Organizational Learning Curves for Customer Dissatisfaction: Heterogeneity Across Airlines

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

Nikos Tsikriktsis
London Business School, Regent's Park, London NW1 4SA, United Kingdom, nikos@london.edu

In the extensive literature on learning curves, scholars have ignored outcome measures of organizational performance evaluated by customers. We explore whether customer dissatisfaction follows a learning-curve pattern. Do organizations learn to reduce customer dissatisfaction? Customer dissatisfaction occurs when customers' ex ante expectations about a product or service exceed ex post perceptions about the product or service. Because customers can increase expectations over time, customer dissatisfaction may not decline even when the product or service improves. Consequently, it is an open question whether customer dissatisfaction follows a learning-curve pattern. Drawing from the literatures on learning curves and organizational learning, we hypothesize that customer dissatisfaction follows a U-shaped function of operating experience (Hypothesis 1), that focused airlines learn faster to reduce customer dissatisfaction than full-service airlines (Hypothesis 2), and that organizational learning curves for customer dissatisfaction are heterogeneous across airlines (Hypothesis 3). We test these hypotheses with quarterly data covering 1987 to 1998 on consumer complaints against the 10 largest U.S. airlines. We find strong support for Hypothesis 1 and Hypothesis 3. Hypothesis 2 is not supported in the sense that the average focused airline did not learn faster than the average full-service airline. However, we do find that the best focused airline learns faster than the best full-service airline. We explore this result by extending a knowledge-based view of managing productivity learning curves in factories to complaint learning curves in airlines.

Key words: learning curve; organizational learning; customer complaints; customer dissatisfaction; airlines

History: Accepted by John Boudreau, organizational behavior, performance, strategy, and design; received September 20, 2004. This paper was with the authors 3 months for 2 revisions.

1. Introduction
The learning-curve phenomenon is well known. As an organization gains experience, organizational performance improves at a decreasing rate. The typical learning-curve example, often observed in manufacturing firms, is the "power curve," which specifies that the logarithm of unit cost decreases linearly as a function of the logarithm of cumulative number of units produced. Scholars have extensively researched learning curves, and managers have often used learning curves for planning purposes (Argote 1999).

Recently, learning-curve scholars have expanded the set of organizational performance measures to include total productivity (Adler 1990), quality (Lapré et al. 2000), timeliness (Argote and Darr 2000), profitability (Ingram and Simons 2002), and organizational survival (Baum and Ingram 1998). Argote (1999) identified the need for learning-curve research to continue to study a wider set of outcome measures. Most learning-curve studies have examined outcome measures evaluated inside the organization. Learning-curve scholars have ignored outcome measures of organizational performance evaluated by customers. In this paper, we study organizational learning curves for customer dissatisfaction, specifically, whether or not organizations learn to reduce customer dissatisfaction.

Quality of a product or service is perceived by customers. Customer dissatisfaction occurs when ex ante expectations about a product or service exceed ex post perceptions of the product or service (Zeithaml et al. 1990). Customer dissatisfaction in turn impedes customer loyalty and repeat purchase (Heskett et al. 1997). Given customers' subjective assessment of quality, it becomes important to study how customer dissatisfaction evolves. Because customers can
increase ex ante expectations, improvements in the product or service may not translate into reduced customer dissatisfaction. Consequently, it is an open question whether customer dissatisfaction follows a learning-curve pattern—decreasing at a decreasing rate. In this paper, we study one outcome of customer dissatisfaction—customer complaints against airlines filed with the government.

This paper is organized as follows. In §2, we first discuss how our outcome measure of customer dissatisfaction differs from traditional learning-curve outcome measures. Next, we develop our hypotheses. In §3, we motivate our learning-curve model. Section 4 describes our data and methodology. Section 5 presents the empirical results. Section 6 concludes with a discussion of our results, limitations, and questions for future research.

2. Learning to Reduce Complaints

2.1. Complaints: A Measure of Customer Dissatisfaction

To appreciate a fundamental difference between customer dissatisfaction and outcome measures previously employed in learning-curve research, consider the following scenario. In 1997, a customer flies on a transatlantic flight in coach on a Boeing 767. She watches a movie projected on the central screen and is satisfied with the in-flight entertainment provided. One year later, she crosses the Atlantic on a Boeing 777. She is pleasantly surprised to find that her seat in coach is equipped with her own personal TV screen, choice of channels, and video games. In 2000, she flies the same transatlantic route and finds to her dismay that her plane is a Boeing 767 not equipped with personal visual in-flight entertainment. The bundle of services changed from 1997 to 2000, even though the specific service of transportation did not change. This customer had increased her expectations about the service. Even though the airline provided the same transportation service in 2000 as in 1997, the customer was dissatisfied because her increased expectations were not met. While this customer may not file an actual complaint, the takeaway from this example is that customer dissatisfaction is a function of the gap between ex ante customer expectations about a product or service and ex post customer perceptions about the product or service (Zeithaml et al. 1990). So, the more perceptions fall short of expectations, the higher customer dissatisfaction will be. As customers can change their expectations over time, dissatisfaction can change even if the product or service remains the same. Therefore, customer dissatisfaction is fundamentally different from previous learning-curve outcome measures such as unit cost, where identical performance by an organization results in identical, objective outcomes. Even if an organization delivers identical results over time, customer dissatisfaction may increase as customers increase their expectations. As a result, it is imperative for organizations to manage the balance of customer ex ante expectations and ex post perceptions.

A learning-curve effect for complaints would capture an organization’s ability to better manage the balance of expectations and perceptions over time. In the introduction, we mentioned that customer dissatisfaction is an important impediment to customer loyalty and repeat purchase. Why is it interesting to study complaints as a measure of customer dissatisfaction? All major U.S. airlines that achieved sustainable cost reductions achieved lasting complaint reductions first (Lapré and Scudder 2004). Hence, while reducing complaints is no guarantee for organizational success, reducing complaints is necessary to achieve lasting cost reduction—which certainly in the airline industry is a must. Next, we develop hypotheses regarding the shape of complaint learning curves, the impact of focus (as opposed to full service) on complaint learning curves, and the variation in complaint learning curves across airlines. Our hypotheses draw predominantly from the literatures on learning curves and organizational learning.

2.2. Hypotheses

March (1991) has argued that organizations face trade-offs between exploration and exploitation. Exploration requires investments in discovery, innovation, and experimentation to address new demands or opportunities. Exploitation, on the other hand, requires investments in refinement and efficiency to get better at executing a given set of routines. Short-term gains from exploitation can make exploration seem less attractive. Conversely, exploring new alternatives leaves an organization with less energy to invest in exploitation. Hence, organizations face a trade-off between exploration and exploitation.

Organizational learning curves have typically focused on exploitation—practice with an existing set of routines makes perfect. Organizations learn by doing the same things over and over again. Exploitation builds both internal and external organizational capabilities (Ingram and Baum 1997). By accumulating operating experience, organizations should become more efficient—the internal capability. There is extensive learning-curve evidence that organizations improve efficiency by gaining experience (Argote 1999). Learning by doing is inferred as efficiency improves as a function of organizational experience, typically measured by the cumulative number of units produced. Cumulative volume is a proxy variable for knowledge acquired through repeated
production. According to Ingram and Baum (1997), operating experience also contributes to the external capability of an organization. Operating experience provides an organization with an opportunity to learn about customer preferences. What do customers want? In the context of airlines, each time an airline operates a flight, the airline gains experience in a variety of activities including checking in passengers, turning around an aircraft, handling baggage, boarding passengers, departing on time, and serving passengers on board. Hence, by repeatedly operating flights, an airline learns how to reduce customer dissatisfaction related to all aspects of flying such as mishandled baggage, denied boarding, late arrivals, and poor customer service. Exploitation in the short run should therefore reduce customer dissatisfaction.

Focus on exploitation at the expense of exploration can hurt organizations in the long run. Focus on exploitation 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. Hence, Ingram and Baum (1997) argue that an organization's rate of failure is a U-shaped function of operating experience. Initially, operating experience decreases the rate of failure (short-run gains of exploitation). Yet, in the long run too much operating experience with a set of outdated routines increases the rate of failure, as outdated routines are no longer adequate to address new demands placed on the organization. Ingram and Baum (1997) and Baum and Ingram (1998) both found empirical evidence supporting a U-shaped relationship between the failure rate of hotel chains and their operating experience. The authors do point out that they expected efficiency to have continued to increase with operating experience. Efficiency is an internal capability—fostered by better knowledge about internal operations, whereas organizational survival is an external capability—fostered by better knowledge about customer preferences. As reducing customer dissatisfaction requires acquiring better knowledge about customer preferences, we expect customer dissatisfaction to follow a U-shaped function of operating experience, just like the failure rate of an organization. As we discussed in §2.1, customers can "raise the bar"—increase expectations based on past experience (Zeithaml et al. 1990). There are two types of customer experience that can lead customers to increase their expectations: past experience with the same organization and past experience with other organizations (competitors). Increasing expectations will increase customer dissatisfaction ceteris paribus. Hence, customer dissatisfaction is a relative measure of performance like organizational survival (Ingram and Baum 1997) or profitability (Ingram and Simons 2002).

In the airline industry, the U.S. Department of Transportation (DOT) released a study in 1993 stating "the dramatic growth of Southwest [Airlines] has become 'the principal driving force' in changes occurring in the airline industry" (Hallowell and Heskett 1993, p. 25). Contrary to the popular hub-and-spoke model in the 1980s, Southwest had substantially grown its model of offering low fares on short-haul, point-to-point routes without operating any hubs. Yet, none of the major hub-and-spoke carriers abandoned their operating model. So, during the timeline of our study (1987-1998), most airlines focused on exploitation as opposed to exploration.

**HYPOTHESIS 1.** Airlines learn from operating experience to reduce customer dissatisfaction in the short run. However, in the long run, customer dissatisfaction follows a U-shaped function of operating experience.

Skinner (1974) introduced the notion of a “focused factory.” A factory that focuses on a narrow product mix for a particular market niche will outperform a plant that attempts to achieve a broader mission. Skinner based his claim on a study of approximately 50 plants in six industries. Simplicity, repetition, experience, and task homogeneity breed competence resulting in better customer service and competitive position. Benefits of focus are not limited to manufacturing. In the service management literature, Heskett (1986) introduced the strategic service vision. The vision comprises the strategic integration of four elements: (i) target market—who are the right customers, (ii) service concept—what are the elements of the service that produce results for the right customers, (iii) operating strategy—what elements of the strategy maximize the difference between perceived value of the service and cost of the service, and (iv) service delivery system—how will the service be delivered consistent with the operating strategy. Every time a customer interacts with a service provider (a "moment of truth"), a customer can judge aspects of the service offering as well as the way the service is delivered. Consequently, every service encounter affects customer dissatisfaction. According to Heskett et al. (1997), companies achieve high profitability by having either market focus (in the target market) or operating focus (in the service delivery system). Moreover, "organizations that achieve both market and operating focus are nearly unbeatable" (p. 8).

Market focus can result in faster learning. In the U.S. hotel industry, Ingram and Baum (1997) found that geographic specialists learned more from their own operating experience than geographic generalists: "Hotel chains that cover many geographic markets might be seen as aggregations of differentiated specialists that learn individually to succeed in local markets, but are unable to successfully transfer
knowledge to each other because they face different market demand” (p. 79).

In the U.S. airline industry, focused airlines such as Southwest achieve both market and operating focus. Focused airlines achieve market focus by cherry picking routes in North America only. Focused airlines achieve operating focus by flying only a few plane types. Simplified operations make it easier to achieve high levels of coordination and teamwork exemplified by fast turnaround times—times planes spend at gates between successive flights. Focus virtually ensures that employees are able to repeat the same actions over and over again, thus moving more quickly down their learning curves both individually and collectively. Through this repetition, employees become entrained to respond to each other and to the uncertainties that impact the flying operation. In contrast, full-service airlines such as American offer both continental and intercontinental services to first-class, business-class, and coach passengers. Full-service airlines operate several hubs. Hub-and-spoke systems provide large domestic coverage and feed passengers into intercontinental routes. Many different types of planes within a fleet, connecting passengers, and longer flights with higher in-flight service requirements make for complicated operations (Heskett et al. 1997, Gittell 2003, Lapré and Scudder 2004). Complicated operations make it much harder to achieve high levels of coordination and teamwork. As Haunschild and Sullivan (2002) note, complex organizations tend to be more political, more hierarchical, and more compartmentalized, which may dampen learning from experience. The authors found that heterogeneity in causes of errors is better for learning, and that focused airlines learn more from heterogeneity in causes of errors. Because focused airlines can achieve higher levels of coordination and teamwork, focused airlines should be able to learn more from customer dissatisfaction caused by errors such as mishandled baggage, denied boarding, and late arrivals. Consequently, focused airlines are likely to learn faster to reduce customer dissatisfaction than full-service airlines.

**Hypothesis 2.** Focused airlines learn faster to reduce customer dissatisfaction than full-service airlines.

Prior learning-curve studies have typically assumed identical learning rates across firms. Dutton and Thomas (1984) compared over 200 learning-curve studies to conclude that a learning rate should not be viewed as a given constant but as a dependent variable to be influenced by management. A few studies in manufacturing settings have demonstrated that learning rates show considerable variation within industries, within firms, and even within plants (Levy 1965, Lapré and Van Wassenhove 2001). A notable exception in learning-curve research in service settings is Pisano et al. (2001). The authors study learning curves for surgery times at 16 different hospitals allowing for a different learning rate for each hospital. The 16 hospitals were the first to implement a new technology for minimally invasive cardiac surgery. Hence, learning rates can differ significantly within the same industry, be it in manufacturing or service settings.

The organizational learning literature provides a rationale for heterogeneous learning rates. Cohen and Levinthal (1990) define absorptive capacity as “the ability to recognize the value of new information, assimilate it, and apply it to commercial ends.” Just like people, organizations have different absorptive capacities. Consequently, organizations will differ in their ability to reflect on new experience and take appropriate actions based on new experience, resulting in heterogeneous learning rates. There are at least two organizational characteristics that might cause heterogeneity in learning curves across airlines.

First, if organizations are faced with problems that are inherently cross-functional, differences in cross-functional absorptive capacity can contribute substantially to heterogeneity in learning rates (Lapré and Van Wassenhove 2001). Gittell (2003, p. 20) identifies 12 distinct functions in the flight departure process. Functions such as pilots, flight attendants, gate agents, fuelers, and baggage transfer agents perform a complex set of tasks. Yet, communication between functions differs greatly between airlines. Because the flight departure process is complex, many things can go wrong, including late departures, mishandled baggage, and boarding problems. Any of these problems can lead to customer dissatisfaction. Southwest assigns each individual flight its own on-site operating agent to facilitate cross-functional communication and coordination. At American, on the other hand, operations agents are located off-site. At any point in time, an operations agent might be involved with 15 different flights. As a consequence, communication across functional boundaries is impaired and joint problem solving and learning suffer. A customer service supervisor at American exemplifies this (Gittell 2003, p. 131):

> Here you don’t communicate. And sometimes you end up not knowing things [...] Everyone says we need effective communication. But it’s a low priority in action. On the gates, I can’t tell you the number of times you get the wrong information from ops [...] We call it the creeping delay. The hardest thing at the gates with off-schedule operations is to get information. They are leery to say the magnitude of the problem.

Second, Edmondson (1999) conceptualizes learning as “an ongoing process of reflection and action,
characterized by asking questions, seeking feedback, experimenting, reflecting on results, and discussing errors or unexpected outcomes of actions" (p. 353). Discussing errors and learning from errors requires what Edmondson (1999) calls psychological safety "a shared belief that the team is safe for interpersonal risk taking" (p. 354). Psychological safety also differs greatly between airlines. Gittell (2003) found that American Airlines’ employees tended to hide information to avoid blame for a delay. An American gate agent stated, “Unfortunately, in this company when something goes wrong, they need to be able to pin it on someone. You should hear them fight over whose department gets charged for the delay” (p. 28). In the early 1990s, Southwest realized that finger pointing and blame avoidance impeded learning (Gittell 2003). Southwest instituted a “team delay” which allowed less precise reporting of the cause of delays, with the goal of diffusing blame and encouraging learning. A station manager at Southwest stated (Gittell 2003, p. 141):

‘[W]e track the source of delays. Usually it’s a situation rather than a person who is at fault. We take a delay when the delay warrants it [...] We try to figure out what caused a delay, but we don’t do much finger pointing. We find that the more you point fingers, the more problems go underground rather than getting solved.

Because of variation in cross-functional communication and psychological safety across airlines, we expect variation in learning curves across airlines.

**HYPOTHESIS 3.** Organizational learning curves for customer dissatisfaction are heterogeneous across airlines.

3. 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, $CR(E)$ the complaint rate after the organization has accumulated $E$ experience, $\mu$ the learning rate, and $CR^*$ the optimal target level for the complaint rate. According to Lapré et al. (2000, p. 600):

$$dCR(E)/dE = \mu[CR(E) - CR^*].$$

A natural target level for $CR^*$ is 0 because we are looking at complaints filed with a third party—the government. The solution for this differential equation is the exponential form

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

or, for estimation purposes

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

Given the theoretical interpretation of the exponential form rooted in the organizational learning literature, we will employ the exponential form in this paper.

Unrelated to the theoretical justification for the exponential form, there is a second benefit of using the exponential form regarding the inclusion or omission of prior experience. For the power form, accounting for prior experience is a major concern in learning-curve estimations. If organizations operated before the availability of data on the outcome measure, omission of prior experience will bias learning-rate estimates. Formally, let $PE_0$ represent Prior Experience up to period 0, and let $E_t$ be Cumulative Experience from period 1 through period $t$. Estimating a learning rate using $\ln(E_t)$ will yield a different coefficient than using $\ln(PE_0 + E_t)$. Loosely put, with the power form, omission of prior experience estimates a different part of the learning curve. With the exponential form, on the other hand, estimating a coefficient with $E_t$ or $PE_0 + E_t$ will yield the same learning rate. In other words, for the exponential form, accounting for prior experience is a nonissue—omission of prior experience will not bias learning-rate estimates. In the next section, we describe our data and research methodology.

4. Data and Method

4.1. Consumer Complaints

We use consumer complaints filed with DOT as our measure for customer dissatisfaction (part of this section draws on Lapré and Scudder (2004)). From 1987 to 1998, passengers could file complaints with DOT
in writing, by telephone, or in person. Complaint categories included flight problems, oversales, reservations/ticketing/boarding, fares, refunds, baggage, customer service, smoking, advertising, credit, tours, and other. Several factors led to a surge of complaints against airlines in 1987: In August 1987 complaints were up by almost 500% 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, starting October 1987, DOT required major airlines to report mishandled baggage, involuntary denied boarding, and on-time arrival statistics. DOT, subsequently, started to publish these statistics along with complaints per 100,000 passengers (the broadest measure of quality) in DOT Air Travel Consumer Reports. In 1999, DOT introduced an e-mail address as an additional channel for filing consumer complaints.

To measure customer dissatisfaction, we use quarterly data on consumer complaints filed with DOT from the fourth quarter of 1987 through the fourth quarter of 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 1987–1998 period, 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 our data set to 1987–1998 controls for ease of reporting complaints. By starting in the fourth quarter of 1987 as opposed to the first quarter of 1987, we also control for increased awareness of filing complaints with DOT.
- We study consumer complaints filed directly with DOT. Airlines were not involved in collecting and reporting these complaints. So, our analysis is not confounded by any changes in reporting patterns by airlines.

Let Complaint Rate, denote the number of consumer complaints per 100,000 passengers against airline i in quarter t. Figure 1 shows the complaint rate evolutions on a common scale. Consequently, Figure 1 allows us to assess individual learning-curve patterns, while Figure 2 allows us to compare complaint rates across airlines.

### 4.2. Independent Variables

We control for factors that provide more opportunities for service failures: long flights and connections. Flight Length,, is the average flight length for airline i in quarter t; Connections,, is the percentage of passengers connecting to another flight for airline i in quarter t. Both variables are constructed using data on travel statistics reported to DOT. The impact of factors such as weather and holidays on complaint reasons can exhibit seasonal trends. We control for any seasonality with dummy variables. We define Quarter2 as 1, if t is the second quarter of a calendar year, 0 otherwise. Similarly, we define Quarter3 and Quarter4.

To measure operating experience, learning-curve scholars typically use cumulative production volume (Argote 1999). As airlines transport "batches" of passengers on flights, airlines need to master all sorts of operations pertaining to flights including turning around a plane for its next flight. We use cumulative number of flights to measure experience. Formally, Experience,, denotes the cumulative number of flights operated by airline i from October 1987 until the previous quarter (t – 1). To construct this experience variable, we obtained flight data from Form 41—which airlines must file with DOT. Later in the paper, we assess the robustness of our findings if we use a different experience variable.

### 4.3. Two Subgroups of Airlines

DOT classifies an airline as major if the airline has at least 1% of total U.S. domestic passenger revenues. Our data set includes the 10 major airlines for the entire 1987–1998 period: Alaska, America West, American, Continental, Delta, Northwest, Southwest, TWA, United, and US Airways. The only other major airlines operating during the 1987–1998 period ceased operations well before 1998: Eastern in 1990 and Pan Am in 1991. Combined, the major airlines account for more than 93% of revenue passenger miles for all U.S. airlines. (One revenue passenger mile is transporting one passenger over one mile in revenue service.)

Alaska, America West, and Southwest focused on routes in North America only. We will refer to these three airlines as "focused airlines." The other seven major airlines offered both continental and intercontinental services. For example, all seven operated transatlantic routes. We will refer to these seven air-
lines as "full-service airlines." Haunschild and Sullivan (2002) used variation in types of aircraft to construct a measure of specialization. We collected data on the number of types of aircraft in an airline’s fleet for each airline for each quarter. The summary statistics showed that the median and the mode for the three focused airlines were 3 or 4. The median and the mode for the full-service airlines ranged between 9 and 14. Furthermore, the variation between the two subgroups was significant, while the variation within
subgroups was not significant. These differences in types of aircraft lend further support to the way we...
in turning around aircraft. Fast turnaround times increase fleet utilization. Lapré and Scudder (2004) showed that the three focused airlines were able to significantly outperform the seven full-service airlines in fleet utilization.

4.4. Method

Our base model is the exponential form without any experience variables. The base model includes control variables for factors that provide more opportunities for service failures, 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 ten airline dummy variables, we need to drop one airline dummy variable. Without loss of generality, we do not include a dummy variable for Southwest:

$$\ln(\text{Complaint Rate}_{ij}) = \beta_0 + \beta_1 \text{Flight Length}_{ij} + \beta_2 \text{Connections}_{ij} + \beta_3 \text{Quarter2}_{ij} + \beta_4 \text{Quarter3}_{ij} + \beta_5 \text{Quarter4}_{ij} + \sum_{j=1}^{10} \beta_{6j} D_{ij} + u_{ij}. \quad (1)$$

In Model (1), $\beta_{6j}$ captures how much the average logarithm of the complaint rate for airline $j$ deviates from the average logarithm of Southwest’s complaint rate $\beta_0$, holding other variables constant. To test the first part of Hypothesis 1 (airlines learn to reduce complaints), we include $\text{Experience}_{i-1}$:

$$\ln(\text{Complaint Rate}_{ij}) = \beta_0 + \beta_1 \text{Flight Length}_{ij} + \beta_2 \text{Connections}_{ij} + \beta_3 \text{Quarter2}_{ij} + \beta_4 \text{Quarter3}_{ij} + \beta_5 \text{Quarter4}_{ij} + \sum_{j=1}^{10} \beta_{6j} D_{ij} + \beta_7 \text{Experience}_{i-1} + u_{ij}. \quad (2)$$

A negative estimate for $\beta_7$ would imply that airlines did learn to reduce complaints. Note that the interpretation of $\beta_{6j}$ changes from Model (1) to Model (2). In Model (2), $\beta_{6j}$ captures how much the intercept, or starting point, of airline $j$’s learning curve deviates from Southwest’s intercept. The second part of Hypothesis 1 stated that a learning-curve effect for customer dissatisfaction is U-shaped. Following Ingram and Baum (1997), we include the square of our experience variable ($\text{Experience}_{i-1}^2$). In Model (2), $\beta_8$ captures how a U-shaped learning curve for full-service airlines differs from the U-shaped learning curve for focused airlines. Lastly, in the full model we allow for each airline to not only have a different intercept, but also a different U-shaped learning curve (Hypothesis 3):

$$\ln(\text{Complaint Rate}_{ij}) = \beta_0 + \beta_1 \text{Flight Length}_{ij} + \beta_2 \text{Connections}_{ij} + \beta_3 \text{Quarter2}_{ij} + \beta_4 \text{Quarter3}_{ij} + \beta_5 \text{Quarter4}_{ij} + \sum_{j=1}^{10} \beta_{6j} D_{ij} + \beta_7 \text{Experience}_{i-1} + \beta_8 \text{Experience}_{i-1}^2 + u_{ij}. \quad (3)$$

Significant estimates for $\beta_7$ and $\beta_8$ would support Hypothesis 3 that organizational learning curves for customer dissatisfaction are heterogeneous across airlines.

For Models (1) through (5), $i = 1, \ldots, N$; $t = 1, \ldots, T$. We have $N = 10$ airlines, and $T = 45$ quarters. Hence, we have time-series cross-section (TSCS) data, not to be confused with “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 Southwest compare with American?"). The asymptotics are in $T$. Consequently, there are three important considerations for the error term $u_{ij}$:
5. Empirical Results

Table 1 shows the regression results for Models (1) through (5). The base model explains 33% of the variation. All nine airline dummy variables are positive and significant indicating that the average complaint rate for each of these airlines was significantly higher than Southwest's average complaint rate.

The estimates for Model (2) support the first part of Hypothesis 1—airlines learned from operating experience to reduce customer complaints. Inclusion of a single experience variable common across all airlines (Experience) explains an additional 3% of the variation. The F statistic for the increase in $R^2$ is 22.48, which is significant at 0.001. The negative estimate for Experience and the positive estimate for (Experience)$^2$ in Model (3) support the second part of Hypothesis 1: Complaints show a U-shaped relationship with experience. The F statistic for the increase in $R^2$ of 18% due to the inclusion of (Experience)$^2$ is 170.54, which is significant at 0.001. So, the U-shaped learning curve in Model (3) explains 21% of the variation (compared to the base model), whereas a single experience variable in Model (2) explains only 3% of the variation. These estimates provide strong support for Hypothesis 1.2

The estimates for Model (4) show that the U-shaped learning curve for full-service airlines does not differ from the U-shaped learning curve for focused airlines. Neither Experience $\times$ Full Service nor (Experience)$^2$ $\times$ Full Service has a significant impact on complaint rate. Consequently, Hypothesis 2 is not supported.

In contrast, estimates for Model (5) show a significant increase in $R^2$ of 20%. The $F$ statistic for the increase in $R^2$ is 21.20, which is significant at 0.001. Several airlines have significantly different slope parameters for Experience and/or (Experience)$^2$. Hence, Hypothesis 3 is supported—organizational learning curves for customer dissatisfaction are heterogeneous across airlines.

Every service encounter with a customer is a moment of truth. Every service encounter offers a firm an opportunity to learn about customer expectations and how to serve customers better. We use cumulative number of passengers as another measure of experience to capture "learning by serving customers." Estimating Models (2) through (5) with cumulative passengers yields similar $R^2$ values for each model compared to Table 1. Moreover, Hypothesis 1 and Hypothesis 3 are supported, while Hypothesis 2 is rejected. In addition, the increase in $R^2$ going from Model (2) to Model (3) is larger than going from Model (1) to Model (2). Hence, none of our conclusions would change.

Why was Hypothesis 2 not supported? We can use the estimates for Model (5) to further explore learning curves for full service versus focused airlines. Figure 3 shows estimated U-shaped learning curves for selected airlines. In order to compare different learning-curve patterns without making Figure 3 too crowded, we excluded the following airlines:

- United and US Airways, which like American have similar comparisons with Southwest: a higher intercept ($\beta_0 > 0$), a slower learning rate ($\beta_1 < 0$), and a slower "relapse rate" ($\beta_2 < 0$).

The 1988-1998 year dummies overlapped, while the estimates for Experience and (Experience)$^2$ did not change. So, the U-shaped learning curve as a function of experience remained, even if we included year fixed effects. Second, instead of year fixed effects, we introduced calendar time in Model (3). Calendar time was not significant. Third, we introduced other time-varying measures that may affect customer dissatisfaction: load factor and real prices. Load factor, a standard utilization measure in the airline industry, is calculated as revenue passenger miles divided by available seat miles for each airline, for each quarter. (One revenue passenger mile is transporting one passenger over one mile in revenue service; one available seat mile is flying one seat over one mile available for revenue service.) To calculate real prices, we expressed average one-way fares for each airline, for each quarter, in 1987 dollars using the Airline Composition Cost Index from the Air Transportation Association (Lapré and Scudder 2004). Inclusion of either load factor or real prices in Model (3) did not change the results—neither variable was significant. Finally, if we included all three variables (calendar year, load factor, and real prices) in Model (3), none of the three variables were significant.

---

1 We ran additional analyses to test if the U-shaped effect is caused by historical factors associated with the passage of time rather than operating experience. First, we introduced dummy variables for each calendar year in Model (3). All 95% confidence intervals for
Table 1  Complaint Rate Learning-Curve Estimates

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
<th>Model (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.155</td>
<td>-0.588</td>
<td>-0.156</td>
<td>-0.290</td>
<td>0.802*</td>
</tr>
<tr>
<td>(0.355)</td>
<td>(0.318)</td>
<td>(0.266)</td>
<td>(0.286)</td>
<td>(0.347)</td>
<td></td>
</tr>
<tr>
<td>Flight length</td>
<td>-0.239**</td>
<td>-0.022</td>
<td>0.054</td>
<td>-0.076</td>
<td>-0.130</td>
</tr>
<tr>
<td>(0.080)</td>
<td>(0.072)</td>
<td>(0.060)</td>
<td>(0.059)</td>
<td>(0.080)</td>
<td></td>
</tr>
<tr>
<td>Connections</td>
<td>1.034*</td>
<td>0.634</td>
<td>0.342</td>
<td>0.247</td>
<td>0.024</td>
</tr>
<tr>
<td>(0.499)</td>
<td>(0.479)</td>
<td>(0.422)</td>
<td>(0.420)</td>
<td>(0.488)</td>
<td></td>
</tr>
<tr>
<td>Quarter2</td>
<td>-0.185***</td>
<td>-0.191**</td>
<td>-0.189***</td>
<td>-0.189***</td>
<td>-0.179***</td>
</tr>
<tr>
<td>(0.062)</td>
<td>(0.061)</td>
<td>(0.051)</td>
<td>(0.050)</td>
<td>(0.050)</td>
<td></td>
</tr>
<tr>
<td>Quarter3</td>
<td>-0.026</td>
<td>-0.043</td>
<td>-0.045</td>
<td>-0.045</td>
<td>-0.012</td>
</tr>
<tr>
<td>(0.074)</td>
<td>(0.073)</td>
<td>(0.061)</td>
<td>(0.060)</td>
<td>(0.062)</td>
<td></td>
</tr>
<tr>
<td>Quarter4</td>
<td>-0.159*</td>
<td>-0.160**</td>
<td>-0.171***</td>
<td>-0.173***</td>
<td>-0.181***</td>
</tr>
<tr>
<td>(0.061)</td>
<td>(0.060)</td>
<td>(0.050)</td>
<td>(0.049)</td>
<td>(0.049)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Dependent variable ln(Complaint Rate). Panel-corrected standard errors in parentheses.

- Continental and TWA, which have a similar U-shaped learning curve as Southwest except for a higher intercept ($\beta_{x_2} > 0$).

While the learning curves for focused airlines Southwest and Alaska outperform the learning curves for all full-service airlines, the learning curve for focused airline America West is in the range of the full-service airline learning curves. So, in Model (4), the learning curve for the average focused airline did not outperform the learning curve for the average full-service airline. As Figure 3 shows, there is significant overlap between the two groups of airlines. In fact, one full-service airline (Delta) even outperforms one focused airline (America West).

Figure 3 demonstrates that organizations need to manage two elements of U-shaped complaint learning curves: the learning rate associated with Experience ($\beta_2 + \beta_7$ in Model (5)) as well as the relapse rate.
6. Discussion and Conclusion

The contributions of our study are threefold. First, we extend learning-curve analysis to a measure of customer dissatisfaction. Second, we provide evidence of a U-shaped learning curve in the context of customer dissatisfaction. Third, we find heterogeneity in organizational learning curves for customer dissatisfaction across airlines. For each contribution, we discuss our results, limitations, and questions for future research.

6.1. Extension of Learning-Curve Analysis to Customer Dissatisfaction

Most learning-curve studies have focused on outcome measures evaluated inside organizations. Most learning-curve scholars have ignored organizational performance evaluated by customers. We estimate learning curves for customer dissatisfaction. Because customers can increase expectations over time, one cannot a priori expect a learning-curve pattern for customer dissatisfaction. Our study shows that airlines did exhibit learning curves for complaint rates, suggesting that airlines did learn to reduce customer dissatisfaction, at least in the short run. Just like organizations need to learn proactively about operations to improve measures of internal capabilities such as efficiency, organizations need to engage in “continuous learning about markets” (Day 1994) to improve measures of external capabilities such as customer dissatisfaction. Day (1994, p. 115) gives an example of the continuous experimentation required to learn about markets:

American Airlines found that customer perceptions of on-time arrival performance improved markedly if the plane doors were opened less than 25 seconds after gate arrival. The key to this insight was their ability to measure how quickly they opened the doors and their follow-up with telephone surveys of the passengers on the flight. In effect, they took advantage of a series of natural experiments.

One limitation of our study is the focus on an outcome measure of customer dissatisfaction without addressing the underlying gap between expectations and perceptions. Future research is needed to disentangle ex ante expectations and ex post perceptions. Which way is most effective to manage customer dissatisfaction learning curves—manage expectations, perceptions, or both?

A related limitation of our study concerns our traditional interpretation of customer dissatisfaction: ex post perceptions falling short of ex ante expectations about a single transaction with a customer. Recently, marketing scholars have started to explore alternative theories for customer (dis)satisfaction. Fournier and Mick (1999) suggest that transaction-specific assessments of (dis)satisfaction could be incomplete. The authors find that (dis)satisfaction with new product technologies is a dynamic process, often has a social dimension, and is affected by emotion. Longitudinal dissatisfaction data on individual customers would enrich our understanding of dissatisfaction learning curves. It should be fruitful, for example, to study how airlines manage dissatisfaction of its frequent fliers.

A promising avenue for future research would be to study organizational learning curves for service failure and recovery efforts. Although organizations...
should avoid failures when possible, failures contain many informative lessons. However, in most organizations "there are few incentives to study failures carefully" (Day 1994). According to Tax and Brown (1998), the majority of customers are dissatisfied with the way companies resolve customer complaints. Because customer (dis)satisfaction is an important driver of customer loyalty and profitability, it is imperative for companies to not only recover but also learn from service failures (Tax and Brown 1998).

6.2. U-shaped Learning Curves for Customer Dissatisfaction

We found strong evidence that complaint rates followed a U-shaped function of experience. Thus, we contribute to the literature on U-shaped learning curves documented for organizational survival rates (Ingram and Baum 1997, Baum and Ingram 1998) and profitability (Ingram and Simons 2002). Firms who are serious about learning from service failures should encourage dissatisfied customers to voice their complaints, so that firms can actually learn from service failures (Tax and Brown 1998). A U-shaped effect for complaints received directly by a firm is not necessarily undesirable, because customers are at least increasingly informing a firm about service failure. It is therefore important to note that our measure of customer dissatisfaction does not concern complaints filed with airlines, but complaints filed with a third party—the government. Moreover, DOT uses these consumer complaints to rank individual airlines and publishes these rankings every month.

One implication of the U-shaped function is that eventually complaint rates would continue to get worse. We had to limit the study period to 1987–1998 to create a controlled experiment as much as possible. DOT's introduction of its e-mail address as well as posting preformatted complaint forms on its website in 1999 hugely facilitated customers to file complaints with DOT. There is clearly a need to study longer periods to research which organizations are able to improve again and under what conditions.

Future research on organizational learning curves for service failure and recovery efforts could further tease out our U-shaped learning-curve finding. In a study of bank customers who complained twice with the same bank within 20 months, Maxham and Netemeyer (2002) found that customers reporting two failures have higher recovery expectations for the second failure than for the first failure. Hence, it is possible that customers' raising the bar for recovery expectations contributes more to the U-shaped effect than service failure does.

6.3. Heterogeneity in Learning Curves for Customer Dissatisfaction

We did not find support for Hypothesis 2—the average focused airline did not learn faster than the average full-service airline. However, we did find that the best focused airline learned faster than the best full-service airline. While there are no guarantees for a focused airline to outperform full-service airlines, focus provides a potential benefit. Heterogeneity in organizational learning curves across airlines implies that it depends on effective management of organizational learning curves to realize the potential benefit of focus. Laprée and Van Wassenhove (2003) argued that successful management of productivity learning curves in factories requires careful management of knowledge creation and transfer. Similarly, in the strategy field, the knowledge-based view argues that knowledge is the most critical resource of the firm (Grant 1996). Consequently, heterogeneity in knowledge resources across firms is a key determinant of superior performance and sustained competitive advantage. We found strong support for heterogeneity in complaint learning curves (Hypothesis 3). We also found superior complaint performance and sustained competitive advantage in the area of customer dissatisfaction for Southwest. As Figure 2 shows, Southwest has the best complaint performance among the major U.S. airlines: the lowest starting point in 1987, the lowest ending point in 1998, and in between the lowest complaint rate for more quarters than all other major airlines combined. Estimates in Table 1 confirmed that, for 1987 through 1998, Southwest has the superior learning curve. Next, we explore some of Gittell's (2003) observations to extend a knowledge-based view of managing productivity learning curves in factories to complaint learning curves in airlines.

First, Gittell (2003) discusses how knowledge sharing by employees differs between American and Southwest. At American, employees showed little awareness of the overall process. They typically explained their jobs without referring how their jobs contributed to the overall process.

By contrast, interviews with Southwest frontline employees revealed that they understood the overall work process—and the links between their own jobs and the jobs performed by their counterparts in other functions. When asked to explain what they were doing and why, the answers were typically couched in reference to the overall process. These descriptions by Southwest employees typically took the form, "The pilot has to do A, B, and C before he can take off, so I need to get this to him right away." Rather than just knowing what to do, Southwest employees knew why, based on shared knowledge of how the overall process worked. (Gittell 2003, p. 32).
Just like in manufacturing, the mix of know-why, or causal understanding, and know-how, or operational skills is important for sharing knowledge (cf. Lapré and Van Wassenhove 2003).

Second, a critical activity in developing shared knowledge is joint problem solving. At American, Gittell (2003) observed that 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? (Gittell 2003, p. 37)

In contrast, Southwest employees conduct root-cause analysis together. A Southwest pilot:

We figure out the cause of the delay [...] It is a matter of working together. Figuring out what we can learn. Not finger pointing. (Gittell 2003, p. 37)

So, joint problem solving at Southwest is geared toward increasing the shared knowledge base.

Third, solving problems that cross departmental boundaries can be a tough challenge. In factories, Lapré and Van Wassenhove (2003) found that seasoned engineers with a rich diverse knowledge base across production departments could really accelerate learning curves. At Southwest, Gittell (2003) observed the critical role of “boundary spanner” played by the operations agent responsible for turning around a flight. The boundary spanner communicated face to face with each function, thus facilitating joint problem solving across functions. Moreover, boundary spanners also reduced customer dissatisfaction. Yet, when other airlines heard about Southwest’s staffing levels for the operations agents, their response was “What a waste!” or “That is so inefficient!” (Gittell 2003, p. 258). However, Southwest’s superior performance in cost, customer complaints, and profitability shows there are clear payoffs to investing in boundary spanners.

Gittell’s (2003) study seems to support a knowledge-based view of managing learning curves. The same knowledge aspects that are important for managing productivity learning curves in factories seem to have contributed to Southwest’s superior complaint learning curve. The observed differences in learning rates, relapse rates, and starting points for complaint learning curves indicate that competitive advantages can accrue from managing complaint learning curves better than others. Just like productivity learning curves, a complaint learning curve is not a given; instead, it needs to be managed proactively.

A limitation of our study is the lack of longitudinal variables describing learning processes or knowledge creation processes. Future research that incorporates learning or knowledge creation to “go behind the complaint learning curve” like Adler and Clark (1991) and Lapré et al. (2000) did for productivity and quality measured inside organizations would be a significant contribution. Future research is also needed to assess the generalizability of our findings obtained in a single industry in a single country. We hope that research along these lines will advance our understanding of how organizations can better manage learning curves.

Acknowledgments
The associate editor, two anonymous reviewers, participants at the 2004 INFORMS meeting in Denver, seminar participants at Vanderbilt University, and Gary Scudder provided very 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


