Organizational Learning and the Transfer of Knowledge: An Investigation of Quality Improvement

Daniel Z. Levin

Organization Management Department, Faculty of Management, Rutgers University, 111 Washington Street,
Newark, New Jersey 07102, levin@business.rutgers.edu

Abstract

Whereas most prior research on the learning curve has focused on improvements in efficiency, this paper deals with the impact of learning on product quality. The key data are measures of automobile reliability published in *Consumer Reports*. Analysis yields three findings: (1) Quality improves over the production life of a car model with the same kind of regularity as an efficiency learning curve. Thus, there is a quality learning curve. (2) Unlike in the efficiency domain, however, learning in the domain of product reliability is primarily a function of time, and not of how many cars have gone down the assembly line. Thus, quality depends not on the accumulation of production experience per se, but on the intensity of "off-line" quality improvement activities and on the transfer of knowledge from the general environment over time. (3) In contrast to the traditional injunction, "do not buy a new car in its first year of production," the opposite advice actually seems to apply: In any given year, the newest car models have the best quality. That is, new car-model designs typically include significant quality improvements that are more than enough to outweigh any disruption created in manufacturing by the new model's introduction and that even surpass the incremental improvements made to older, existing car models.
(*Learning Curve*; *Knowledge Transfer*; *Product Reliability*; *Quality*; *Cars*; *Auto Industry*)

The fact that people get better at doing things as they gain experience, but at a decreasing rate, is captured in what is often called a learning curve. This type of improvement has been well documented at many different levels of analysis: for individuals (Mazur and Hastie 1978), small groups (Leavitt 1951), work shifts within a factory (Epple et al. 1991), entire factories (Argote et al. 1990), multifactory organizations (Irwin and Klenow 1993), and industries (Udayagiri and Balakrishnan 1993). Yet despite the long history of research on organization-based learning curves going back at least to Wright (1936), the phenomenon's theoretical underpinnings have remained murky (Kantor and Zangwill 1991).

The present study contributes to the effort to identify the factors underlying the learning curve—to show that the learning curve does not just "happen" (Adler and Clark 1991, Zangwill and Kantor 1998). To help outline these theoretical underpinnings, it may be useful to review briefly the "five W's" (*who*, *what*, *where*, *when*, and *why*) of organizational learning curves. This study suggests we need to rethink in particular the *what* and *when* of learning curves.

Who does the learning? Argote (1993) focuses on three broad categories: individual employees, organizational systems (e.g., coordination, technology), and actors in the organization's environment (e.g., suppliers, competitors). According to this classification, then, the "learning" underlying a learning curve is achieved by a combination of employees, organizational systems, and outside actors.

What is learned? In most learning curve studies, the measure that shows improvement is productivity, which is an important element of organizational effectiveness, but only one element. A key goal of this study is to extend the boundaries of the learning curve concept to test its applicability to another crucial element of performance, quality improvement. While some might assume that whatever research applies to productivity learning must also apply to all other aspects of organizational performance, such an assumption is questionable. After all, organizational performance itself is not a single coherent construct; rather, it is an "umbrella" construct that includes many distinct organizational goals (Hirsch and Levin 1999). Thus, what we know about productivity-based learning may or may not carry over into the quality domain.

Where does the learning take place? In answering this question, researchers have emphasized the important role of cognitive factors, such as search algorithms (Muth 1986), and of behavioral factors, such as engineering changes and training (Adler and Clark 1991). More recently, Zangwill and Kantor (1998, p. 913) have proposed that learning and improvement, though often observed at the aggregate level, take place in ''numerous small activities [such as] entering sales data, billing, contacting clients, training, telephoning.''

When does organizational learning take place? This question has not been widely or systematically addressed by the literature. This study tries to fill this gap by extending the theoretical argument advanced by Zangwill and Kantor (1998). Specifically, this paper argues that, when an organization's main learning processes for noticing and improving problems are a function of time rather than accumulated production experience, then the overall learning curve will be a function of time rather than cumulative experience. This argument, presented in more detail below, runs counter to the usual finding in the learning curve literature, where ''studies have found that calendar time becomes statistically insignificant once cumulative output is included in the analysis'' (Lieberman 1987, p. 442).

This study will not only explore learning *during* the learning curve, but also any learning *before* the learning curve even begins. For many years it has been routinely recommended never to buy a new car model in its first year of production (*Consumer Reports* 1982, p. 11; Harbour and Associates 1990, p. 52). This advice implies that quality-related learning proceeds faster during production than in the new car model design process. Consistent with the arguments of Tyre and Orlikowski (1994), however, I find evidence that the opposite is true: namely, that managers typically make the largest quality improvements during certain ''windows of opportunity,'' and one of those key windows is the period before product introduction. This early knowledge determines the starting point of a product's eventual learning curve. The current study extends this idea of a window of opportunity by proposing that the larger the opportunity, the greater the amount of learning and improvement activities that can occur. For a car model's minor update or facelift every year (the learning curve), the improvement will be relatively small; for a major model change or all-new car model introduction, the learning will be larger.

Why does learning occur? Few researchers have asked this question because it is hard to imagine why people would not want to improve. As long as learning curve research focuses only on costs, researchers will only find instances in which the motivation to learn is present. Yet,

while firms typically have a strong economic incentive to reduce their costs, they do not always have an incentive to improve product reliability. However, by the 1980s, U.S. automakers finally had a major strategic incentive in the form of Japanese competition. As one industry analyst noted, ''The decade [of the 1980s] started with product quality being the domestic industry's number one problem. It was a major reason why it was losing market share to Japanese imports at an alarming rate'' (Harbour and Associates, 1990 p. 142). The auto industry, then, should be an especially good context in which to find and examine learning curves for quality.

In sum, this study extends our understanding of learning curves in three main ways. First, it expands the *what* of learning curves beyond efficiency to include a new area, quality improvement. Second, in terms of *when* (and perhaps *how*) learning occurs, it provides evidence that at least some kinds of learning are time-based instead of experience-based. Third, by comparing the two times *when* learning can occur, this study shows that considerably more improvement occurs as a result of new product introduction than during any year of subsequent production (i.e., the learning curve).

Hypotheses

Quality improvement in the car industry arises out of two processes occurring at the product level. For a given car model at a given point in time, quality is a function of (1) improvement made during the car model's production life, and (2) the car model's initial baseline level of quality. Most learning curve research focuses on the first process (''the learning curve'' itself), but, especially in the car industry, the second process (new product development) is also important. I begin by investigating the learning curve itself, then step back to see the effects of new product introduction, and, finally, I compare the effects of each process. The first three hypotheses (H1–H3) address the rate of improvement for the learning curve: Does a learning curve for quality exist? If so, what form does it take? What factors influence it? Then, separate from these questions of what happens during the learning curve itself, Hypothesis 4 asks what happens during new product introduction, before the learning curve even starts: Is there improvement in the ''starting points'' of these learning curves? Finally, Hypothesis 5 compares these two types of learning—annual learning-curve improvements versus annual starting-point improvements—to see which has a greater impact.

Quality can be defined in many different ways (Garvin 1988). The definition of quality used in this study is a product's reliability, measured by frequency-of-repair

rates. While it is interesting to look at ''internal'' measures of quality, such as conformance to specifications, it is arguably even more interesting to take the customers' point of view and focus on their experience of quality. Improvements in quality, and more specifically auto repair rates, should take the form of a learning curve. After all, ''we might expect that when a car has been in production several years, management should have invested in identifying and solving most of the product's quality problems. In other words, there ought to be a learning curve'' (Cole 1990, p. 80). We expect this quality improvement to take the shape of a learning *curve* for the same reason that we see a learning curve for productivity improvement: namely that, after a while, ''the easiest gains have already been made—the cream has been skimmed. New gains in product quality [then come] more slowly and appear to [get] more expensive'' (Cole 1990, p. 77). The implication is that quality improvement in automobiles will show a learning curve for quality, as measured by repair rates.

HYPOTHESIS 1. *The average car model improves in quality during its production life, but in decreasing increments.*

Note that Hypothesis 1 (H1) leaves open the question of whether this quality-improvement learning curve is a function primarily of experience gained over *time* or as a result of cumulative *production volume* experience. Granted, these two possibilities are highly correlated, and their effects are not always so easy to tease apart (Adler 1990). Yet, from a theoretical perspective, this open question is an interesting one because it helps us to get at some of the underlying sources of organizational learning. In the well-documented domain of productivity-based learning curves, Argote (1993, p. 41) notes, ''empirical studies have demonstrated that calendar time is generally not a significant predictor of organizational learning [once one takes into account the role of] cumulative output at the particular group or organization.'' We might therefore expect a similar effect, in which cumulative experience (based on production volume) is more important than the ''mere'' passage of time, for quality as well. After all, the growing popularity in the manufacturing domain of continuous improvement and statistical process control as a means of improving quality in the auto industry suggests that cumulative production experience should be the driver of quality improvement.

HYPOTHESIS 2. *Quality improvement during the average car model's production life is a function of cumulative production output.*

There is, however, reason to question this expectation,

for in the quality domain, one might expect differences in (1) how people *notice* problems of poor reliability and (2) how they *solve* such problems.

There is, by definition, a time lag between when a product is produced and when producers start to notice how reliable it is. That is, producers get less immediate feedback for reliability than they do for efficiency, or even for defects. Fortunately for automakers, it does not always take years to get feedback on product reliability. For example, they can get early warning signs from warranty claims, high-mileage vehicles like taxis and rental cars, and from performing their own endurance tests. However, this type of feedback does not depend especially on how many units of a given car model were made. Rather, it depends on how intensively, during a given period of time, people focus on gaining this knowledge.

The use of this feedback to solve problems of poor reliability may also be more a function of time than of production experience. The total quality management (TQM) movement has relied heavily on off-line improvement teams. These teams first identify the root causes of problems, then propose, design, test, and implement solutions to the most important root cause until it is corrected; then move on to the next most important source of problems. Indeed, this improvement cycle ''is the iterative learning loop at the heart of TQM'' (Sterman et al. 1997, p. 504). As a consequence, learning in the quality domain is likely to come not so much from how many cars have gone down the assembly line, but from the intensity of ''off-line'' activities. And indeed, past studies of TQM have assumed that the form of improvement for quality improvement activities is largely a function of time (Schneiderman 1988, Sterman et al. 1997).

Given the importance of time-dependent quality-improvement cycles for solving reliability problems, as well as the time lag for how these problems get noticed in the first place, a competing hypothesis to H2 would be:

HYPOTHESIS 2–ALT. *Quality improvement during the average car model's production life is a function of time.*

Learning in an organization can be a function not only of the organization's own internally generated knowledge, but also of knowledge transferred from the organization's environment (Argote 1993). What are some likely sources of transferred knowledge that would enhance a car model's learning curve for quality? One likely source might be the makers of other, closely related car models. Prior research, after all, has found that knowledge transfer occurs within similar categories of nuclear power plants (Lester and McCabe 1993), pharmaceutical research programs (Henderson and Cockburn 1996), and semiconductors (Udayagiri and Balakrishnan 1993). In

the automobile industry, the opportunity to share similar information and technology could potentially occur among car models within the same product ''family,'' called a vehicle platform. Even car models from different product families often share technological processes, suppliers, and certain common parts like seatbelts. It is also easier to transfer knowledge across groups or organizations if they have a shared culture (Argote 1993), so among car models in the same division or company, we might expect to see some knowledge transfer. Among U.S. automakers, too, there is a shared ''Detroit'' culture of sorts (Ingrassia and White 1994), as well as some shared suppliers (Harbour and Associates 1990).

Along these lines, several prior studies have investigated cumulative output-based knowledge transfer (also called ''spillovers''), where the cumulative production output of one group or organization benefits another (Argote 1993, Darr et al. 1995, Irwin and Klenow 1993). Thus, it may be possible to detect production-based knowledge transfers in the quality domain from the following sources:

HYPOTHESIS 3. *The average car model's quality improves as a function of the cumulative production output of its (a) platform, (b) division, (c) company, and (d) industry.*

We should note an alternative view of these spillover effects: It is also possible that the external knowledge available to stimulate quality improvements accumulates in a manner more related to the passage of time; e.g., by reverse engineering a competitor's product, or by learning from a supplier's technological breakthrough, or by avoiding the mistakes of earlier projects within the firm. We might therefore imagine an H3–ALT, focusing on *time.* This process, however, is already captured by H2–ALT, because car models from the same platform typically debut at the same time and because U.S. automakers and their divisions began operations too many decades ago for any differences in founding dates to be meaningful in this context. Indeed, prior research (Argote 1993) has identified the passage of time (e.g., as in H2–ALT) with a generalized knowledge transfer, or ''march of progress.'' H2–ALT thus reflects, in a general way, the benefits to car model producers from others' knowledge of how to solve reliability problems.

So far we have focused on the rate of improvement for existing products; i.e., the learning curve. Over time, however, companies introduce new products with major—even dramatic—improvements already built in. Surely, then, no picture of quality improvement and learning in the auto industry is complete without considering the learning that takes place as a result of new product introductions, before a car model's learning curve

even begins. For example, previous studies (*Consumer Reports* 1991, p. 248; Harbour and Associates 1990) have documented the fact that each of the Big Three U.S. automakers—Chrysler, Ford, and General Motors (GM)—made significant *overall* strides in improving quality during the 1980s. But, if we return to the ''five *W's*'' of learning curves, we must ask *when* this improvement takes place. That is, do these overall improvements come from (1) incremental improvements made to existing car models (the learning curve examined in H1–H3), or from (2) the introduction of all-new car models to replace poor-quality ones? This second type of improvement affects the ''starting point,'' or baseline performance level, of the learning curve. The null hypothesis would be that, over time, all of an automaker's new car models, on average, have the same baseline quality level; i.e., a debuting car model must start from scratch, with no carryover of learning from previous car models and with no benefit from any newly created knowledge. In contrast, a ''march of progress'' view would suggest that it *is* likely that makers of debuting car models learn from the past. Moreover, in this view, the debut of a new model provides a window of opportunity (Tyre and Orlikowski 1994) for manufacturers to incorporate the very latest knowledge on design, materials, assembly, etc., because the usual constraints (e.g., the large costs in changing any production line technologies) are temporarily lifted when a new product line debuts.

HYPOTHESIS 4. *The later a car model debuts, the better is its baseline quality level.*

This hypothesis is tested in terms of time; i.e., it focuses on the year a car model debuts. (Note that it might also be interesting, as part of a future research study with the benefit of a longer time series, to test a comparable H4–ALT, based on cumulative production output, as well.)

If H4 is supported, then we can say that there is at least *some* learning present before a new learning curve begins. But how does this new-and-improved starting point compare to the benefits of being at the end of a long learning curve? Is this ''debut-year'' learning enough to compensate for the lack of production experience? Prior research (e.g., Hayes and Clark 1986) tells us that the introduction of a new product into a repetitive manufacturing context can be highly disruptive, impairing efficiency and sometimes dramatically so. Such disruption is also likely to impair quality. This ''disruptive learning'' perspective would suggest that, even if the baseline quality improves, it will still take time before product quality reaches (and eventually surpasses) its previous level. After all, one could argue, any learning requires at least some ''unlearning,'' and so some disruptions are probably inevitable.

This view is expressed in the old industry adage, "Don't buy cars in their first year" (*Consumer Reports* 1982, p. 11; Harbour and Associates 1990, p. 52). This saying is based on the premise that the learning that ultimately occurs during production will outweigh any reliability enhancements embedded in a brand-new product.

By contrast with this adage and with the cases discussed by Hayes and Clark (1986), Garvin (1988), who investigated the defect rates of factories making room air conditioners, found that although defects did worsen for a day or so after the annual shutdown period, no effect was detectable in the quarterly data. That is, new models pick up the quality levels right where the older models leave off. According to this "constant learning" view, debut-year improvements and learning-curve improvements are equal, so that there is a constant rate of learning for automakers as they transition from older car models to newer ones.

A third alternative derives from a theoretical perspective we might call "enhanced learning." This perspective suggests that debuting car models will benefit from some combination of the carryover from previous learning, minus a few minor disruptions, *plus* the exploitation of the larger window of opportunity provided by the debut. This last point implies that more extraordinary or exceptional learning will occur during product debuts, and have a greater effect, than the more incremental changes made during a product's production life (Lant and Mezias 1992, Tushman and Romanelli 1985). Although the basic design of cars has not changed fundamentally during the past century, the introduction of a new product is nonetheless an important discontinuity in design, manufacturing, and sales. Thus, relative to day-to-day changes, or even to annual model changes, "major" model changes or debuts can indeed involve more "radical" rethinking. Assuming firms are motivated to improve quality, as U.S. automakers surely were in the 1980s, then it follows that the larger the opportunity to improve, the greater the learning and improvement will be. That is, although the window of opportunity to improve a car model *during* its production life occurs at least every year, these opportunities are relatively small in scope. Change too much and factory costs will rise; take too long to make changes while the assembly line is shut down, and lost product sales will mount. But the window of opportunity is much larger when a new product first debuts. So much is changing anyway that the marginal cost of making major quality improvements is likely to be relatively low. Furthermore, to minimize costs and disruptions, organizations may wait until the next new product is introduced to incorporate new manufacturing and product technologies that will enhance reliability. Thus, the large window of

opportunity makes possible a greater improvement in product quality from the organizational learning embedded in new products.

So is this debut-year learning enough to compensate for the associated disruption? Figure 1 shows the three possible answers we have discussed. (Note that the vertical axis in Figure 1 is in logarithmic form, so the learning curves appear linear.) The final hypothesis, which is based on the "enhanced learning" perspective, frames the key test:

HYPOTHESIS 5. *In any given year, car models with the latest debuts will have the best quality, despite having a shorter production life in which to make improvements.*

Methods

Unit of Analysis

Unlike many learning curve studies that focus on a company or factory, the unit of analysis in this study is the car model, which may be built at multiple sites. One reason for this choice is that analyzing the data at the level of the product or product line allows better control for the many technological changes and innovations that occurred over the course of this study's time frame (Abernathy and Utterback 1978). Moreover, the designers, managers, and assemblers associated with a single car model share both the same corporate and task environment. Hence, they make up a kind of organizational

Figure 1   Three Alternatives for How a New Car Model's Debut Might Affect the New Learning Curve's Starting Point
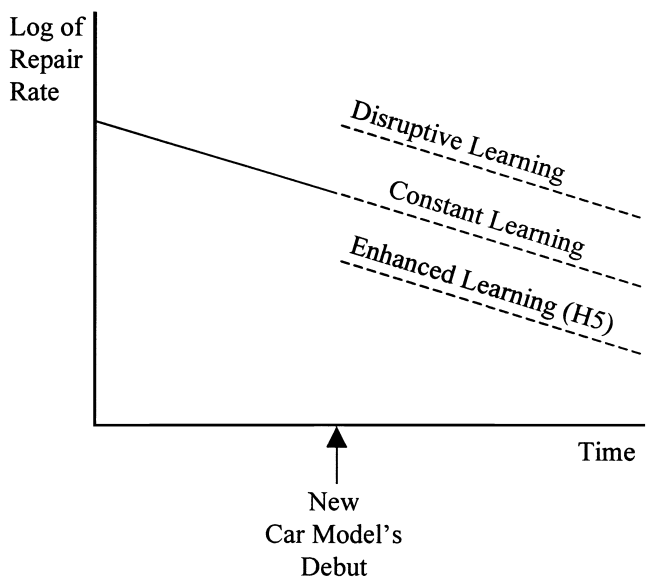
unit, regardless of how geographically spread out they may be. One drawback to this approach is that quite a number of car models in the 1980s were built at multiple plants, but the *Consumer Reports* data on car model quality,[1] used in this study, cannot be broken down by plant. Thus, there is unobserved variation in plant performance across sites building the same model.

This study's sample includes all Big Three passenger car (not truck) models assembled exclusively in the United States and Canada. This group was chosen because North American carmakers formed an important and cohesive industry group facing, to one degree or another, a common set of competitive and quality problems in the 1980s (Harbour and Associates 1990, Ingrassia and White 1994). Although every car model is redesigned slightly from year to year, the unit of analysis for this study is defined as the car model produced over the four- to thirteen-year period following a major overhaul. Because the difference between a major and a minor overhaul is not always obvious, and only sometimes coincides with a name change, the determination of a major redesign or introduction was done by coding the descriptions in *Ward's Automotive Yearbook* (1974–1989) of each year's car models. Typically, *Ward's* described major changes with interchangeable phrases like *all-new* or *debuting* or *bowing*. Coding was done separately by the author and an assistant blind to the hypotheses, with agreement in 93.9% of the cases (Cohen's kappa = 0.75). Agreement increased to 100% after attempts by the two coders to reconcile their lists. Each car model has also been designated as belonging to a narrowly defined car platform, containing several ''sibling'' models, and to a broadly defined platform, containing several *sets* of siblings (''cousins''). For example, according to *Ward's* (1977–1989), in 1977 GM began building large rear-wheel drive cars off of its (broadly defined) ''B/D'' platform. The narrowly defined ''B'' platform included several siblings—the Buick LeSabre, Chevrolet Caprice, Oldsmobile 88, and Pontiac. Their slightly larger cousins, built off of the ''D'' platform, were the Buick Electra, Cadillac, and Oldsmobile 98.

Outcome Variable

The repair rate data for this study come from the annual reader survey conducted by the nonprofit magazine, *Consumer Reports*, and include car models built from 1981 to 1989. In two separate sets of analyses, each designed to validate the other, this study takes these car models, built in the 1980s, and measures their repair record at two different milestones: during the third year of ownership and during the sixth year of ownership. This three- and six-year measurement lag is used to gauge the reliability

of cars manufactured in each year of a car model's production life; i.e., for cars from a car model's debut year in production, from its second year in production, from its third year in production, etc.

One potential concern here is that measuring quality in the sixth year of ownership, or even in the third year of ownership, is too far removed from what took place in the factory. However, the median age of cars in operation in the United States was six or seven years throughout the 1980s, and more than 90% of the cars built between 1981 and 1989 were still in operation six years later (*Ward's Automotive Yearbook* 1996). Thus, six years seems a reasonably relevant and meaningful time frame for measuring a car's reliability. Furthermore, some quality problems, like rust, may not even appear on a car right away.

Another potential pitfall in measuring quality after so many years of ownership is the possibility that repair rates improved during this period because owners neglect their cars today less than they used to, so that their cars no longer need to be repaired as much. While it may be difficult to completely rule out the possibility of an overall reduction in owner neglect, there is some circumstantial evidence to contradict this alternative explanation. First, the initial quality (90 days after purchase) of Big Three cars improved throughout the 1980s (Harbour and Associates 1990), so it seems probable that their long-term reliability would improve as well, apart from anything owners did. Second, a trend of less owner neglect might imply that, all else equal, owners would become less likely to ''junk'' their cars after six years, but this did not occur (*Ward's Automotive Yearbook* 1996). In any event, this study's use of statistical controls and dummy variables may help to some extent in easing this concern. It may also be the case that customers' reporting of reliability problems has been affected by increasing customer expectations for quality. Nevertheless, if anything, this possible bias merely strengthens most of the study's results, because it should be harder, not easier, to detect improvement in customer-perceived quality over time if customer expectations of quality are rising.

This study's repair ratings do, of course, have the same problems as other survey-based research, although with hundreds of thousands of respondents per year, sampling error is probably small. One potential limitation of this study is that the recollection of car owners about recent repairs may somehow be biased by the media or other outside influences in ways that would affect parameter estimates. Even if *Consumer Reports* survey respondents do not represent the general public, though, it seems reasonable to assume that any overall bias would affect the ratings of all North American car models about equally

(Eberts and Eberts 1995, Ginter et al. 1987). Moreover, the analyses for H1–H3 statistically control for repair rate differences *across* car models, including those due to owner neglect, by using a dummy variable in the regression analysis for each car model. The analysis for H4 and H5 also includes an effort to control for extraneous factors like owner neglect and differences in customer bases (see below).

This study also makes an effort to ensure year-to-year consistency in the data set. When the repair record of only one version (e.g., V6 or V8 engine) of a car model was published by *Consumer Reports*, then that rating has been used to represent the entire car model for that year. In those cases where the frequency-of-repair data in a given year were published for two or three versions of the same car model, this study uses a weighted average based on production volume. Any biases from aggregating by engine type should be lessened by the fact that this study has excluded all engine-related repairs (see below). For repair rates measured during the third year of ownership, this study's initial sample includes 317 observations of car models by year (1984–1989); these data come from the Big Three firms, their nine divisions, and 76 car models. Unfortunately, *Consumer Reports* only began publishing repair rates in a format usable for this study in its 1987 issue; as a result, the earliest repair rates measured during the third year of ownership only go back as far as cars manufactured in 1984. For cars in their sixth year of ownership, though, the data set can begin with cars made in 1981. Thus, for repair rates measured during the sixth year of ownership, this study's sample includes 376 observations of 78 car models by year (1981–1989).

Consumer Reports publishes frequency-of-repair data for more than a dozen "trouble spots" (e.g., fuel system, air conditioning) on a five-point scale, where each of the five rating categories corresponds to a percentage range. These ratings have been adjusted by *Consumer Reports* for mileage differences across models. For this study the frequency-of-repair ratings have been converted to the midpoint of the range indicated for each trouble spot. For the first four categories on the five-point scale, this method will lead to frequency-of-repair estimates that are somewhat imprecise, but not systematically biased in either direction (King et al. 1994). However, for the fifth category (those cars with the worst repair rates in that year), *Consumer Reports* only indicates that the repair rate was above a particular percentage. Using a rule of thumb suggested by a source at *Consumer Reports*, an upper bound (15 percentage points above the industry average) has been placed on these "Category 5" data, and the midpoint of the resulting range used. The issue of whether or not this procedure can potentially lead to

biased parameter estimates (i.e., a "ceiling effect") is addressed in more detail in the section on analysis techniques (see below).

Because this study will test the effects of time versus cumulative production output, and because most prior research has focused on cumulative output, a conservative test should use a measure of quality more closely related to cumulative output. Thus, I have selected reliability factors that relate as closely as possible to manufacturing in general and to assembly in particular (Krafcik 1988), which are the primary subject of cumulative production output-based learning curve theory. In the data from *Consumer Reports*, at least four of the various trouble spots fit this description:

(1) Body rust, e.g., pitting, corrosion, and perforation;
(2) Paint and exterior trim, e.g., fading, discoloring, chalking, peeling;
(3) Body integrity, e.g., air or water leaks, wind noise, rattles and squeaks;
(4) Body hardware, e.g., window, door, and seat mechanisms, locks, belts.

Statistical analysis allows us to confirm that these trouble spots tap an underlying construct at the car model level of analysis. For the sample of repair rates measured during the third year of ownership, a factor analysis of these trouble spots indicates the presence of only one underlying factor, onto which all four items have factor loadings above 0.4. These four items also form a statistically reliable scale, with a Cronbach's alpha of 0.70. In addition, this four-item scale, computed as the average (Harman 1976) of the four repair rates, is relatively stable over time; i.e., it has a similar factor structure and statistical reliability for every year of the sample. For the sample of repair rates measured during the sixth year of ownership, however, the factor analyses do not always yield a one-factor solution with all four trouble spots, but consistently do so when the data on body hardware are dropped. Because this study's theory does not require the inclusion of the body hardware trouble spot, and a more valid and reliable measure can be constructed by deleting it, a three-item measure is used here instead for cars in their sixth year of ownership. All factor loadings are well above 0.4, and the three-item scale has a reliability of 0.75. This three-item scale, computed as the average of the three repair rates, is also stable over time in terms of both its factor structure and statistical reliability. To ensure comparability across the two data sets, most of the regression analyses were also conducted on a three-item repair measure for cars in their third year of ownership (Cronbach's alpha = 0.66), with substantially similar results.

Predictor Variables

Data on production output and on market segments (sub-compact, compact, intermediate, full-size, and luxury) come from *Ward's Automotive Yearbook* (1975–1989). Because the repair rate data are reported by model year, which begins approximately every August, cumulative production volume is also computed based on model-year output. Platform designations are based on information reported in *Ward's* (1976–1989); in a handful of cases, this information was supplemented or confirmed by checking *Consumer Reports* (1976–1989) and popular-press articles in LEXIS-NEXIS (1983–1989). Industry-wide production is defined in this study as all United States and Canadian output produced by the Big Three.

The starting point used for division, company, and industry production output is 1975, a date which should be early enough to capture most cumulative production knowledge, since all of the car models in the sample did not even debut until 1976 or later. Furthermore, recent empirical work (Argote et al. 1990, Darr et al. 1995, Epple et al. 1991, Henderson and Cockburn 1996) has found that production knowledge tends to depreciate; i.e., more recent production experience is more useful than earlier experience. So, because all of the Big Three have been making cars for many decades, it seems reasonable to assume that any pre-1975 learning that is carried forward into the 1980s will be modest and, in any case, largely common across the companies.

Analysis Techniques

This study uses panel data. The regression equations are based on ordinary least squares (OLS) with fixed effects, i.e., with a dummy variable included for each car model. This approach assumes that each car model has its own baseline level of quality—i.e., its own intercept, or starting point, on the learning curve—but has the same rate of learning as all other car models. Accounting for fixed effects, especially because these data consist of many short time series, is assumed to eliminate the major sources of autocorrelation and heteroscedasticity.

The traditional shape of the learning curve is $y = ax^{-b}$, where y is the outcome measure (repair rate) of a car model produced in a given year, a is a constant, x is that car model's cumulative output up through the previous year, and b is the slope of the learning curve. By taking the natural logarithm of both sides of this equation, it can be rewritten as $ln(y) = ln(a) - b \, ln(x)$. Typically, additional predictor variables are included in this second equation, already in log form. While researchers have also used other specifications for the learning curve, this functional form, the power curve, remains the most common and robust (Kantor and Zangwill 1991). Note that when

one includes dummy variables corresponding to each car model, the predictor variables in the regression equation are only able to explain the variation in repair rates *within* a car model.

The passage of time is also included in this study's statistical model. A preliminary analysis, using a subset of this study's sample, determined that the year of production was a critical predictor of improvements in repair rates. The functional form of this relationship, however, is difficult to predict from theory. The current study takes the somewhat novel approach of using the least restrictive method for incorporating a time trend: dummy variables are used for each year that a car model has been in production, up through the 13th year (the highest). Another advantage to using dummy variables for the year of production is that it helps rule out the alternative explanation that the industry's reduction in repair rates is not a learning curve at all, but just a sign that car models with bad repair rates are taken off the market, and that surviving car models have better quality than ones that exit the market. There is no evidence that the Big Three discontinue car models early because of poor repair rates (Cole 1990). Nevertheless, the dummy variables make it possible to examine year-to-year changes in repair rates even early on in the production life of the average car model before any car models start exiting the market. Note that because the cumulative output measures are computed using a one-year lag, there is no output to predict a car model's first-year repair rate; the reference category for the time trend dummy variables in the initial regression equations (Tables 2 and 3) is therefore the second year of production.

One pitfall to the procedure (described earlier) of placing an upper limit on repair rate data from Category 5 (on the five-point scale) is that it may lead to a ''ceiling effect'' and thereby bias the regression results (Kennedy 1985). To the extent that this problem exists, it should be more pronounced for repair rates measured during the sixth year of ownership, when the least reliable cars may start to have repair rates that are not just bad, but truly awful—i.e., above the repair rate ''ceiling'' imposed on the data set. For example, in its sixth year of ownership, the 1989 Hyundai Sonata, although not included in this study's data set, was singled out by *Consumer Reports* (1995, p. 286) for having the worst repair rate for body hardware, 37%, well above that year's Category 5 range of 14.8% and above. Censoring models, such as TOBIT or PROBIT, are sometimes used to address this problem, but these statistical models are not designed for either fixed effects or the use of a multimeasure scale as the outcome variable. So instead, this study uses a procedure to test to see if including the Category 5 observations does

in fact create a ceiling effect, i.e., make the regression slopes flatter. To do so, a dummy variable (''ceiling'') is created that corresponds to a Category 5 rating for any of the three or four trouble spots used to construct the overall repair rate measure for a given car model in a given year. An observation is thus assigned a *zero* for the ''ceiling'' variable only if the car model in question was able to score below a five on *all* the relevant trouble spots in that year. In contrast, observations ''tainted'' with at least one trouble-spot rating of five (on the five-point scale) are assigned a *one* for the ''ceiling'' variable. Unfortunately for U.S. carmakers, this latter group of observations is sizable: For repair rates measured during the third year of ownership, the ''ceiling'' observations (i.e., *ceiling* = 1) make up 36% of the sample; for repair rates measured during the sixth year of ownership, they make up 42%. The ''ceiling'' dummy variable is then interacted with the relevant predictor variables. If the interaction effects are significant, then caution dictates limiting any generalizations of exactly how steep or flat the learning curve is only to certain car models; namely, those models *not* associated with the worst-possible repair rating in a given year.

Unlike the tests for learning curve parameters (H1–H3), the tests for differences in baseline repair rates by debut year (H4 and H5) *do* make comparisons across car models. In particular, to avoid any linear dependence among the predictor variables, the car model fixed effects have been replaced in Equation 11 with 13 control variables that correspond to the five market segments (subcompact, compact, intermediate, full-size, and luxury), the Big Three automakers (Chrysler, Ford, and GM), and the interaction terms among these variables (e.g., Ford*luxury). There are 13, not 15, controls because there is a reference category, and no Chrysler luxury cars are in the sample. If debut year, which ranges from 1976 to 1988, has a significantly negative regression coefficient in Equation 11, then H4 is supported. If it is indeed negative but significantly less so than the coefficient for year of production, then a ''disruptive learning'' perspective is supported; if it is significantly *more* negative than the coefficient for year of production, however, then an ''enhanced learning'' perspective (H5) is supported. Note that compared to the fixed effects model, however, these 13 control variables may not statistically control as well for all of the differences in repair rates across car models that might be due to extraneous factors. This omitted variable bias does not appear to be much of a problem because the coefficient for year of production seems fairly similar in Equation 9 and Equation 11. However, some people may still wish to consider this analysis suggestive.

Results

Table 1 shows the means, standard deviations, and simple correlations among the variables used in the regression equations in Table 2. Because each of the paper's four tables of regression results (Tables 2–5) has a slightly different sample size, this paper, in the interests of parsimony, shows only the correlations corresponding to Table 2. Although some of the variables are logarithmically transformed in the regression equations, they are shown in their untransformed (and more interpretable) state in Table 1. Note that the simple (zero-order) correlations among the variables shown in Table 1 may seem a bit confusing, even misleading, because they lack the statistical controls used in the multiple regression equations below.

H1: Learning Curve

The traditional learning curve equation, taking the natural log of both the left-hand and right-hand variables, is shown for repair rates measured during the third year of ownership (Table 2) and for repair rates measured during the sixth year of ownership (Table 3). There appears to be support for H1 in terms of cumulative production experience. On average, when a manufacturer has previously produced a lot of cars of a given model, that car model's repair rate is lower, whether that repair rate is measured during the third or sixth year of ownership (Equation 2). So it would seem that increased experience leads to improved product quality, yet (because this equation is in log-log form) in decreasing increments. In other words, as predicted, there is a learning curve for car model repair rates.

H2 and H2-ALT: Learning over *Time*

This cumulative output-based effect disappears, however, when the year-of-production dummy variables are taken into account. Although cumulative production experience is highly correlated with the passage of time, the regression analysis is able to distinguish between the two. Once we control for the fact that the average car model's repair rate generally improves from year to year during its production life, the extent of a manufacturer's production experience for a particular car model appears to make no difference (Equation 3).

Additional analyses (not shown here) were conducted to see if incorporating the idea that production knowledge can depreciate might make the production output variable a significant independent predictor of repair rates after all. This alternative explanation was tested by replacing cumulative production output with variables corresponding

Table 1   Means, Standard Deviations, and Simple Correlations[a]

	Mean	Std. Dev.	1	2	3	4	5	6	7
1. Repair rate[b]	0.077	0.025							
2. Years in Production	5.84	2.78	0.17**						
3. Car model output[c]	750	818	0.24**	0.72**					
4. Sibling output[c]	2,515	2,859	0.28**	0.67**	0.71**				
5. Cousin output[c]	4,122	3,460	0.13*	0.62**	0.61**	0.80**			
6. Division output[c]	10,700	5,940	0.17**	0.18**	0.48**	0.20**	0.07		
7. Firm output[c]	39,371	19,352	0.19*	0.10	0.24**	0.35**	0.15**	0.31**	
8. Big Three output[c]	92,394	13,146	−0.25**	0.24**	0.14*	0.01	0.02	0.25**	0.31**

[a]N = 317.
[b]Repair rate is measured during the *third* year of ownership.
[c]Computed as cumulative production output (lagged by one year). Shown in 000s.
*$p < 0.05$
**$p < 0.01$

Table 2   Regression Results for Cumulative Production Output and Year of Production[a]

Outcome Variable = ln (Repair rate, measured during the *third* year of ownership)

	Equation 1		Equation 2		Equation 3		Equation 4		Equation 5	
76 fixed effects	signif.***		signif.***		signif.***		signif.***		signif.***	
Car model output[b]			−0.128***	(0.022)	0.003	(0.065)	−0.037	(0.092)	0.015	(0.099)
3rd Year of Production					−0.201**	(0.063)	−0.216**	(0.067)	−0.156	(0.085)
4th Year of Production					−0.222*	(0.089)	−0.248*	(0.098)	−0.114	(0.148)
5th Year of Production					−0.269*	(0.107)	−0.302*	(0.119)	−0.088	(0.212)
6th Year of Production					−0.305*	(0.124)	−0.342*	(0.138)	−0.048	(0.273)
7th Year of Production					−0.283*	(0.140)	−0.323*	(0.154)	0.042	(0.336)
8th Year of Production					−0.340*	(0.152)	−0.383*	(0.167)	0.051	(0.394)
9th Year of Production					−0.311	(0.164)	−0.355*	(0.179)	0.175	(0.468)
10th Year of Production					−0.294	(0.172)	−0.340	(0.187)	0.273	(0.532)
11th Year of Production					−0.326	(0.177)	−0.376	(0.194)	0.320	(0.591)
12th Year of Production					−0.550**	(0.185)	−0.601**	(0.202)	0.175	(0.652)
13th Year of Production					−0.487*	(0.203)	−0.536*	(0.218)	0.283	(0.702)
Sibling output[b]							0.064	(0.101)	−0.004	(0.118)
Cousin output[b]									0.046	(0.076)
Division output[b]									0.532	(0.802)
Firm output[b]									−1.434	(1.62)
Big Three output[b]									−0.037	(1.64)
$R^2 =$	0.720		0.755		0.779		0.780		0.783	
adjusted-$R^2 =$	0.633		0.677		0.695		0.695		0.694	

[a]N = 317. Standard errors are in parentheses.
[b]Computed as the log of cumulative production output (lagged by one year).
*$p < 0.05$
**$p < 0.01$
***$p < 0.001$

Table 3   Regression Results for Cumulative Production Output and Year of Production[a]

Outcome Variable = ln (Repair rate, measured during the *sixth* year of ownership)

	Equation 1		Equation 2		Equation 3		Equation 4		Equation 5	
78 fixed effects	signif.***		signif.***		signif.***		signif.***		signif.***	
Car model output[b]			−0.093***	(0.020)	0.114	(0.066)	0.046	(0.088)	0.067	(0.092)
3rd Year of Production					−0.122	(0.066)	−0.145*	(0.069)	−0.160	(0.099)
4th Year of Production					−0.244**	(0.093)	−0.283**	(0.098)	−0.304	(0.176)
5th Year of Production					−0.287*	(0.112)	−0.336**	(0.119)	−0.361	(0.249)
6th Year of Production					−0.409**	(0.127)	−0.463***	(0.135)	−0.494	(0.322)
7th Year of Production					−0.476***	(0.141)	−0.537***	(0.150)	−0.569	(0.396)
8th Year of Production					−0.477**	(0.150)	−0.541***	(0.160)	−0.571	(0.464)
9th Year of Production					−0.526**	(0.160)	−0.592***	(0.169)	−0.633	(0.545)
10th Year of Production					−0.501**	(0.173)	−0.568**	(0.182)	−0.611	(0.615)
11th Year of Production					−0.460*	(0.180)	−0.535**	(0.192)	−0.598	(0.676)
12th Year of Production					−0.495*	(0.192)	−0.570**	(0.203)	−0.634	(0.740)
13th Year of Production					−0.772***	(0.197)	−0.843***	(0.206)	−0.861	(0.798)
Sibling output[b]							0.107	(0.091)	0.069	(0.100)
Cousin output[b]									0.053	(0.071)
Division output[b]									−0.456	(0.524)
Firm output[b]									2.126	(1.54)
Big Three output[b]									−1.680	(1.67)
R^2 =	0.773		0.788		0.805		0.806		0.808	
adjusted-R^2 =	0.714		0.732		0.744		0.744		0.743	

[a]N = 376. Standard errors are in parentheses.
[b]Computed as the log of cumulative production output (lagged by one year).
*$p < 0.05$
**$p < 0.01$
***$p < 0.001$

to a given car model's (non-cumulative) output from the previous year, from the year before that, from two years before that, etc. At a minimum, to estimate a knowledge depreciation parameter, the coefficients of these variables should be statistically significant and decreasing in magnitude the farther back one goes (Darr et al. 1995); in fact, they were neither. Thus, there is no evidence of production knowledge depreciation in this study's data. Taking knowledge depreciation into account, then, does not change the central finding from Equation 3—namely, that the *year* in a car model's production life, not cumulative production experience up to that point, best predicts a car model's ultimate repair rate (confirming H2–ALT).

So far, because the cumulative experience variable is computed using a one-year lag, this study has excluded repair rate data based on a car model's first year in production. Because measures of lagged production output are not statistically significant, however, it now becomes possible to expand this study's sample size to include these data. Doing so adds 39 more observations and seven more car models for repair rates measured during the third

year of ownership (Table 4); for repair rates measured during the sixth year of ownership, 67 more observations and ten more car models can be added (Table 5). As before, the year of production is examined using the least restrictive approach of having dummy variables which correspond to the number of years since a car model began its production life (Equation 7). In this case, the reference category for the dummy variables is the first year of a car model's production. The coefficients for Equation 7 in both Tables 4 and 5 seem to indicate a gradual reduction in repair rates for each subsequent year of a car model's production life. When we compare the R^2 figures for Equation 6 versus Equation 7, we see that this learning curve explains an additional 10% of the total variance in repair rates and about one-third of the within-car model variance in Table 4. In Table 5, it explains an additional 4% of the total variance and one-sixth of the within-car model variance.

To test for a possible "ceiling effect" for these data, each of the time trend dummy variables is interacted (in Equation 8) with the "ceiling" dummy variable, de-

Table 4   **Regression Results for Year of Production and Debut Year**[a]

Outcome Variable = ln (Repair rate, measured during the *third* year of ownership)

	Equation 6	Equation 7	Equation 8	Equation 9	Equation 10	Equation 11
83 fixed effects	signif.***	signif.***	signif.***	signif.***	signif.***	
2nd Year of Production		−0.112** (0.040)	−0.054 (0.046)			
3rd Year of Production		−0.318*** (0.042)	−0.142** (0.045)			
4th Year of Production		−0.337*** (0.043)	−0.178*** (0.046)			
5th Year of Production		−0.386*** (0.046)	−0.217*** (0.050)			
6th Year of Production		−0.416*** (0.049)	−0.284*** (0.052)			
7th Year of Production		−0.395*** (0.056)	−0.248*** (0.058)			
8th Year of Production		−0.452*** (0.061)	−0.272*** (0.061)			
9th Year of Production		−0.422*** (0.071)	−0.311*** (0.075)			
10th Year of Production		−0.405*** (0.074)	−0.399*** (0.084)			
11th Year of Production		−0.437*** (0.078)	−0.348*** (0.085)			
12th Year of Production		−0.661*** (0.090)	−0.433*** (0.079)			
13th Year of Production		−0.598*** (0.123)	−0.487*** (0.116)			
"Ceiling" effect			0.378*** (0.055)	0.430*** (0.040)	0.420*** (0.041)	0.496*** (0.043)
Ceiling * Yr. of Production dummy variables			n.s.	not reported	not reported	not reported
Year of Production				−0.038*** (0.006)	−0.037*** (0.008)	−0.030*** (0.006)
Ford * Yr. of Production					0.007 (0.012)	
Chrysler * Yr. of Production					−0.020 (0.013)	
Debut Year						−0.043*** (0.006)
13 control variables for all company by market segment combinations						signif.***
R^2 =	0.688	0.786	0.881	0.877	0.878	0.744
adjusted-R^2 =	0.595	0.709	0.830	0.831	0.832	0.722

[a] N = 356. Standard errors are in parentheses.
*$p < 0.05$
**$p < 0.01$
***$p < 0.001$
n.s. = not significant

scribed above. For repair rates measured during the third year of ownership (Table 4), the overall interaction is not significant. Nevertheless, the ceiling effect is statistically significant for one or two individual years (not shown), and omitting the ceiling effect variables seems to distort several of the point estimates; this analysis therefore retains these interaction variables in subsequent analyses. For repair rates measured during the sixth year of ownership (Table 5), the overall interaction effect is statistically significant (p = 0.022), so these interaction variables are also included in subsequent analyses. The "ceiling"-related data points are not dropped altogether, however, because they help identify the coefficients of the fixed effects dummy variables. In sum, although there is some evidence of a ceiling effect, it is important to note that, with or without these potentially biased observations, Equations 7 and 8 both indicate a similar downward

trend in repair rates over time. That is, this problem only really affects what our estimate of the learning curve's slope should be, and thus does not alter this study's central findings on evidence of learning.

The decline in repair rates for a car model during its production life is plotted for a so-called "no ceiling effect" line, i.e., for car models in a given year that did not receive the worst-possible repair rating of five (on a five-point scale) for any relevant trouble spot. These plots are based on the coefficients from Equation 8 for repair rates as measured during the third year of ownership (Figure 2) and during the sixth year of ownership (Figure 3). In both graphs, the lines suggest a relatively straight trend. (The fluctuations at the far right of Figure 3 are probably due to a small sample size for car models after the tenth year of production. Note the large corresponding standard errors in Equation 8 of Table 5). Because the repair rate

Table 5   Regression Results for Year of Production and Debut Year[a]

Outcome Variable = ln (Repair rate, measured during the *sixth* year of ownership)

	Equation 6	Equation 7	Equation 8	Equation 9	Equation 10	Equation 11
88 fixed effects	signif.***	signif.***	signif.***	signif.***	signif.***	
2nd Year of Production		−0.092* (0.040)	−0.084* (0.042)			
3rd Year of Production		−0.128** (0.040)	−0.132** (0.041)			
4th Year of Production		−0.200*** (0.040)	−0.272*** (0.043)			
5th Year of Production		−0.199*** (0.042)	−0.293*** (0.048)			
6th Year of Production		−0.289*** (0.047)	−0.366*** (0.052)			
7th Year of Production		−0.335*** (0.053)	−0.475*** (0.061)			
8th Year of Production		−0.320*** (0.056)	−0.410*** (0.067)			
9th Year of Production		−0.362*** (0.066)	−0.476*** (0.080)			
10th Year of Production		−0.320*** (0.075)	−0.389*** (0.102)			
11th Year of Production		−0.274** (0.088)	−0.722*** (0.137)			
12th Year of Production		−0.305** (0.107)	−0.497** (0.190)			
13th Year of Production		−0.586*** (0.119)	−0.754*** (0.116)			
''Ceiling'' effect			0.188** (0.058)	0.222*** (0.051)	0.218*** (0.051)	0.413*** (0.051)
Ceiling * Yr. of Production dummy variables			signif.*	not reported	not reported	not reported
Year of Production				−0.062*** (0.006)	−0.058*** (0.008)	−0.057*** (0.007)
Ford * Yr. of Production					−0.001 (0.011)	
Chrysler * Yr. of Production					−0.026 (0.014)	
Debut Year						−0.070*** (0.005)
13 control variables for all company by market segment combinations						signif.***
R^2 =	0.763	0.802	0.867	0.860	0.862	0.758
adjusted-R^2 =	0.705	0.745	0.821	0.819	0.819	0.742

[a]N = 443. Standard errors are in parentheses.
*p < 0.05
**p < 0.01
***p < 0.001

measure is in log form, a straight line essentially means that during a car model's production life quality improved, but in decreasing amounts (confirming H1 and H2–ALT). Because Equation 8 only includes dummy variables, the corresponding graphs are not in log-log form but in semilog form. So, technically speaking, Figures 2 and 3 show an exponential curve, not a power curve. To test the linearity of the trend line shown in Figures 2 and 3 formally, Equation 8 can be compared to Equation 9. For repair rates as measured during the third year of ownership (Table 4), $F_{11,248} < 1$; for repair rates as measured during the sixth year of ownership (Table 5), $F_{11,330} = 1.490$, $p = 0.133$. Thus, we cannot reject the presumption that the lines in Figures 2 and 3 are essentially linear.

The slopes (shown in Equation 9) of these lines represent the annual rate of learning for the average car

Figure 2   Learning Curve (on Log Scale) for Car Model Repair Rate as Measured During the *Third* Year of Ownership[a]

[a]Plotted points correspond to the coefficients listed in Equation 8 of Table 4.

Figure 3 Learning Curve (on Log Scale) for Car Model Repair Rate, as Measured During the *Sixth* Year of Ownership[a]

[a]Plotted points correspond to the coefficients listed in Equation 8 of Table 5.

model that did not receive the worst-possible repair rating in a given year (i.e., for the ''no ceiling effect'' observations). Specifically, when repair rates are measured during the third year of ownership (Table 4), every additional year that a car model is produced results in a 3.8% drop in the repair rate measure. When repair rates are measured during the sixth year of ownership (Table 5), every additional year that a car model is produced results in a 6.2% drop.

Thus far, we have assumed that these learning rates are an industrywide phenomenon. To test this assumption, the year of production variable (for the ''no ceiling effect'' observations) is interacted with two dummy variables (Ford and Chrysler, with GM as the reference category) representing the Big Three automakers. The results, shown in Equation 10 of Tables 4 and 5, indicate that these learning *rates* are essentially the same for all three firms (for Table 4, $F_{2,257} = 1.740$, p = .178; for Table 5, $F_{2,339} = 1.818$, p = .164).

H3: Transfer of Production-Based Knowledge

When lagged cumulative production experience for a car model's platform, division, firm, or industry are added sequentially as predictors in Equations 4 and 5 of Tables 2 and 3, and in various other combinations not shown here (including each variable on its own), none of the individual variables have statistically significant coefficients. As can often occur, however, the very high correlations among some of the cumulative output measures, as well as the year of production, introduce a great deal of multicollinearity into the data. The result can be erratic parameter and standard error estimates, as seen in

Equation 5, where every single predictor variable goes ''haywire'' and loses its statistical significance; the multicollinearity is simply too great. Nevertheless, it turns out that the five ''knowledge transfer'' variables in Equation 5 do not even have a *combined* statistically significant effect on repair rates over and above that of the predictor variables in Equation 3: for Table 2, $F_{5,224} < 1$; for Table 3, $F_{5,281} < 1$.

These data thus provide no evidence for the transfer of *production*-based knowledge on product quality, as measured by repair rates (H3). Indeed, taken with the support for H2–ALT, this study finds no benefit to a car model's repair rate from *any* measure of cumulative production experience, not for the car model itself, its siblings, its cousins, its division, or its firm. Furthermore, additional analyses (not shown here) designed to account for a possible ceiling effect for the production output measures (by interacting these measures, in numerous combinations, with the ''ceiling'' variable) yield similar results.

H4: ''Debut-Year'' Learning

As shown in Equation 11 of Tables 4 and 5, even after controlling for a combination of all companies by market segment, the later a car model begins its production life, the lower its baseline repair rate is, thus confirming H4.

H5: Debut-Year Learning Versus the Learning Curve

According to Equation 11 in Table 4, the average car model debuts with a baseline repair record (as measured in the third year of ownership) that is 4.3% better than the previous year's debuting car model. In contrast, that previous year's debuting car model improved incrementally only an average of 3.0% during the one intervening year. In other words, an extra year of ''debut-year'' learning (as embedded in a new car model debut) leads to better repair rates than does an extra year of the incremental learning curve during a car model's production life. Thus, on balance, a car model's debut is not a disruptive or even neutral event, but instead gives rise to ''enhanced'' learning (confirming H5). As shown in Figure 4, the incremental improvements during the learning curve (3.0% per year on average) do add up considerably over the years, but the larger improvements (4.3% per year) made to debuting car models mean that the gap will never be closed. These overall results are statistically significant for repair rates measured during either the third year of ownership ($F_{1,327} = 8.626$, p = 0.004) or the sixth year of ownership ($F_{1,414} = 5.232$, p = 0.023). There is no evidence of a ceiling effect for debut year in Equation 11 in Table 4, although there is some evidence

Figure 4 **"Debut-Year" Learning Versus the Learning Curve**[a]
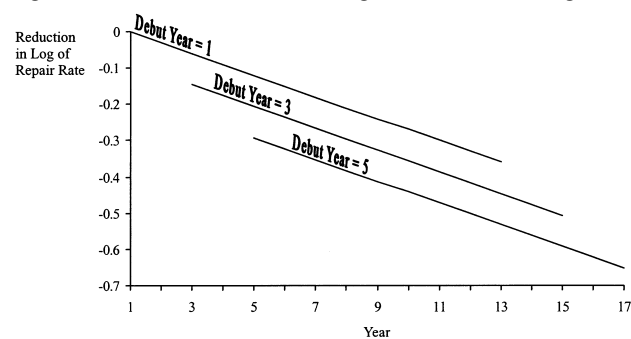
[a]For car model repair rate, as measured during the *third* year of ownership. Plotted points correspond to the coefficients listed in Equation 11 of Table 4.

for it in Table 5. Nevertheless, correcting for this possible ceiling effect yields similar results.

Discussion

The findings from this study suggest a number of specific ways in which we may need to rethink our understanding of organizational learning and of learning curves. These new insights, which both expand and refine prior research, deal primarily with the *what* and *when* of organizational learning curves. In particular, the evidence suggests three main conclusions: (1) that stable learning curves are not limited to the cost or efficiency domain; (2) that, contrary to prior research in the efficiency domain, some learning curves appear to be more a function of time than a function of cumulative experience; and (3) that improvements to the starting point of some learning curves, when a product is first introduced, are even more important than improvements made during subsequent production.

This study presents systematic empirical evidence that, as predicted by H1, there is indeed a learning curve for quality. However, researchers and others may need to be cautious in generalizing efficiency-based findings to other aspects of performance like quality, because the sources and forms of organizational learning, especially the *when* of learning, can vary depending on the outcome measure chosen. Learning curve researchers may be somewhat surprised by this study's finding that, as predicted by H2–ALT, the main source of quality improvement appears to be the passage of time, and not cumulative production experience. After all, most studies of productivity-related learning have found cumulative experience to be more,

not less, important than the passage of time (Argote 1993, Argote and Epple 1990, Lieberman 1987). However, to those familiar with the "quality movement," this finding is consistent with the notion that quality improvement occurs as a result of broad organizational practices that are not directly tied to production experience (Benson et al. 1991, Shortell et al. 1995). Moreover, not only are "off-line" quality-improvement efforts likely to be a function of time (Schneiderman 1988, Sterman et al. 1997), but so too is the design cycle in this industry. The findings here thus also suggest that learning in the area of product reliability may be heavily based on the design group and its annual window of opportunity (Tyre and Orlikowski 1994) to update and improve car models at the end of every summer.

If annual product updates provide organizations with a small window of opportunity to keep moving along the learning curve for quality, then major model changes (or debuts) provide an even larger window of opportunity to *begin* the learning curve at an improved baseline level by incorporating new knowledge that improves quality. Indeed, as predicted by H4, this study finds that more recently introduced car models have better baseline repair records than do models that debuted earlier. Moreover, as predicted by H5, the learning embedded in later-debuting car models, rather than having a disruptive effect, on balance actually *enhances* organizational learning. That is, the learning embedded in later-debuting car models (what we might call "exceptional" learning) leads to better repair rates at any given point in time than the more incremental organizational learning that typically occurs during a car model's production life ("everyday" learning). For consumers, this finding suggests that, all else being equal, and contrary to conventional wisdom, the latest-debuting model is likely to be a more reliable product. Wait a year and that debuting car model will likely be better in quality (H1), but not as good on average as that next year's debuting model (see Figure 4).

Besides these three main findings—that there is a learning curve for quality; that it is a function of time; and that, above all, improvements to the learning curve's starting point are particularly critical—we must now consider the issue of knowledge transfer. The lack of support for H3 suggests that know-how brought in from outsiders does not accumulate as a function of their production experience, but the favorable results for H2–ALT are consistent with the idea that this outside knowledge accumulates with the passage of time. Thus manufacturers probably *do* share quality-related knowledge across product families, divisions, and firms. For example, the invention of two-sided galvanized steel has helped reduce rust problems in cars (Berry 1986). What this study's data

cannot say, however, is to what extent quality improvements made with the passage of time, whether for a product's debut (H4) or subsequent learning curve (H1), are due to inventions by that product line's design or research engineers; to knowledge transferred from others in the same company (Argote 1993); to knowledge "diffused" across the industry (Lieberman 1987); or to technological breakthroughs occurring in related industries. Indeed, if we think back to the earlier theoretical question of *who* does the learning—employees, organizational systems, and outside actors (Argote 1993)—then any or all of these may be contributing to the reduction in repair rates over time. That is, it was not possible to trace this kind of knowledge transfer back statistically to any specific source, as H3 attempted to do. Nevertheless, some "researchers use time as a proxy variable for [knowledge transfer], assuming that the acquisition of knowledge in the general environment would be associated with the passage of time" (Argote 1993, p. 41). As a result, we might plausibly interpret the support for H2–ALT (learning as a function of year in production) and for H4 (learning as a function of debut year) as evidence of a kind of generalized knowledge transfer, a technological "march of progress." Thus, in a sense, we do find evidence of knowledge transfer after all.

Conclusion

In sum, the process of learning, in an industrial setting, is more complex than previously thought. First, a learning curve exists not only for efficiency but also for quality. Second, organizational learning in the area of product reliability depends on the passage of time rather than on cumulative production experience. This result is the reverse of what researchers have found for efficiency learning curves. Third, while it is common wisdom not to buy a new car in its first year of production, the opposite advice actually seems to apply. That is, more learning takes place as a result of new product introduction, before the learning curve even begins, than in any year of production. These findings expand the traditional concept of learning, especially our understanding of *what* gets learned *when*.

It is worth noting some of the limitations of this study. For example, it may be the case that automobile quality, in general or as measured in this study, is a unique situation. Further study is needed to determine if the time-dependent learning curve proposed here can be generalized to other types of organizations, since this study is the first large-scale, industrywide investigation of organizational learning and learning curves in the area of re-

pair rates. In addition, it is possible that a learning curve based on production experience does exist for quality but cannot be detected using only annual data because the curve simply flattens out in the early months after a car model's debut or redesign.

Future research might focus usefully on the *why* of learning curves. The motivation to learn in the quality domain was strong among North American automakers during this study's time frame due to competitive pressure, but motivation is not a constant. When we focus on variables like cost, we can safely assume that the people in the organization want to learn, especially in for-profit firms. With product reliability, however, as in some other domains like research and development (Levin 1999), the motivation to learn and improve is not always as strong. This same point also applies to a performance indicator like customer satisfaction, which organizations do not always have a strong economic or strategic incentive to improve. From a conceptual perspective, then, organizational learning theory might benefit from incorporating the insights of "desire-and-ability" models of organizational change (Levin and Shortell 1996, Zajac and Kraatz 1993).

Another fruitful area for future research would be the *how* of organizational learning curves. By using a more qualitative research methodology of in-depth field studies, this alternative approach would complement the quantitative approach used here to uncover this area's underlying mechanisms, patterns, and routines. Such studies of organizational processes (Levin 1999; Pentland 1992, 1995) enable us to probe more deeply inside the "black box" of organizational learning. Along similar lines, although the current study used public data pragmatically, it was not possible to detect all of the subtleties (e.g., manufacturing process changes) underlying learning rates within the industry. By gaining access to one or more companies, however, future studies could explore such subtleties in more depth.

Endnote

[1]All car reliability data used in this article are copyrighted ©1976–1995 by Consumers Union of U.S., Inc., Yonkers, NY 10703-1057, a

nonprofit organization. Adapted with permission from issues of *Consumer Reports* for educational purposes only. No commercial use or photocopying permitted. To subscribe, call 1-800-234-1645 or visit us at <www.ConsumerReports.org>.

References

Abernathy, W. J., J. M. Utterback. 1978. Patterns of industrial innovation. *Tech. Rev.* **80**(7) 40–47.

Adler, P. S. 1990. Shared learning. *Management Sci.* **36** 938–957.

——, K. B. Clark. 1991. Behind the learning curve: A sketch of the learning process. *Management Sci.* **37** 267–281.

Argote, L. 1993. Group and organizational learning curves: Individual, system and environmental components. *British J. Soc. Psych.* **32** 31–51.

——, S. L. Beckman, D. Epple. 1990. The persistence and transfer of learning in industrial settings. *Management Sci.* **36** 140–154.

——, D. Epple. 1990. Learning curves in manufacturing. *Science* **247** 920–924.

Benson, P. G., J. V. Saraph, R. G. Schroeder. 1991. The effects of organizational context on quality management: An empirical investigation. *Management Sci.* **37** 1107–1124.

Berry, B. H. 1986. Detroit tries out the new electrogalvanized steel. *Iron Age* **229** (July 4) 44+.

Cole, R. 1990. U.S. quality improvement in the auto industry: Close but no cigar. *California Management Rev.* **32**(4), 71–85.

Consumer Reports. 1976–1995. Consumers Union, Yonkers, NY.

Darr, E. D., L. Argote, D. Epple. 1995. The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises. *Management Sci.* **41** 1750–1762.

Eberts, R., C. Eberts. 1995. *The Myths of Japanese Quality*. Prentice Hall, Upper Saddle River, NJ.

Epple, D., L. Argote, R. Devadas. 1991. Organizational learning curves: A method for investigating intra-plant transfer of knowledge acquired through learning by doing. *Organ. Sci.* **2** 58–70.

Garvin, D. A. 1988. *Managing Quality: The Strategic and Competitive Edge*. Free Press, New York.

Ginter, J. L., M. A. Young, P. R. Dickson. 1987. A market efficiency study of used car reliability and prices. *J. Consumer Affairs* **21** 258–276.

Harbour and Associates. 1990. *The Harbour Report: A Decade Later: Competitive Assessment of the North American Automotive Industry, 1979–1989*. Harbour & Associates, Berkley, MI.

Harman, H. H. 1976. *Modern Factor Analysis*. University of Chicago Press, Chicago, IL.

Hayes, R. H., K. B. Clark. 1986. Why some factories are more productive than others. *Harvard Bus. Rev.* **64**(5) 66–73.

Henderson, R., I. Cockburn. 1996. Scale, scope and spillovers: The determinants of research productivity in drug discovery. *Rand J. Econom.* **27** 32–59.

Hirsch, P. M., D. Z. Levin. 1999. Umbrella constructs versus validity police: A life cycle model. *Organ. Sci.* **10** 199–212.

Huber, G. P. 1991. Organizational learning: The contributing processes and the literatures. *Organ. Sci.* **2** 88–115.

Ingrassia, P., J. B. White. 1994. *Comeback: The Fall and Rise of the American Automobile Industry*. Simon & Schuster, New York.

Irwin, D. A., P. J. Klenow. 1993. Learning-by-doing spillovers in the semiconductor industry. *J. Polit. Economy* **102** 1200–1226.

Kantor, P. B., W. I. Zangwill. 1991. Theoretical foundation for a learning rate budget. *Management Sci.* **37** 315–330.

Kennedy, P. 1985. *A Guide to Econometrics* 2nd ed. MIT Press, Cambridge, MA.

King, G., R. O. Keohane, S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ.

Krafcik, J. F. 1988. Comparative analysis of performance indicators at world auto assembly plants. Master of Science thesis, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

Lant, T. K., S. J. Mezias. 1992. An organizational learning model of convergence and reorientation. *Organ. Sci.* **3** 47–71.

Leavitt, H. J. 1951. Some effects of certain communication patterns on group performance. *J. Abnormal and Soc. Psych.* **46** 38–50.

Lester, R. K., M. J. McCabe. 1993. The effect of industrial structure on learning by doing in nuclear power plant operation. *Rand J. Econom.* **24** 418–438.

Levin, D. Z. 1999. Transferring knowledge within the organization in the R&D arena. Ph.D. dissertation, Northwestern University, Evanston, IL.

——, S. M. Shortell. 1996. Desire to implement and ability to implement: The case of total quality management. Paper presented at the annual meeting of the Academy of Management, Cincinnati, OH.

Lieberman, M. 1987. The learning curve, diffusion, and competitive strategy. *Strategic Management J.* **8** 441–452.

Mazur, J. E., R. Hastie. 1978. Learning as accumulation: A reexamination of the learning curve. *Psych. Bull.* **85** 1256–1274.

Muth, J. F. 1986. Search theory and the manufacturing progress function. *Management Sci.* **32** 948–962.

Pentland, B. T. 1992. Organizing moves in software support hot lines. *Admin. Sci. Quart.* **37** 527–548.

——. 1995. Grammatical models of organizational processes. *Organ. Sci.* **6** 541–556.

Schneiderman, A. M. 1988. Setting quality goals. *Quality Progress* **21**(4) 51–57.

Shortell, S. M., D. Z. Levin, J. L. O'Brien, E. F. X. Hughes. 1995. Assessing the evidence on CQI: Is the glass half empty or half full? *Hospital & Health Services Admin.* **40**(1) 4–24.

Sterman, J. D., N. P. Repenning, F. Kofman. 1997. Unanticipated side effects of successful quality programs: Exploring a paradox of organizational improvement. *Management Sci.* **43** 503–521.

Tushman, M. L., E. Romanelli. 1985. Organizational evolution: A metamorphosis model of convergence and reorientation. L. L. Cummings, B. M. Staw, eds. *Research in Organizational Behavior*, vol. 7. JAI Press, Greenwich, CT, 171–222.

Tyre, M. J., W. J. Orlikowski. 1994. Windows of opportunity: Temporal patterns of technological adaptation in organizations. *Organ. Sci.* **5** 98–118.

Udayagiri, N. D., S. Balakrishnan. 1993. Learning curves and knowledge spillovers: The case of semiconductor memories. Working paper, Reginald H. Jones Center, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Ward's Automotive Yearbook. 1974–1989, 1996. Detroit, MI.

Womack, J. P., D. T. Jones, D. Roos. 1990. *The Machine That Changed the World: The Story of Lean Production*. HarperCollins, New York.

Wright, T. P. 1936. Factors affecting the cost of airplanes. *J. Aeronautical Sci.* **3** 122–128.

Zajac, E. J., M. S. Kraatz. 1993. A diametric forces model of strategic change: Assessing the antecedents and consequences of restructuring in the higher education industry. *Strategic Management J.* **14**(S) 83–102.

Zangwill, W. I., P. B. Kantor. 1998. Toward a theory of continuous improvement and the learning curve. *Management Sci.* **44** 910–920.