Managing Customer Outrage: Focus Organizational Learning Efforts on Service Failure or Recovery?

Michael A. Lapré
Vanderbilt University, Owen Graduate School of Management, Nashville, TN 37203,
michael.lapre@owen.vanderbilt.edu

As service failures are inevitable, firms must be prepared to recover from service failures, thereby turning angry, frustrated customers into loyal customers. Despite the compelling economics of customer loyalty, firms continue to struggle with service recovery. Should firms focus organizational learning efforts on reducing service failure or on reducing dissatisfaction with recovery? Drawing from the literatures on organizational learning, learning curves, and marketing, I hypothesize that dissatisfaction with recovery contributes more to the variation in customer outrage across firms than service failure does (H1), that a U-shaped function of operating experience explains more variation in dissatisfaction with recovery than in service failure (H2), and that heterogeneity in organizational learning curves explains more variation in dissatisfaction with recovery than in service failure (H3). The hypotheses are tested with quarterly data for nine major U.S. airlines over 11 years. All three hypotheses are supported. In the context of mishandling baggage, dissatisfaction with recovery explains 88% of the variation in customer outrage, whereas service failure explains only 12%. The empirical results suggest firms should pay more attention to organizational learning curves for service recovery.

Key words: Learning Curve; Organizational Learning; Marketing; Service Failure; Recovery; Customer Complaints; Customer Dissatisfaction; Airlines

To err is human; to recover, divine. (Hart et al. 1990, p.156)

1. Introduction

In 1990, Hart et al. published "The profitable art of service recovery." According to the authors, service failures are inevitable. Hence, firms must be prepared to recover from service failures, thereby turning angry, frustrated customers into loyal ones. After experiencing a service failure, customers will typically give firms a second chance for recovery. It is crucial for service firms to get it right the second time. Poor recovery leads to customer dissatisfaction, and possibly customer outrage (Schneider and Bowen 1999). Hart et al. (1990, p.149) pointed out that "No business can afford to lose customers, if only because it costs much more to replace a customer than it does to retain one – five times more, most industry experts agree." Customer (dis)satisfaction is an important driver of customer loyalty and repeat purchase (Heskett et al. 1997). Given the economics of customer loyalty, the importance of recovering and learning from service failure seems self-evident. Yet eight years after the publication of "The profitable art of service recovery", Tax and Brown (1998) conclude that the majority of customers are still dissatisfied with the way companies resolve their complaints. The question arises whether firms should focus their learning efforts on service failure or on recovery.

Service marketing scholars have explored the "service recovery paradox": are customers who experience a service failure followed by excellent recovery at least as satisfied as or more satisfied than customers who do not experience failure? If a service recovery paradox exists, firms should focus learning efforts on recovery. Smith and Bolton (1998) surveyed students and business travelers to obtain self-reports of service failure and recovery episodes. Smith and Bolton (1998, p.77) found that "excellent service recovery can lead to increased customer satisfaction and repatronage intentions. However, the result was only obtained at the very high levels of customers' recovery ratings." The authors suggested that organizations might be better off reducing service failures rather than using recovery as an opportunity to increase satisfaction. McCollough et al. (2000) conducted scenario-based experiments with passengers from 50 departures in one airport. The authors found no evidence for a recovery paradox. Even excellent recovery did not yield higher satisfaction levels compared to error-free service delivery.

While these assessments of a service recovery paradox suggest firms should focus their learning efforts on reducing service failure, the literature lacks evidence regarding the area in which organizational learning efforts might provide more opportunity to distinguish a firm from competitors – service failure or recovery? This paper is the first empirical comparison of organizational learning curves for service failure and recovery.

A previous paper (Lapré and Tsikriktsis 2006) studied organizational learning curves for customer dissatisfaction measured by consumer complaints filed with the government divided by passengers. The findings indicated that (1) customer dissatisfaction with airlines followed a U-shaped function of operating experience, and (2) organizational learning curves across airlines were heterogeneous. In the current paper, I tease out where the U-shaped learning-curve effect and learning-curve heterogeneity originate: service failure or recovery? To do so, I decompose customer outrage into service failure and dissatisfaction with recovery. Service failure is measured by mishandled bags divided by passengers; dissatisfaction with recovery by baggage-related consumer complaints filed with the government divided by mishandled bags. First, I hypothesize that recovery (as opposed to service failure) provides more opportunity to distinguish a firm from competitors. Hence, I expect more variation between airlines in dissatisfaction with recovery than in service failure. Next, I hypothesize that the U-shaped learning-curve effect explains more variation in dissatisfaction with recovery than in service failure, and that heterogeneity in organizational learning curves explains more variation in dissatisfaction with recovery than in service failure. I test the hypotheses with quarterly data for nine major U.S. airlines over 11 years. Smith and Bolton (1998, p.70) note:

> Service recovery research is particularly challenging because the activities associated with recovery are triggered by a service failure, making systematic empirical studies (in a lab or field environment) very difficult to conduct. Thus prior studies on service failure/recovery

encounters have been limited to descriptive research based primarily on retrospective self-reports of service failure and recovery episodes.

Given the focus on self-reports in past service failure/recovery research, the airline dataset provides a unique opportunity to systematically assess service failure and recovery performance at the firm level over time. This paper is organized as follows. In Section 2, I develop my hypotheses. Section 3 describes the data and methodology. Section 4 presents the empirical results, which are further discussed in Section 5. Section 6 concludes with contributions, limitations, and questions for future research.

2. Focus Learning Efforts on Service Failure or Recovery?

2.1. Service Failure and Recovery Statistics in the Airline Industry

The U.S. Department of Transportation (DOT) defines mishandled baggage as an instance where an airline receives a report from a passenger concerning lost, damaged, delayed or pilfered baggage. In many industries, the most common response for a dissatisfied customer is to do nothing (Oliver 1997). Airline passengers, however, tend to care enough about their personal belongings to file mishandled baggage reports with airlines. DOT requires airlines to report mishandled baggage data, which DOT publishes every month in Air Travel Consumer Reports. From 1988 to 1998, the 9 largest U.S. airlines combined received over 25.7 million mishandled baggage reports, on average 5.8 mishandled baggage reports per 1,000 passengers. After airlines' recovery attempts, some customers are so dissatisfied (outraged) that they file a baggage complaint with DOT. From 1988 to 1998, DOT received 8,638 baggage complaints against the 9 largest U.S. airlines, on average 0.34 baggage complaints per 1,000 mishandled baggage reports. While this number might seem low, Lapré and Scudder (2004) showed the importance of reducing customer complaints filed with DOT. All major U.S. airlines that achieved sustainable cost reductions achieved lasting complaint reductions first. Therefore, reducing complaints is necessary (though not sufficient) to achieve lasting cost reduction – which certainly in the airline industry is a must. For most of the years between 1988 and 1998, baggage complaint was the second largest complaint category behind flight problems. However, the category "flight problems" lumps three types of problems together (delays, cancellations, and misconnections). If these three problems all had their own category, baggage complaints would be the number 1 complaint category. Either way, there is a need for airlines to reduce customer outrage measured by baggage complaints filed by passengers with DOT. Next, I develop my hypotheses. The hypotheses development draws from the literatures on learning curves, organizational learning, and marketing.

2.2. Hypotheses

Define the overall rate of customer outrage as number of complaints filed with a third party divided by the number of customers served. This rate is the product of (the number of service failures over the number of customers served) and (the number of complaints filed with a third party over the number of service failures) – rates of service failure and dissatisfaction with recovery respectively. Which rate (service failure or dissatisfaction with recovery) provides more opportunity for competitive advantage? Crandall, former CEO of American Airlines noted that "The physical aspects of our company are very much like our competitors'. After all, we all use the same airports, the same airplanes; we buy food from the same caterers, we use the same computers, and so on. It's pretty clear that the only difference between "us" and "them" is our ability … the ability of all of American's people to provide superior service." (Lorsch et al. 1991, p.10). Airlines use common systems and resources to handle baggage such as automated generation of baggage tags, conveyer belts, and baggage handlers. Not only are these resources common across airlines, airlines often share these resources. Given the similarity of resources, and by extension, causes for service failure, I expect little variation in the rates of service failure across airlines.

After reporting service failures with service providers, customers pay close attention to firms' recovery attempts. In particular, customers expect to be treated fairly on three dimensions: outcome, process, and interaction (Tax et al. 1998). Tax and Brown (1998) found that these three fairness dimensions explain 85% of the variation in satisfaction with recovery efforts. "Further, performing poorly on even one fairness dimension severely limits the potential for customer satisfaction." (Tax and Brown 1998, p.81). Effective frontline employees play a key role in providing fair outcomes, processes, and interactions during service recovery. So, hiring, training, and empowerment are critical in preparing employees for their service recovery role. Going back to Crandall's quote above this is where airlines can distinguish themselves from each other: "the only difference between "us" and "them" is … the ability of all of American's people to provide superior service." Gittell (2003) demonstrated that large variation exists across airlines in terms of investments in frontline leaders and hiring and training practices for relational competence. Hence, I expect more variation in dissatisfaction with recovery than in service failure.

HYPOTHESIS 1 (H1): *Dissatisfaction with recovery explains more of the variation in customer outrage across airlines than service failure does.*

Recent studies have found a U-shaped learning-curve effect for organizational survival rates (Ingram and Baum 1997, Baum and Ingram 1998), profitability (Ingram and Simons 2002), and customer dissatisfaction (Lapré and Tsikriktsis 2006): initially, organizational performance improves as a function of operating experience, but eventually deteriorates. The rationale offered for a U-shaped effect is rooted in the trade-off between exploration and exploitation (March 1991). As organizations focus on exploitation, repeated practice with an existing set of routines makes perfect. Organizations learn by doing the same things over and over again. Focus on exploitation at the expense of exploration can constrain organizations to "competency traps" (Levitt and March 1988). The set of routines perfected by the organization is no longer adequate in a changing environment.

It is important to note that learning by doing builds both internal and external organizational capabilities (Ingram and Baum 1997). By accumulating operating experience, organizations should become more efficient–the internal capability. When efficiency improves as a function of operating experience, learning by doing is inferred (Argote 1999). At the same time, operating experience also contributes to the external capability of an organization. Specifically, organizations learn about consumers' preferences–what do consumers want? Each time an airline operates a flight, the airline gains experience in handling baggage. By repeatedly operating flights, an airline also learns how to reduce customer dissatisfaction resulting from mishandled baggage. The empirical evidence of U-shaped learning curves concerns relative measures of organizational performance reflecting external capabilities: organizational survival, profitability, and customer dissatisfaction all capture performance evaluated by the market place. Absolute measures of organizational performance reflecting internal capabilities such as efficiency are expected to continue to improve (e.g. Ingram and Baum 1997).

Customer dissatisfaction occurs when customers' ex ante expectations about a product or service exceed ex post perceptions about the product or service. Customer dissatisfaction is a relative measure of performance, because customers can "raise the bar" – increase expectations based on past experience either with the same organization (Grayson and Ambler 1999) or with other organizations (Zeithaml et al. 1990). Lapré and Tsikriktsis (2006) found a U-shaped relationship between customer dissatisfaction with airlines and airlines' operating experience. In this paper, I decompose customer outrage into service failure and dissatisfaction with recovery. Reducing mishandled baggage (service failure) is an internal capability – fostered by better knowledge about internal operations, whereas dissatisfaction with recovery is an external capability – fostered by better knowledge about customer preferences. Hence, I expect a U-shaped effect for dissatisfaction with recovery. Furthermore, Maxham and Netemeyer (2002) studied bank customers who complained twice with the same bank within 20 months. The authors found that customers reporting two failures have higher recovery expectations for the second failure than for the first failure. Thus, I expect a U-shaped effect to explain more variation in dissatisfaction with recovery than in service failure.

HYPOTHESIS 2 (H2): *The U-shaped learning-curve effect explains more variation in dissatisfaction with recovery than in service failure.*

Prior learning-curve studies in service settings have typically assumed identical learning rates across firms. For example, Ingram and Baum (1997) and Baum and Ingram (1998) assumed identical learning rates across hotel chains. However, scholars have accumulated considerable evidence that learning rates show substantial variation across organizations (Argote 1999). The more organizations rely on unique knowledge resources, the more variation in learning rates can be expected (Lapré et al. 2000). A notable exception in learning-curve research in service settings is Pisano et al. (2001). The authors studied learning curves for surgery times after implementing a new technology for minimally invasive cardiac surgery at 16 different hospitals. The authors found that learning rates differed significantly across hospitals. Lapré and Tsikriktsis (2006) also found heterogeneity in organizational learning curves for customer dissatisfaction across airlines.

Edmondson et al. (2003) found that learning rates across organizations are more heterogeneous for dimensions of performance that rely on tacit knowledge than for those that rely on codified knowledge. Systems for baggage handling –such as automated printing of baggage tags, destination codes, and reading of these codes– rely more on codified knowledge. On the other hand, recovery attempts by frontline employees –including fair treatment– rely more on tacit knowledge. As mentioned above, hiring, training, and empowerment prepare employees for service recovery. Effective frontline employees should enhance improvement rates for the more tacit task of providing recovery that customers perceive as fair. Gittell (2003) examined hiring and training practices at Southwest, Continental, United and American. She found significant variation in airlines' hiring and training practices for relational competence – the ability to relate effectively with others. Southwest had the highest performance in terms of hiring for relational competence, American the lowest, and Continental and United in between. Variation in hiring for relational competence can be expected to contribute to variation in recovery performance, especially in frontline employees' ability to provide fair treatment. Thus, I expect heterogeneity in organizational learning curves to explain more variation in dissatisfaction with recovery than in service failure.

HYPOTHESIS 3 (H3): *Heterogeneity in organizational learning curves explains more variation in dissatisfaction with recovery than in service failure.*

3. Data and Method

The U.S. Department of Transportation (DOT) classifies an airline as major if the airline has at least one percent of total U.S. domestic passenger revenues. The dataset includes the 9 major airlines for the entire period 1988-1998: American, America West, Continental, Delta, Northwest, Southwest, TWA, United, and US Airways. Two other major airlines operating in part of 1988-1998 ceased operations well before 1998: Eastern in 1990 and Pan Am in 1991. One airline, Alaska, became a major airline in 1995. Because data on Alaska's baggage complaints is not consistently available, Alaska is excluded from the analysis. After Eastern's cessation in 1990, the nine major airlines included in this study account for more than 90% of revenue passenger miles for all U.S. airlines. (One revenue passenger mile is transporting one passenger over one mile in revenue service.) This section describes the data and research methodology, drawing partly from Lapré and Scudder (2004) and Lapré and Tsikriktsis (2006).

3.1. Service Failure: Mishandled Baggage

The service failure I focus on is an airline's failure to properly handle passengers' baggage. In October 1987, DOT required the largest U.S. airlines to file the rate of mishandled baggage reports per 1,000 passengers each month for publication in DOT's Air Travel Consumer Reports. The mishandled baggage rate is based on the total number of reports each airline received from passengers concerning lost, damaged, delayed or pilfered baggage. Initially, airlines used different systems to track baggage problems leading to variations in what was reported to DOT. Starting in December 1987, DOT made certain that rates could be compared across airlines (DOT Air Travel Consumer Reports):

> Some carriers have more comprehensive reporting systems than others. In order to establish a uniform system from which useful comparisons can be made, reports filed with carriers about courtesy tracers, voluntary separations, carry-on baggage and double counts are not included in calculating the rate of complaint reports. This allows the data for each carrier to be listed in rank order, based on the net number of reports filed per 1,000 passengers.

In July 1992, all reporting airlines had comprehensive systems in place, so DOT switched from publishing net number of baggage reports to total baggage reports. As rates based on total baggage reports are slightly higher than rates based on net number of baggage reports, I control for this reporting change in the analysis.

Let Mishandled Baggage Reports$_{it}$ denote the number of mishandled baggage reports received by airline i in quarter t, and let Passengers$_{it}$ be the number of enplaned passengers for airline i in quarter t. I measure service failures with Mishandled Baggage Rate$_{it}$ = Mishandled Baggage Reports$_{it}$ / (Passengers$_{it}$ /

1,000), the number of mishandled baggage reports per 1,000 passengers. Figure 1 shows the mishandled baggage rates for the 9 major airlines.

Figure 1 about here.

3.2. Recovery: Consumer Complaints

After receiving mishandled baggage reports from passengers, airlines get their chance to recover passengers. I will use baggage complaints filed by consumers with DOT divided by mishandled baggage reports to measure customer dissatisfaction with airlines' recovery efforts. From 1987 to 1998, passengers could file baggage complaints with DOT in writing, by telephone, or in person. Several factors led to a surge of baggage complaints against airlines in 1987: in August 1987 baggage complaints were up by 450% over January 1987.

First, in early 1987, airlines' performance as well as DOT's consumer phone number and address were given widespread publicity, which in turn led to increased consumer awareness concerning airline quality and the means to file complaints.

Second, in May 1987, the Secretary of Transportation, Elizabeth Dole, sent a letter to all major airlines concerning consumer dissatisfaction with the airline industry. She asked airlines to consider several steps including reeducation and training of employees, assessment of resources allocated to various sources for dissatisfaction, such as processing refunds and baggage claims, and review of complaint trends and processing times to resolve complaints.

Third, as mentioned above, in October 1987 DOT required major airlines to report mishandled baggage statistics. DOT, subsequently, started to publish these statistics along with the number of baggage complaints in DOT Air Travel Consumer Reports. In 1999, DOT introduced an e-mail address as an additional channel for filing consumer complaints.

I use baggage complaints filed with DOT from 1988 to 1998 for the following reasons.

- Consumer complaints filed with DOT clearly capture customer dissatisfaction: these customers were so dissatisfied that they wanted to tell the government about their service interactions.

- During the period 1988-1998, consumers could only file complaints with DOT in writing, by telephone, or in person. No other channels were added or deleted during this time frame. Restricting the dataset to 1988-1998 controls for ease of reporting complaints. Starting in 1988 as opposed to 1987 also controls for increased awareness of filing complaints with DOT.

When DOT receives a consumer complaint in month *m*, the complaint does not necessarily concern an incident that occurred in month *m*. Typically, 20% of all consumer complaints received in month *m* concern month *m*, about 35% concern month *m*–1, 40 to 45% all prior months, and for a small percentage the occurrence of the incident is unknown. To take the delay between the occurrence of mishandled baggage and receipt of baggage complaints by DOT into account, I match complaints received in February, March, and April with mishandled baggage reports for January, February, and March. I use a similar matching for the other quarters. Let Baggage Complaints_{it} denote the number of baggage complaints against airline *i* matched to quarter *t* (i.e., if *t* is the first quarter in the year, these are the baggage complaints received in February, March, and April). To measure customer dissatisfaction with recovery, I use Baggage Complaint Rate_{it} = Baggage Complaints_{it} / (Mishandled Baggage Reports_{it} / 1,000), the number of baggage complaints per 1,000 mishandled baggage reports. Figure 2 shows the baggage complaint rates for the 9 major airlines.

Figure 2 about here.

3.3. Independent Variables

Connections provide more opportunities for mishandling baggage. I control for connections with the variable Connections_{it}, the percentage of passengers connecting to another flight for airline *i* in quarter *t*. This variable is constructed using data on travel statistics reported to DOT. To control for the reporting change by DOT in July 1992, I introduce the dummy variable Total_t = 1 for quarters after June 1992, 0 for quarters prior to July 1992. The impact of factors such as weather and holidays on service failures can exhibit seasonal trends. I control for any seasonality with dummy variables. Let Quarter2_t be 1, if *t* is the second quarter of a calendar year, 0 otherwise. Similarly, I define Quarter3_t and Quarter4_t. Following

Lapré and Tsikriktsis (2006), I use cumulative number of flights since January 1988 to capture operating experience (Experience_{it}). Later in the paper, I address the robustness of the findings if I use different experience variables.

3.4. Learning-Curve Model

The functional form most commonly used in learning-curve research is the power form. The power form specifies that the logarithm of unit cost decreases as a linear function of the logarithm of cumulative production volume. Despite its frequent use, Lapré et al. (2000) note fundamental shortcomings of the power form. Arguably the biggest shortcoming is that the power form is an empirical observation lacking any theoretical underpinnings.

An alternative functional form for the learning curve is the exponential form, first introduced by Levy (1965) as the "adaptation curve." The adaptation curve follows from the assumption that the rate of improvement of a process is proportional to the amount a process can improve. Lapré et al. (2000) provide a theoretical foundation for Levy's assumption grounded in the organizational learning literature on performance gaps. In the context of complaint rates, let E be a measure for operating experience, $BCR(E)$ the baggage complaint rate after the organization has accumulated E experience, μ the learning rate, and BCR^* the optimal target level for the baggage complaint rate. According to Lapré et al. (2000, p.600):

> "The performance gap [$BCR(E) - BCR^*$] induces an organization to search for alternatives to reduce this gap. A larger discrepancy spurs the organization to exert more effort in searching for better knowledge (e.g., March and Simon 1958). The effectiveness of acquiring new knowledge is determined by the learning rate μ. Consequently, we can model the rate of improvement as the product of the learning rate and the performance gap."

$$\mathrm{d}BCR(E)/\mathrm{d}E = \mu\,[BCR(E) - BCR^*]$$

A natural target level for BCR^* is 0, because the performance measure concerns complaints filed with a third party–the government. The solution for this differential equation is the exponential form:

$$BCR(E) = \exp(a + \mu\, E),$$

or for estimation purposes:

$$\ln(BCR(E)) = a + \mu E.$$

The same logic applies to service failures measured by mishandled baggage rate. In addition to the theoretical interpretation of the exponential form rooted in the organizational learning literature, there is another benefit for using the exponential form as opposed to the power form. Omitting prior experience with the exponential form does not bias learning rates (Lapré and Tsikriktsis 2006). Hence, I will employ the exponential form in this paper.

3.5. Method

Total customer outrage with an airline's baggage performance can be measured with Baggage Customer Outrage$_{it}$ = Baggage Complaints$_{it}$ / (Passengers$_{it}$ / 1,000,000). Total baggage related customer outrage can be decomposed into service failure and recovery components as follows:

$$\ln(\text{Baggage Complaints}_{it} / (\text{Passengers}_{it} / 1{,}000{,}000) = \ln\{[\text{Baggage Complaints}_{it} / (\text{Mishandled Baggage Reports}_{it} / 1{,}000)] \times [\text{Mishandled Baggage Reports}_{it} / (\text{Passengers}_{it} / 1{,}000)]\}$$

$$= \ln(\text{Baggage Complaint Rate}_{it}) + \ln(\text{Mishandled Baggage Rate}_{it}).$$

Consequently, estimating this identity as

$$\ln(\text{Baggage Customer Outrage}_{it}) = \alpha_0 + \alpha_1 \ln(\text{Mishandled Baggage Rate}_{it}) + \alpha_2 \ln(\text{Baggage Complaint Rate}_{it}) + u_{it} \quad (1)$$

should yield parameter estimates of 0 for α_0, 1 for α_1 and α_2, and an R^2 of 1. To test H1, I estimate:

$$\ln(\text{Baggage Customer Outrage}_{it}) = \alpha_0 + \alpha_1 \ln(\text{Mishandled Baggage Rate}_{it}) + u_{it} \quad (1a)$$

and

$$\ln(\text{Baggage Customer Outrage}_{it}) = \alpha_0 + \alpha_2 \ln(\text{Baggage Complaint Rate}_{it}) + u_{it}. \quad (1b)$$

The R^2 of Model (1a) represents the contribution of service failure to the variation in baggage related customer outrage across airlines, while the R^2 of Model (1b) represents the contribution of dissatisfaction with recovery to the variation in baggage related customer outrage across airlines. If the R^2 of Model (1a) is smaller than the R^2 of Model (1b), H1 is supported (dissatisfaction with recovery explains more variation in customer outrage than service failure does).

The base model for service failure is the exponential form without any experience variables. The base model includes control variables for DOT's change in reporting, connections that provide more opportunities for service failure, seasonal dummy variables, and dummy variables for individual airlines (D_{ij} = 1 if $i = j$, 0 otherwise). As the intercept is a linear combination of all nine airline dummy variables, one airline dummy variable needs to be excluded. Without loss of generality, I do not include a dummy variable for Southwest:

$$\ln(\text{Mishandled Baggage Rate}_{it}) = \beta_0 + \beta_1 \text{ Total}_t + \beta_2 \text{ Connections}_{it} + \beta_3 \text{ Quarter2}_t$$
$$+ \beta_4 \text{ Quarter3}_t + \beta_5 \text{ Quarter4}_t + \sum_{j>1} \beta_{6j} D_{ij} + u_{it}. \tag{2}$$

In Model (2), β_{6j} captures how much the average logarithm of the mishandled baggage rate for airline j deviates from the average logarithm of Southwest's mishandled baggage rate β_0, holding other variables constant. To test whether airlines learn to reduce service failures, I include Experience$_{it-1}$:

$$\ln(\text{Mishandled Baggage Rate}_{it}) = \beta_0 + \beta_1 \text{ Total}_t + \beta_2 \text{ Connections}_{it} + \beta_3 \text{ Quarter2}_t$$
$$+ \beta_4 \text{ Quarter3}_t + \beta_5 \text{ Quarter4}_t + \sum_{j>1} \beta_{6j} D_{ij} + \beta_7 \text{ Experience}_{it-1} + u_{it}. \tag{3}$$

A negative estimate for β_7 would imply that airlines did learn to reduce service failures. Note that the interpretation of β_{6j} changes from Model (2) to Model (3). In Model (3), β_{6j} captures how much the intercept, or starting point, of airline j's learning curve deviates from Southwest's intercept. To test whether learning curves for service failures are U-shaped, I include the square of the experience variable (Experience$_{it-1}$)^2:

$$\ln(\text{Mishandled Baggage Rate}_{it}) = \beta_0 + \beta_1 \text{ Total}_t + \beta_2 \text{ Connections}_{it} + \beta_3 \text{ Quarter2}_t$$
$$+ \beta_4 \text{ Quarter3}_t + \beta_5 \text{ Quarter4}_t + \sum_{j>1} \beta_{6j} D_{ij} + \beta_7 \text{ Experience}_{it-1} + \beta_8 (\text{Experience}_{it-1})^2 + u_{it}. \tag{4}$$

A negative estimate for β_7 and a positive estimate for β_8 would indicate a U-shaped relationship between service failure and operating experience. I estimate models for baggage complaint rate similar to Models (2) – (4) to investigate a U-shaped relationship between dissatisfaction with recovery and operating experience. To test H2 (the U-shaped learning-curve effect explains more variation in dissatisfaction with recovery than in service failure), I compare the increase in R^2 from Model (3) to Model (4) for mishandled baggage rate with the increase in R^2 from similar models for baggage complaint rate. Models (3) and (4) assume identical learning rates across all airlines. In the full model, I allow for each airline to not only have a different intercept, but also a different U-shaped learning curve:

$$\begin{aligned}\ln(\text{Mishandled Baggage Rate}_{it}) = {} & \beta_0 + \beta_1 \text{ Total}_t + \beta_2 \text{ Connections}_{it} + \beta_3 \text{ Quarter2}_t \\ & + \beta_4 \text{ Quarter3}_t + \beta_5 \text{ Quarter4}_t + \sum_{j>1} D_{ij}\{\beta_{6j} + \beta_{7j} \text{ Experience}_{it-1} + \beta_{8j} (\text{Experience}_{it-1})^2\} \\ & + \beta_7 \text{ Experience}_{it-1} + \beta_8 (\text{Experience}_{it-1})^2 + u_{it}. \qquad (5)\end{aligned}$$

Significant estimates for β_{7j} and β_{8j} would indicate that organizational learning curves for service failure are heterogeneous across airlines. I also estimate a model for baggage complaint rate similar to Model (5) to investigate heterogeneity in organizational learning curves for dissatisfaction with recovery across airlines. Lastly, to test H3 (heterogeneity in organizational learning curves explains more variation in dissatisfaction with recovery than in service failure), I compare the increase in R^2 from Model (4) to Model (5) for mishandled baggage rate with the increase in R^2 from similar models for baggage complaint rate.

For all models, $i = 1, \ldots, N$; $t = 1, \ldots, T$. The dataset contains $N = 9$ airlines and $T = 44$ quarters. Therefore, the dataset is a case of "time-series cross-section" (TSCS) data as opposed to "panel" data (Beck 2001). In panel data, the units are sampled, and typically observed only a few times. Specific units are of no interest. The asymptotics are in N. Conversely, in TSCS data, units are fixed; there is no sampling, and we are interested in specific units ("how does American compare with Southwest?"). The asymptotics are in T. Consequently, there are three important considerations for the error term u_{it}: heteroscedasticity, contemporaneous correlation, and autocorrelation. To address these three considerations, Parks (1967) developed a generalized least squares method for TSCS data. However, Beck and Katz

(1995) showed that the Parks' method produces standard errors that lead to extreme overconfidence even for data sets with N = 10 and T = 40 – very similar to my dataset. Instead, to deal with heteroscedasticity and contemporaneous correlation, Beck and Katz (1995) developed a superior estimation procedure that uses ordinary least square (OLS) parameter estimates, but replaces the OLS standard errors with panel-corrected standard errors (PCSE). Lastly, if autocorrelation is a concern (typical in learning-curve research) one should use Prais-Winsten estimators with panel-corrected standard errors (Beck and Katz 1995). Unlike the Parks' method, Beck and Katz (1995) make a case against panel-specific autocorrelation, and recommend a single first-order autoregressive process for the error term: $u_{it} = \rho u_{it-1} + e_{it}$. Therefore, the estimation approach for all models is Prais-Winsten regression with panel-corrected standard errors and a single first-order autoregressive process common to all airlines (implemented with the procedure "xtpcse" in Stata).

4. Empirical Results

Table 1 shows the regression results for Models (1a), (1b), and (1). The R^2 for Model (1a) is 0.12; hence, service failure explains only 12% of the variation in baggage customer outrage. On the other hand, the R^2 of Model (1b) is 0.88; so, dissatisfaction with recovery explains 88% of the variation in baggage customer outrage. Note that in the last column estimation of the identity in Model (1) indeed yields estimates of 0 for α_0, 1 for α_1 and α_2, and an R^2 of 1. The estimates in Table 1 provide strong support for H1: the contribution of dissatisfaction with recovery to the variation in customer outrage (88%) is much larger than the contribution of service failure to the variation in customer outrage (12%). In fact, compared with service failure, dissatisfaction with recovery explains 0.88/0.12 = 7.3 times as much of the variation in customer outrage.

Table 1 about here.

Table 2 shows the regression results for Models (2) through (5) for service failure. The base model explains 72.9% of the variation. All eight airline dummy variables are positive and significant, indicating that the average mishandled baggage rate for these airlines was significantly higher than Southwest's average mishandled baggage rate. The estimates in Column (2) indicate that airlines reduced service failure as a function of operating experience. Yet, inclusion of a single experience variable common across all airlines (Experience) explains a modest 1.6% of the variation. The negative estimate for Experience and the positive estimate for $(\text{Experience})^2$ in Column (3) indicate that service failure follows a U-shaped relationship with experience. However, the increase in R^2 of 0.9% due to the inclusion of $(\text{Experience})^2$ is small. Estimates in Column (4) show an additional increase in R^2 of 5.8%. Four airlines have significantly different slope parameters for Experience and/or $(\text{Experience})^2$, indicating that organizational learning curves for service failure are heterogeneous across airlines. Combined though, the increase in R^2 from the base model to the full model is only 8.3%. Hence, both the U-shaped learning curve effect and heterogeneity in organizational learning only make a modest contribution in explaining service failure.

Table 2 about here.

Table 3 shows the regression results for Models (2) through (5) for dissatisfaction with recovery. The base model explains 15.5% of the variation. The estimate for Experience in Column (2) is not significant indicating that airlines did not reduce dissatisfaction with recovery. However, the negative estimate for Experience and the positive estimate for $(\text{Exprience})^2$ in Column (3) do provide strong evidence that dissatisfaction with recovery follows a U-shaped relationship with experience. The increase in R^2 of 15.7% due to the inclusion of $(\text{Experience})^2$ is very significant. Furthermore, the estimates in Column (4) show an additional significant increase in R^2 of 28.9%. Four airlines have significantly different slope parameters for Experience and/or $(\text{Experience})^2$, indicating that organizational learning curves for dissatisfaction with recovery are heterogeneous across airlines. The increase in R^2 from the base model to the full model is 45.5%. Consequently, both the U-shaped learning-curve effect and heterogeneity in organizational learning make a substantial contribution in explaining dissatisfaction with recovery.

Table 3 about here.

The U-shaped effect explained only 0.9% of the variation in mishandled baggage rate (increase in R^2 from Column (2) to Column (3) in Table 2). On the other hand, the U-shaped effect explained 15.7% of the variation in baggage complaint rate (increase in R^2 from Column (2) to Column (3) in Table 3). Thus, H2 is supported: the U-shaped learning-curve effect explains more variation in dissatisfaction with recovery than in service failure. Similarly, H3 is supported – heterogeneity in organizational learning curves explains more variation in dissatisfaction with recovery than in service failure. The increase in R^2 from Column (3) to Column (4) in Table 3 is 28.9% for baggage complaint rate compared to only 5.8% in Table 2 for mishandled baggage rate. Table 4 summarizes the statistics used to test the three hypotheses. Table 4 also gives a clear answer to the research question "Where do the U-shaped learning-curve effect and learning-curve heterogeneity for customer outrage originate: service failure or dissatisfaction with recovery?" The U-shaped learning-curve effect for service failure explains 0.9% × 12% = 0.1% of the variation in customer outrage, whereas the U-shaped learning-curve effect for dissatisfaction with recovery explains 15.7% × 88% = 13.8%. Similarly, learning-curve heterogeneity for service failure explains 5.8% × 12% = 0.7% of the variation in customer outrage, while learning-curve heterogeneity for dissatisfaction with recovery explains 28.9% × 88% = 25.4%.

Table 4 about here.

To assess the robustness of the estimates in Tables 2 and 3 obtained with cumulative number of flights as the measure for experience, I re-estimated all models with two alternative measures for experience: cumulative number of mishandled bags and cumulative number of passengers served. The R^2 values obtained with either measure were similar to or slightly lower than the values reported in Tables 2 and 3. More importantly, all results and conclusions remain the same.

5. Discussion

The results provide strong support for the notion that dissatisfaction with recovery contributes much more to the variation in total customer dissatisfaction than service failure – 88% compared to 12%. To visualize this finding, Figure 3 depicts selected estimated U-shaped learning curves calculated with the estimates obtained in columns (4) of Tables 2 and 3. In order to compare different learning-curve patterns without making Figure 3 too crowded, I only retained American, Continental, Delta, Southwest, and TWA for the following reasons:

- Southwest and TWA represent the best and the worst performance for service failure. Southwest is the benchmark for the time horizon studied, while TWA (i) has the worst starting point, and (ii) is the only airline with a significant U-shaped effect for service failure (Column (4) in Table 2).
- Delta and Continental represent the best and worst starting points for recovery.
- Southwest and American represent the best and worst ending points for recovery.

The two graphs with estimated learning curves have the same spread on the vertical axis (4.0) so that the variation can be compared across the two graphs. Moreover, we know from Equation (1), confirmed by estimates in Table 1, that we can basically add up the two. Clearly, the estimated service failure learning curves show much less variation than the estimated recovery learning curves – the service failure learning curves are flatter, exhibit less learning, less relapse, and are bunched together. In contrast, the estimated recovery learning curves show much more variation across airlines in starting points, learning rates, relapse rates, and ending points.

Figure 3 about here.

Because the variation in relapse rates is much larger for recovery, it is much harder to keep the gains in recovery. The extreme example of this phenomenon is American. American lost all improvements it had gained in recovery to end at the same level in 1998 compared to where the airline started in 1988. Southwest, on the other hand, did not start out as the best airline in terms of customer outrage, but became the best due to its improvements in recovery. Recovery provided Southwest more opportunity to distinguish itself from the competition than service failure did. Indeed, by 1998, many airlines had caught up to Southwest's mishandled baggage rate performance, while Southwest had moved ahead of the competition with its baggage complaint rate.

Why was the observed variation in dissatisfaction with recovery 7.3 times as large as the observed variation in service failure? Why do most service firms have so much trouble learning to improve recovery (Tax and Brown 1998)? Apparently it must be hard for most service firms to hire, train, and empower frontline employees who provide fair outcomes, processes, and interactions for customers. While Southwest instructs frontline employees to "do whatever you feel comfortable doing for a customer" (Heskett et al. 1997, p. 120), other airlines fail to make the right investments in frontline employees. Gittell (2003, p.37) observed that American's employees focused on assigning blame and avoiding blame as opposed to working together to solve problems. A ramp supervisor at American:

> "If you ask anyone here, what's the last thing you think of when there's a problem? I bet your bottom dollar it's the customer. And these are guys who bust their butts every day. But they're thinking, how do I keep my ass out of the sling?"

At Southwest, frontline employees get much more coaching from frontline leaders. On the other hand, at American, Continental, and United, supervisors are spread too thin to coach, causing them to focus on performance measures instead (Gittell 2003). Southwest has always been very careful in its growth strategy – only going into a new destination if they are assured that they can find the right Southwest like employees who relate well to Southwest's customers (Heskett and Hallowell 1993).

Customers typically give a service provider a second chance for recovery, but not delivering on recovery can create "terrorists" (Heskett et al. 1997), customers so outraged that they utilize every opportunity to criticize the service provider. A dissatisfied customer will tell from 10 to 20 other people about his or her problem (Fitzsimmons and Fitzsimmons 2004). With the internet, however, terrorists can reach thousands as the following example shows. Jeremy Cooperstock, a professor at McGill University, wrote a letter to United Airlines to complain about shoddy service on a trip to Tokyo in 1996: lack of coordination in rerouting passengers, not announcing a critical flight delay in English, not being able to satisfy seating reservations made weeks in advance, upgrading late arrival and standby passengers over others who had already been inconvenienced, and mishandling of a garment bag. Not having received a reply from United two months later, he sent a follow-up letter. United responded with a form letter that completely ignored the substance of the initial complaint, prompting Cooperstock to set up web pages at the University of Toronto, where he was a student at the time. Instead of addressing the complaint or the underlying causes, United threatened with legal action to have the web pages removed. In response, Cooperstock started the website untied.com (note the spelling). Many passengers sent Cooperstock complaints about United. He subsequently posted these complaints categorized by rudeness, misinformation, and incompetence, among others. Soon after, employees started to share their negative experiences as well, including complaints about not feeling empowered by management. Cooperstock commented: "… as it turns out, my experience was trivial compared to other horror stories. Having a ripped garment bag was not such a big deal." The site got a lot of media coverage from organizations such as Wall Street Journal, Forbes, Newsweek, USA Today, New York Times, and many others. By 2000, the site got between 15,000 and 33,000 visits a month, half of them by unique users, 1% of the visits from United Airlines headquarters. It is important to note that "It wasn't the poor treatment [Cooperstock] received from United, but rather the subsequent disregard for a serious, polite complaint that led to the creation of the Web pages." (untied.com) In other words, it wasn't the service failure that led Cooperstock to start untied.com, but rather his dissatisfaction with United's recovery attempts (or lack thereof).

6. Conclusion

Argote (1999, pp.201-202) called for organizational learning-curve scholars to examine "a wider set of outcome measures and also different organizational contexts, including more service settings." This paper contributes to the literature on U-shaped learning curves by examining service failure and dissatisfaction with recovery. Specifically, this paper adds a data point to the studies of U-shaped learning curves ranging from internal to external capabilities. Darr et al. (1995) did not observe a U-shaped effect for unit costs in pizza stores (internal efficiency). This paper identified a very modest U-shaped effect for mishandling baggage (internal service failure), only observed for one out of nine airlines (TWA); and a significant U-shaped effect for customer dissatisfaction with recovery (short-term external evaluation of service firms). Ingram and Baum (1997) and Baum and Ingram (1998) found strong evidence for a U-shaped effect for organizational survival of hotel chains (long-term external evaluation of service firms). It seems that the U-shaped learning curve effect becomes more pronounced for longer term external evaluation of organizations.

Secondly, the paper also contributes to the literature on heterogeneity in learning curves. Heterogeneity in learning curves is more important for dissatisfaction with recovery than for service failure. As front-line employees play a key role in providing service recovery, service firms are well advised to help front-line employees learn about improving recovery efforts. Tucker et al. (2002) observed that front-line employees who face customers with different reactions to situations and unique requests often end up solving problems on their own. Combined with pressures to reduce labor costs, long-term learning from service failures and recovery suffers. However, the empirical results regarding the relapse rates for dissatisfaction with recovery underscore the importance of learning. American lost all its improvements in recovery, whereas Southwest substantially outperformed the competition. Interestingly, in a study of four airlines, Gittell (2003) found that Southwest scored highest on hiring for relational competence, American scored lowest, while Continental and United were in between. The recovery learning-curve estimates in Figure 3 show the same ranking as Gittell's (2003) hiring for relational competence data (collected in the mid 1990s).

Thirdly, the paper contributes to the literature on service failure and recovery. Service recovery research has typically been limited to studying self-reports. The airline dataset revealed that -compared to service failure- dissatisfaction with recovery can contribute 7.3 times as much to the variation in customer outrage. This finding sheds new light on the question whether firms should focus their learning efforts on service failure or recovery in order to gain a competitive advantage.

This paper has investigated a single type of service failure, in a single industry, in a single country, using a single measure of customer dissatisfaction (complaints to a third party – the government). Future research is needed to assess the generalizability of this study on any of these dimensions. It should be particularly fruitful to study service failures that occur in the customers' presence (the front-office as opposed to the back-office). Another limitation of the research concerns the simple tally of mishandled baggage reports. It should be interesting to incorporate the severity of service failures in a follow-up study. According to Zeithaml et al. (1990), customer dissatisfaction is a function of the difference between ex ante expectations about a service and ex post perceptions of service delivery. Future research is also needed to disentangle ex ante expectations and ex post perceptions. In learning about service failure and recovery, do firms learn to manage expectations, perceptions, the balance? Which avenue is most conducive to managing service failure and recovery learning curves? Hopefully, research along these lines will advance our knowledge of how service firms can better manage organizational learning curves.

Acknowledgments

Participants at Wharton's Technology Miniconference, in particular discussant George Day, participants at the POMS 2005 conference in Chicago, as well as Piyush Kumar, Rich Oliver, Gary Scudder, and Tim Vogus provided helpful comments. Support for this research was provided by the Dean's Fund for Faculty Research, Owen Graduate School of Management at Vanderbilt University.

References

Argote, L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic Publishers, Norwell, MA.

Baum, J.A.C., P. Ingram. 1998. Survival-enhancing learning in the Manhattan hotel industry, 1898-1980. *Management Science* **44**(7) 996-1016.

Beck, N. 2001. Time-series cross-section data: What have we learned in the past few years? *Annual Review of Political Science* **4** 271-293.

Beck, N., J.N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review* **89**(3) 634-647.

Darr, E.D., L. Argote, D. Epple. 1995. The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science* **41**(11) 1750-1762.

Edmondson, A.C., A.B. Winslow, R.M.J. Bohmer, G.P. Pisano. 2003. Learning how and learning what: Effects of tacit and codified knowledge on performance improvement following technology adoption. *Decision Sciences* **34**(2) 197-223.

Fitzsimmons, J.A., M.J. Fitzsimmons. 2004. *Service Management: Operations, Strategy, and Information Technology*, 4th edition. McGraw-Hill, New York, NY.

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

Grayson, K., T. Ambler. 1999. The dark side of long-term relationships in marketing services. *Journal of Marketing Research* **36**(1) 132-141.

Hart, C.W.L., J.L. Heskett, W.E. Sasser, Jr. 1990. The profitable art of service recovery. *Harvard Business Review* **68**(4) 148-156.

Heskett, J.L., R.H. Hallowell. 1993. *Southwest Airlines: 1993 (A).* Harvard Business School Case 9-694-023. Harvard Business School Publishing, Boston, MA.

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

Ingram, P., J.A.C. Baum. 1997. Opportunity and constraint: Organizational learning from operating and competitive experience. *Strategic Management Journal* **18**(Summer Special Issue) 75-98.

Ingram, P., T. Simons. 2002. The transfer of experience in groups of organizations: Implications for performance and competition. *Management Science* **48**(12) 1517-1533.

Lapré, M.A., A.S. Mukherjee, L.N. Van Wassenhove. 2000. Behind the learning curve: Linking learning activities to waste reduction. *Management Science* **46**(5) 597-611.

Lapré, M.A., G.D. Scudder. 2004. Performance improvement paths in the U.S. airline industry: Linking trade-offs to asset frontiers. *Production and Operations Management* **13**(2) 123-134.

Lapré, M.A., N. Tsikriktsis. 2006. Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Science,* Forthcoming.

Levitt, B., J.G. March. 1988. Organizational learning. *Annual Review of Sociology* **14** 319-340.

Levy, F. 1965. Adaptation in the production process. *Management Science* **11**(6) B136-B154.

Lorsch, J., G. Loveman, J. Horn. 1991. American Airlines (A): Strategy in the 1990s. Case 9-491-044. Harvard Business School Publishing, Boston, MA.

March, J.G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**(1) 71-87.

March, J.G., H.A. Simon. 1958. *Organizations*. Wiley, New York, NY.

Maxham III, J.G., R.G. Netemeyer. 2002. A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing* **66**(4) 57-71.

McCollough, M.A., L.L. Berry, M.S. Yadav. 2000. An empirical investigation of customer satisfaction after service failure and recovery. *Journal of Service Research* **3**(2) 121-137.

Oliver, R.L. 1997. *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw-Hill, New York, NY.

Parks, R. 1967. Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association* **62** 500-509.

Pisano, G.P., R.M.J. Bohmer, A.C. Edmondson. 2001. Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science* **47**(6) 752-768.

Schneider, B., D.E. Bowen. 1999. Understanding customer delight and outrage. *Sloan Management Review* **41**(1) 35-45.

Smith, A.K., R.N. Bolton. 1998. An experimental investigation of customer reactions to service failure and recovery encounters. *Journal of Service Research* **1**(1) 65-81.

Tax, S.S., S.W. Brown. 1998. Recovering and learning from service failure. *Sloan Management Review* **40**(1) 75-88.

Tax, S.S., S.W. Brown, M. Chandrashekaran. 1998. Customer evaluations of service complaint experiences: Implications for relationship marketing. *Journal of Marketing* **62**(2) 60-76.

Tucker, A.L., A.C. Edmondson, S. Spear. 2002. When problem solving prevents organizational learning. *Journal of Organizational Change Management* **15**(2) 122-137.

Zeithaml, V.A., A. Parasuraman, L.L. Berry. 1990. *Delivering Quality Service: Balancing Customer Perceptions and Expectations.* The Free Press, New York, NY.

Table 1: Contribution to Variation in Baggage Customer Outrage

	Service Failure (1a)	Recovery (1b)	Both (1)
Constant	-0.952^{**}	1.719^{***}	0.000
	(0.319)	(0.074)	(0.000)
ln(Mishandled Baggage Reports / 1,000 Passengers)	0.849^{***}		1.000^{***}
	(0.167)		(0.000)
ln(Baggage Complaints / 1,000 Mishandled Baggage Reports)		0.983^{***}	1.000^{***}
		(0.021)	(0.000)
R^2	0.120	0.888	1.000
$N \times T$	396	396	396

Note. Dependent variable ln(Baggage Complaints / 1,000,000 Passengers). Panel-corrected standard errors in parentheses. * Significant at 0.05, ** at 0.01, and *** at 0.001.

Table 2: Service Failure Learning-Curve Estimates for Mishandled Baggage Rate

	(1)	(2)	(3)	(4)		
Constant	1.232***	1.387***	1.497***	1.339***		
	(0.093)	(0.085)	(0.097)	(0.088)		
Total	–0.055	0.088	0.168*	0.201**		
	(0.066)	(0.078)	(0.081)	(0.074)		
Connections	1.521***	1.049***	0.990***	0.756**		
	(0.326)	(0.288)	(0.284)	(0.288)		
Quarter2	–0.224***	–0.225***	–0.224***	–0.223***		
	(0.024)	(0.022)	(0.020)	(0.021)		
Quarter3	–0.173***	–0.170***	–0.171***	–0.167***		
	(0.029)	(0.026)	(0.024)	(0.025)		
Quarter4	–0.018	–0.023	–0.025	–0.026		
	(0.025)	(0.023)	(0.021)	(0.022)		
				Dummy	Experience × Dummy	(Experience)^2 × Dummy
American	0.335***	0.442***	0.454***	0.664***	–0.086	–0.001
	(0.063)	(0.069)	(0.072)	(0.101)	(0.076)	(0.012)
America West	0.261***	0.182*	0.134	0.783***	–0.708***	0.133
	(0.080)	(0.088)	(0.094)	(0.107)	(0.221)	(0.100)
Continental	0.319***	0.332***	0.350***	0.724***	–0.130	–0.009
	(0.071)	(0.072)	(0.075)	(0.100)	(0.096)	(0.019)
Delta	0.281***	0.452***	0.433***	0.620***	–0.020	–0.010
	(0.060)	(0.070)	(0.076)	(0.087)	(0.066)	(0.011)
Northwest	0.356***	0.405***	0.428***	0.690***	–0.177*	0.019
	(0.064)	(0.065)	(0.068)	(0.100)	(0.087)	(0.016)
TWA	0.359***	0.357***	0.341***	0.992***	–0.723***	0.170**
	(0.075)	(0.080)	(0.083)	(0.119)	(0.158)	(0.054)
United	0.444***	0.543***	0.566***	0.756***	–0.194*	0.022
	(0.069)	(0.071)	(0.072)	(0.109)	(0.080)	(0.013)
US Airways	0.353***	0.486***	0.481***	0.713***	–0.016	–0.012
	(0.077)	(0.082)	(0.088)	(0.113)	(0.081)	(0.012)
Experience		–0.064***	–0.173***	–0.099		
		(0.017)	(0.044)	(0.072)		
(Experience)^2			0.012**	0.014		
			(0.004)	(0.011)		
R^2	0.720	0.736	0.745	0.803		
Increase in R^2		0.016***	0.009***	0.058***		
ρ	0.673	0.678	0.688	0.528		
$N \times T$	396	396	396	396		

Note. Dependent variable ln(Mishandled Baggage Reports / 1,000 Passengers). Panel-corrected standard errors in parentheses. * Significant at 0.05, ** at 0.01, and *** at 0.001.

Table 3: Recovery Learning-Curve Estimates for Baggage Complaint Rate

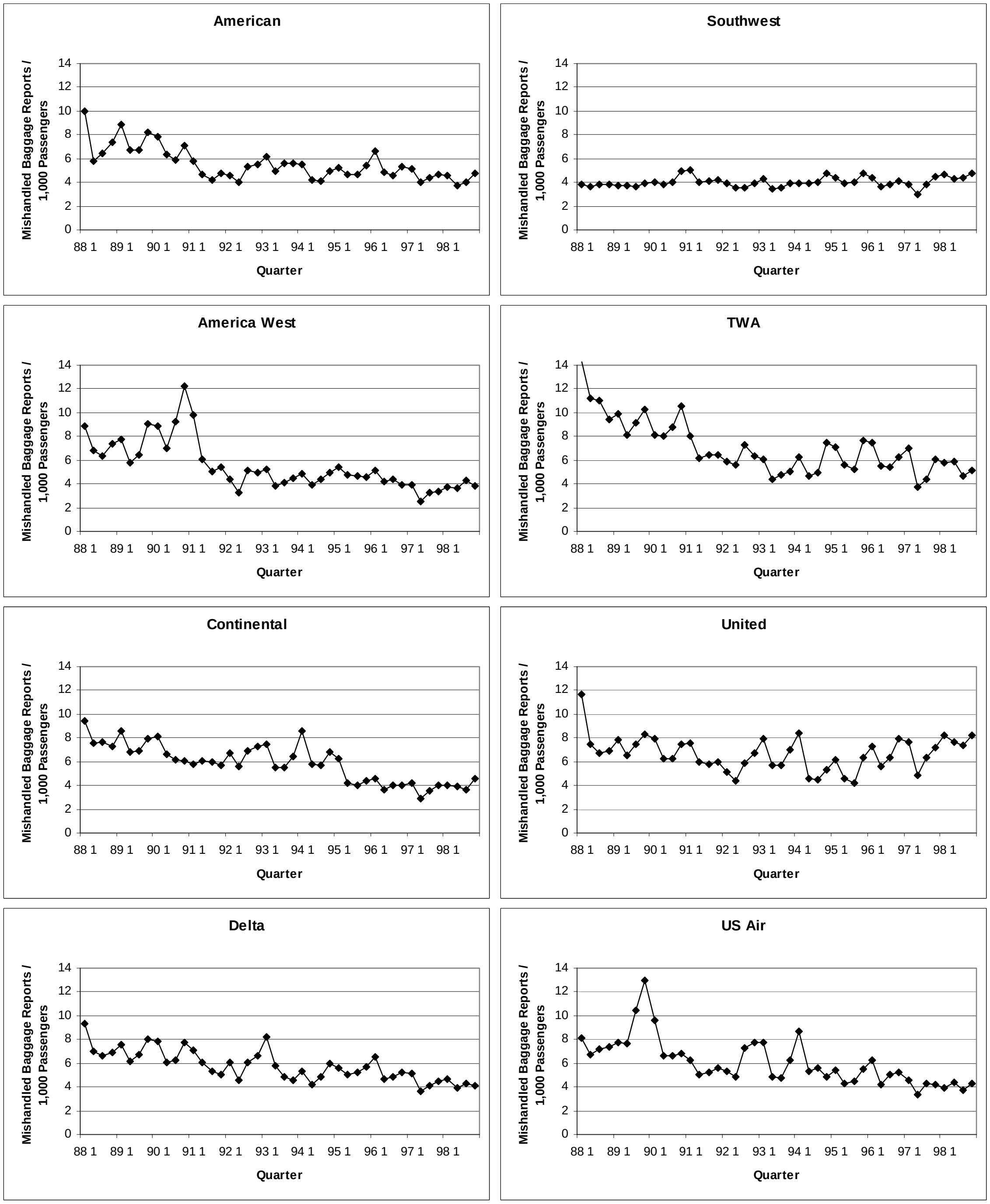
	(1)	(2)	(3)	(4)		
Constant	-1.471^{***}	-1.562^{***}	-1.021^{***}		-0.495	
	(0.362)	(0.368)	(0.321)		(0.372)	
Total	-0.508^{**}	-0.599^{**}	-0.110		0.135	
	(0.170)	(0.213)	(0.167)		(0.133)	
Connections	-0.389	-0.094	-0.276		0.032	
	(0.917)	(0.930)	(0.835)		(0.955)	
Quarter2	0.093	0.094	0.099		0.103	
	(0.071)	(0.073)	(0.057)		(0.058)	
Quarter3	0.057	0.054	0.047		0.038	
	(0.085)	(0.087)	(0.066)		(0.065)	
Quarter4	-0.044	-0.042	-0.048		-0.048	
	(0.074)	(0.075)	(0.058)		(0.059)	
				Dummy	Experience × Dummy	$(\text{Experience})^2$ × Dummy
American	0.779^{**}	0.717^{*}	0.823^{***}	-0.274	0.926^{**}	-0.133^{*}
	(0.293)	(0.300)	(0.241)	(0.353)	(0.306)	(0.054)
America West	1.018^{***}	1.064^{***}	0.794^{**}	0.241	-0.191	0.396
	(0.312)	(0.312)	(0.254)	(0.383)	(0.571)	(0.228)
Continental	0.983^{***}	0.974^{***}	1.066^{***}	1.472^{***}	-0.549	0.133
	(0.290)	(0.283)	(0.232)	(0.383)	(0.368)	(0.071)
Delta	-0.024	-0.126	-0.167	-1.085^{**}	0.979^{***}	-0.147^{**}
	(0.295)	(0.318)	(0.256)	(0.351)	(0.296)	(0.053)
Northwest	0.394	0.363	0.480	0.549	-0.369	0.099
	(0.327)	(0.321)	(0.264)	(0.394)	(0.361)	(0.065)
TWA	1.450^{***}	1.449^{***}	1.341^{***}	1.122^{*}	0.564	-0.220
	(0.303)	(0.296)	(0.249)	(0.439)	(0.458)	(0.125)
United	0.525	0.467	0.613^{**}	-0.141	0.736^{*}	-0.114^{*}
	(0.277)	(0.280)	(0.230)	(0.339)	(0.293)	(0.053)
US Airways	0.196	0.117	0.133	-0.334	0.721^{*}	-0.123^{*}
	(0.322)	(0.332)	(0.268)	(0.402)	(0.330)	(0.056)
Experience		0.038	-0.551^{***}		-1.290^{***}	
		(0.053)	(0.099)		(0.302)	
$(\text{Experience})^2$			0.065^{***}		0.181^{***}	
			(0.010)		(0.053)	
R^2	0.155	0.164	0.321		0.610	
Increase in R^2		0.009	0.157^{***}		0.289^{***}	
ρ	0.576	0.566	0.482		0.262	
$N \times T$	396	396	396		396	

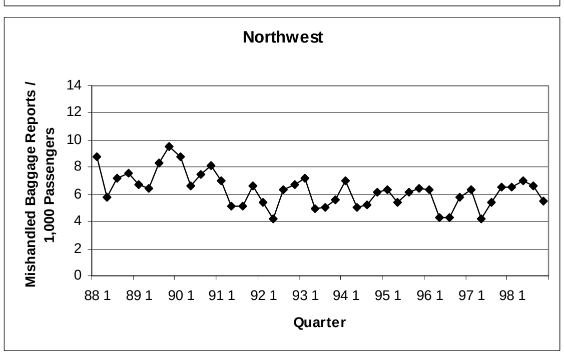
Note. Dependent variable ln(Baggage Complaints / 1,000 Mishandled Baggage Reports). Panel-corrected standard errors in parentheses. * Significant at 0.05, ** at 0.01, and *** at 0.001.

Table 4: Explained Variation in Customer Outrage

		Service Failure	Dissatisfaction with Recovery
(A)	R^2 from Table 1	0.12	0.88
(B)	R^2 explained by U-shaped effect (Column (3) in Tables 2 and 3)	0.009	0.157
(A) × (B)	Combined R^2: contribution of U-shaped learning curve in explaining Customer Outrage	0.001	0.138
(C)	R^2 explained by heterogeneity (Column (4) in Tables 2 and 3)	0.058	0.289
(A) × (C)	Combined R^2: contribution of learning-curve heterogeneity in explaining Customer Outrage	0.007	0.254

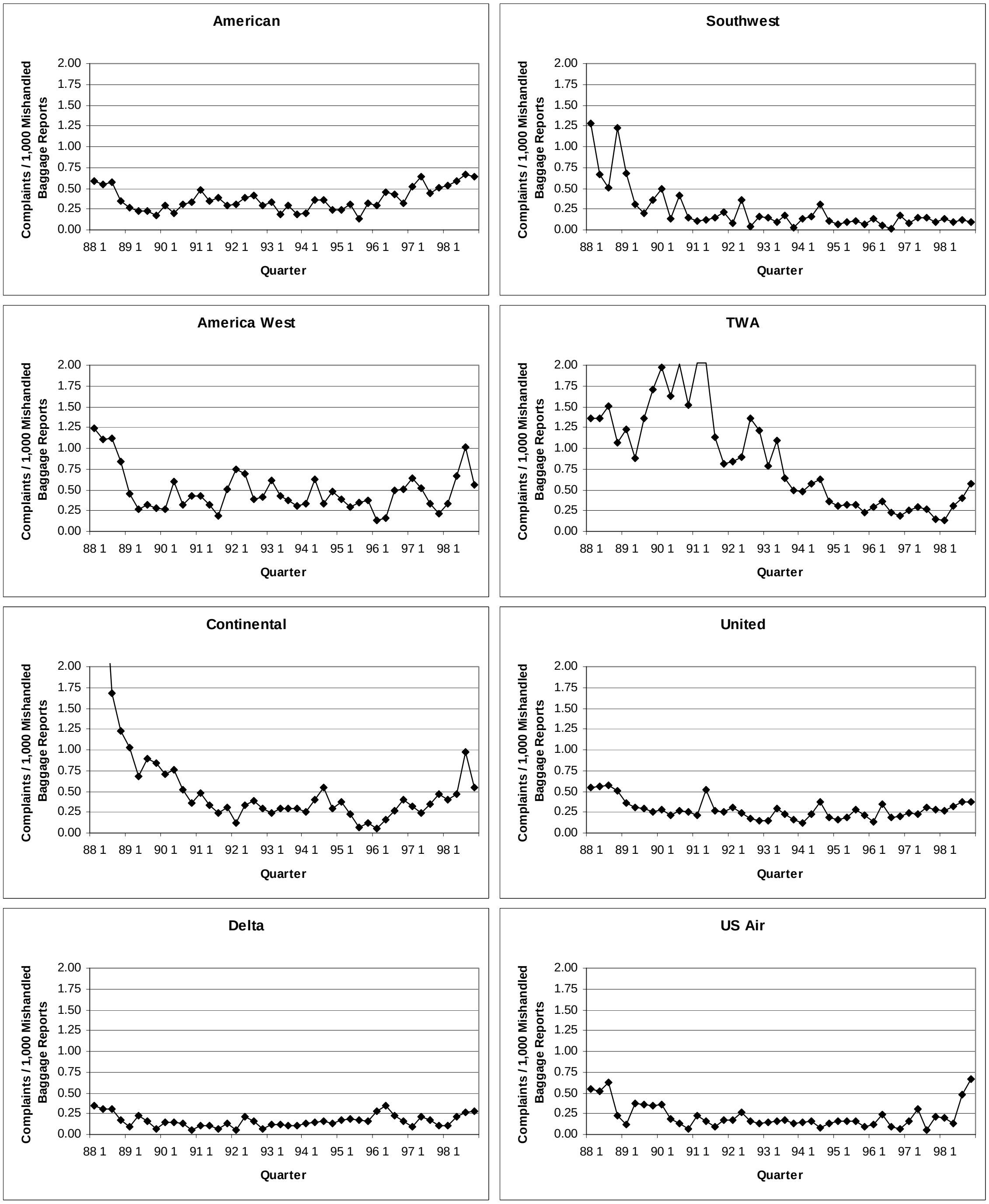
Figure 1: Mishandled Baggage Reports per 1,000 Passengers

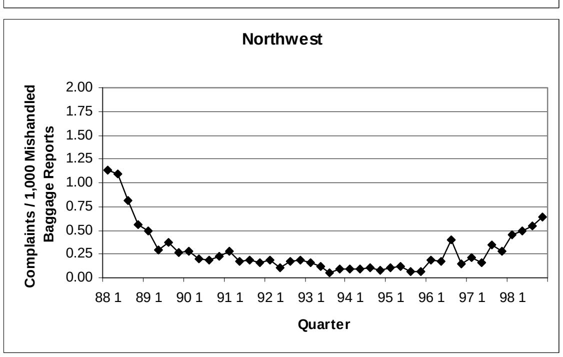




Note. TWA's figure for the first quarter in 1988: 14.4.

Figure 2: Complaints per 1,000 Mishandled Baggage Reports





Note. Continental's figures for the first two quarters in 1988: 2.67 and 3.35.

Figure 3: Selected Estimated U-shaped Learning Curves