source of delay, which would limit switching to other carriers or away from air travel altogether based on this information. Given that those causes of delays readily observable to passengers, i.e., weather, represent only 5%–7% of all delays, this seems plausible.

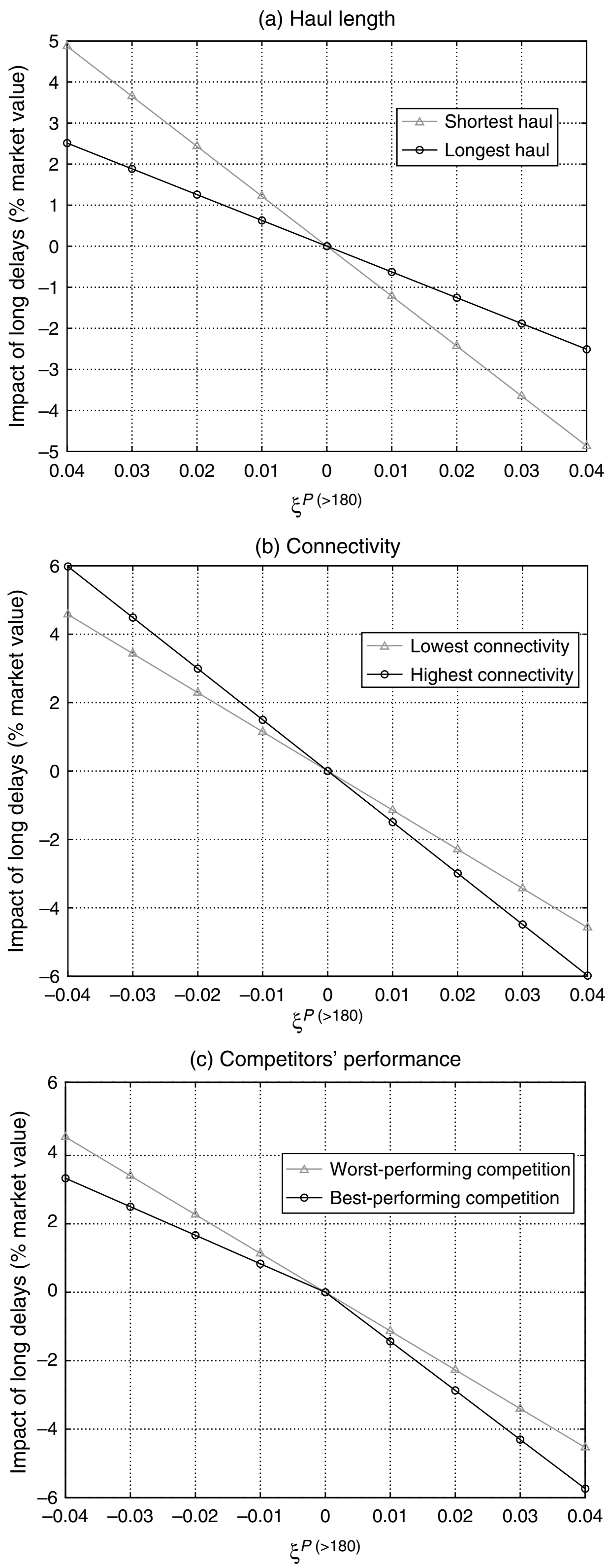
Figures 2(a)–2(c) summarize our results by plotting the difference in the marginal impact of missing expectations on long delays for the three statistically significant moderators: haul length, connectivity, and competitors' performance.

6. Discussion and Conclusions

Information is now widely recognized as a crucial source of competitive advantage. We suggest that managers may be able to use stock price reactions to provide insight into appropriate operating decisions. Our particular approach has some important advantages for our application and can be viewed as a complement to more typical event studies that look at information releases at very specific times. When information is diffuse, yet readily available to consumers and investors, and there is no one *event* to identify the impact of operations on market value, one must rely on broader covariation between returns and performance (e.g., Roll 1984, Bartov et al. 2002). A market-based approach is also advantageous when the full costs of operational performance (e.g., disgruntled passengers switching to a competitor) may not be realized for many months into the future. In these situations, identification of the impact of operational performance on firm performance from published financial statements is confounded because the current period's performance may be reflected in numerous future statements.

Our results suggest that airline managers should focus on eliminating long delays and cancellations, even at the cost of modest increases in short delays. Deshpande and Arıkan (2012) note that increasing scheduled flight times will decrease the percentage of late flights. Clearly, doing this will have less impact on long delays and cancellations. On the other hand, there are a number of alternative strategies available to carriers that target long delays specifically. Our conversations with operation managers and industry professionals revealed a few ways in which airlines limit the frequency of long delays: ensuring that aircraft of the same type depart at similar times to increase opportunities to substitute aircraft and crews, increasing access to boarding gates and other essential airport facilities, and limiting the possibility of delay

Figure 2 **Factors Altering the Impact of Long Delays on Market Value**



propagation by isolating delays to those parts of the network where they are most likely to occur. Atkinson et al. (2012) study the first two of these approaches and find both to be very effective for reducing long delays. A strategy of substituting crews and aircraft is often used to parse one long delay—e.g., due to an aircraft incurring a mechanical failure—into many short delays. Obviously, this approach is only appropriate if the marginal cost of a minute of long delay exceeds that of a minute of short delay. The effectiveness of these strategies, along with our finding that long delays disproportionately impact market value, suggests that the development of more ambitious scheduling algorithms that explicitly account for opportunities for recovery following a disruption is a worthwhile pursuit for operations research modelers.

Our regression-based approach also allows for identification of carrier-specific factors that alter the impact of delays. We find that short-haul carriers face nearly twice as high a penalty for long delays, consistent with a high elasticity of demand with respect to delays on routes with more attractive alternatives to air travel. An increased degree of connectivity in a carrier's network also significantly increases the costs of delays, due to the additional costs associated with accommodating passengers who miss connections and greater customer dissatisfaction that could negatively affect future buying behavior. Finally, we show that the performance of a carrier's competitors can significantly alter the impact of long delays, as a carrier that misses expectations for long delays will be penalized less if its competition does the same, and more if its competition performs better than expected in the same time period.

Our results are not without qualifications. For one, to the extent that our predictions for operational variables differ from those of market participants, the unpredicted residuals used in our returns regressions are measured with error. To the extent that this measurement error is random, our coefficient estimates would be biased toward zero. Since we have found economically and statistically significant estimates even in the presence of the potential downward biases, we are confident that our results and their implications are likely valid at least qualitatively. To the extent that the market does not reflect the full impact of operational variables on value, which may be the case if it imperfectly or partially aggregates available information, our estimates of the effect of these variables will be attenuated. Both of these qualifications imply that our results are an underestimate of the impact of operational variables on firm value.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2013.0448.

Acknowledgments

The authors are grateful to Marianne Corbishley, Haiyan Liu, Paul Landefeld, Ashima Sood, and Vaidheeshwaran Venkatram for their assistance with data collection. For useful comments, the authors thank Gérard Cachon, Federico Ciliberto, Sanjay Jain, and Vinod Singhal. Stuart Thomas, Lonnie Hurwitz, and Alex Heinold of Southwest Airlines and Liying Gu of the Airports Council International–North America provided valuable industry insight that significantly improved this paper.

References

Aaker DA, Jacobson R (1994) The Financial information content of perceived quality. *J. Marketing* 58:191–201.

Abarbanell J, Bushee B (1997) Fundamental analysis, future earnings, and stock prices. *J. Accounting Res.* 35:1–24.

Anderson SW, Baggett LS, Widener SK (2009) The impact of service operations failures on customer satisfaction: Evidence on how failures and their source affect what matters to customers. *Manufacturing Service Oper. Management* 11:52–69.

Anderson SW, Pearo LK, Widener SK (2005) Different strokes for different folks: Linking customer satisfaction to the service concept and customer characteristics. Working paper, University of California, Davis, Davis.

Angrist J, Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

Arıkan M, Deshpande V, Sohoni M (2013) Building reliable air-travel infrastructure using empirical data and stochastic models of airline networks. *Oper. Res.* 61:45–64.

Atkinson S, Ramdas K, Williams J (2012) An empirical investigation into the tradeoffs that impact on-time performance in the airline industry. Working paper, University of Georgia, Athens.

Ball M, Barnhart C, Dresner M, Hansen M, Neels K, Odoni A, Peterson E, Sherry L, Trani A, Zou B (2010) Total delay impact study: A comprehensive assessment of the costs and impacts of flight delay in the United States. Working paper, National Center of Excellence for Aviation Operations Research, University of Maryland, College Park.

Barnett A, Shumsky R, Hansen M, Odoni A, Gosling G (2001) Safe at home? An experiment in domestic airline security. *Oper. Res.* 49:181–195.

Bartov E, Givoly D, Haynes C (2002) The rewards to meeting or beating earnings expectations. *J. Accounting Econom.* 33:173–204.

Berry S, Jia P (2010) Tracing the woes: An empirical analysis or the airline industry. *Amer. Econom. J.: Microeconomics* 2:1–43.

Bratu S, Barnhart C (2005) An analysis of passenger delays using flight operations and passenger booking data. *Air Traffic Control Quart.* 13:1–27.

Cannon J, Randall T, Terwiesch C (2008) Evaluating operations-based data in earnings prediction: A study of the US airline industry. Working paper, University of Utah, Salt Lake City.

Carhart M (1997) On persistence in mutual fund performance. *J. Finance* 52:57–82.

Ciliberto F, Schenone C (2012) Are the bankrupt skies the friendliest? *J. Corporate Finance* 18:1217–1231.

Cragg J, Donald S (1993) Testing the identifiability and specification in instrumental variable models. *Econometric Theory* 9:224–240.

Deshpande V, Arıkan M (2012) The impact of airline flight schedules on flight delays. *Manufacturing Service Oper. Management* 14:423–440.

Fama E (1970) Efficient capital markets: A review of theory and empirical work. *J. Finance* 25:383–417.

Fama E, French K (1993) Common risk factors in the returns on stocks and bonds. *J. Financial Econom.* 33:3–56.

Ferguson J, Kara A, Hoffman K, Sherry L (2012) Estimating domestic US airline cost of delay based on European model. *Transportation Res. Part C: Emerging Technologies*, ePub ahead of print January 11, http://dx.doi.org/10.1016/j.trc.2011.10.003.

Folkes VS, Koletsky S, Graham JL (1987) A field study of causal inferences and consumer reaction: The view from the airport. *J. Consumer Res.* 13:534–539.

Forbes S (2008) The effect of service quality and expectations on customer complaints. *J. Indust. Econom.* 56:190–213.

Forbes S (2008) The effect of air traffic delays on airline prices. *Intl. J. Indust. Organ.* 26:1218–1232.

Forbes S, Lederman M, Tombe T (2011) Do firms game quality ratings? Evidence from mandatory disclosure of airline on-time performance. Working paper, Case Western Reserve University, Cleveland.

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

Hamilton VL (1980) Intuitive psychologist or intuitive lawyer? Alternative models of the attribution process. *J. Personality Soc. Psych.* 39:767–772.

Hauser R (2002) A matter of priority. Presentation, Booz-Allen and Hamilton, On Time Performance Conference, London.

Hendricks KB, Singhal VR (2003) The effect of supply chain glitches on shareholder value. *J. Oper. Management* 21:501–522.

Hendricks KB, Singhal VR (2005) An empirical analysis of the effect of supply chain disruptions on long run stock price performance and risk of the firm. *Production Oper. Management* 14:35–52.

Kasznik R, McNichols MF (2002) Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *J. Accounting Res.* 40:727–759.

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

Maloney M, Mulherin H (2003) The complexity of price discovery in an efficient market: The stock market reaction to the challenger crash. *J. Corporate Finance* 9:453–479.

Mazzeo M (2003) Competition and service quality in the U.S. airline industry. *Rev. Indust. Organ.* 22:275–296.

Roll R (1984) Orange juice and weather. *Amer. Econom. Rev.* 74:861–880.

Rubin D, Talarico J (2003) Confidence, not consistency, characterizes flashbulb memories. *Psych. Sci.* 14:455–461.

Rupp N, Holmes G (2006) An investigation into the determinants of flight cancellations. *Economica* 73:749–783.

Russon MG, Bowers JP (1990) Convenience, distance decay, and capacity of service in a short-haul market for air transportation. *Transportation J.* 30:36–43.

Shumsky R (1993) Response of U.S. air carriers to on-time disclosure rule. *Transportation Res. Record* (1379):9–16.

U.S. General Accounting Office (2003) Airline labor relations: Information on trends and impact of labor actions. Report, U.S. General Accounting Office, Washington, DC.

Williams JW (2012) Capacity investments, exclusionary behavior, and welfare: A dynamic model of competition in the airline industry. Working paper, Terry College of Business, University of Georgia, Athens.