All About Learning Curves

by Evin Stump P.E.
Senior Consultant
Galorath Incorporated

Analysts who use or plan to use learning curves will find that this paper covers the subject in considerable detail. The two most popular models, unit and cumulative average, are covered exhaustively. The math needed is completely exposed, and many tips are given with regard to practice. Twenty-two learning curve related equations are developed in this paper, and there are 41 fully worked examples.

In addition to the basics, some of the more advanced topics covered include:

- Error analysis
- Aggregated learning curves
- Segmented learning curves
- Interruption of production
- Fitting learning curves to production data
- Tradeoff analysis with learning curves
- Local rate of learning
- Production rate effects

To get maximum benefit from this paper, you should be comfortable with exponents and logarithms, and with simple algebraic manipulations. It also helps, but is not necessary, that you know a bit about ordinary least squares regression analysis.

Two topics in this paper require some knowledge of differential and integral calculus for complete understanding, but understanding need not be complete to make use of the results. Readers whose math is limited can simply plug the equations into a spreadsheet to get the results they need.

This paper comprises thirteen sections and an appendix. The appendix collects into one place all of the 22 formulas developed in the paper. The sections and the page numbers where they begin are:

I. Introduction

This paper discusses learning curves and their application to real world problems. Learning curves have several alternate names, such as improvement curves, progress curves, startup functions, and efficiency curves, but your author prefers the original historical name "learning curves" to the more recent alternatives.

Learning curves are mathematical models used to estimate efficiencies gained when an activity is repeated. The "learning effect" was first noted in the 1920s in connection with aircraft production. It use was amplified by experience in connection with aircraft production in WW II. Initially, it was thought to be solely due to the learning of the workers as they repeated their tasks. Later, it was observed that other factors probably entered in, such as improved tools and working conditions, and various management initiatives. But to your author's mind, it does little or no damage to a proper understanding of the phenomenon to group all of these factors together under the general heading of "learning." Management can learn, too.

The underlying notion behind learning curves is that when people individually or collectively repeat an activity, there tends to be a gain in efficiency. Generally, this takes the form of a decrease in the time needed to do the activity. Because cost is generally related to time or labor hours consumed, learning curves are very important in industrial cost analysis. A key idea underlying the theory is that every time the production quantity doubles, we can expected a more or less fixed percentage decrease in the effort required to build a single unit (the Crawford theory), or in the average time required to build a group of units (the Wright theory). These decreases occur not in big jumps, but more or less smoothly as production continues.

Learning tends to be "lost" when there is a break in repetitions of the activity, or a change in the nature of the activity. These subjects are explored in this paper, because such breaks are a common occurrence.

The usual use of learning curves is to estimate the labor hours, and thereby much of the cost, of a manufactured good that is built in significant quantities. Other uses include planning of the manufacturing work force, and estimating costs of construction when structures are repeated, as in like houses or like spans of a bridge. Learning curves have been applied to creation of documents, boring of tunnels, drilling of wells, upgrades of previously manufactured products, and many other repetitive activities. Some manufacturing companies also apply "learning" to the purchase of raw materials and also to the purchase of manufactured components from other companies. The former is usually justified

on the basis that as work progresses, scrap rates will decrease, yield rates will improve, and purchasing practices will become more efficient. The latter is justified on the basis that other companies will probably experience learning also, so why shouldn't that be a part of the price negotiation?

Learning curves are not applicable where there is no opportunity for increased efficiency. For example, if a part is made by a fully automated machine at the rate of ten per minute, then there may be no way to improve efficiency beyond that point, at least in the short run. Of course in the long run, that machine might be replaced with one that has twice the rate, but whether or not that qualifies as learning is problematic. Learning curves are also not applicable when an activity is sporadic, such as engine overhauls that occur irregularly.

Learning effects are greatest in operations that have considerable "touch" labor, that is, when people in the production process frequently touch the product. Learning tends to be less in semi-automated processes, and it may be non-existent in fully automated processes. There may be some "learning" in certain sporadic processes, in the sense of a gain in efficiency, but few would argue that the methods discussed in this paper are suitable for modeling it.

Several learning curve models have been proposed, but only two are in widespread use. Those are the "unit" (U) model, due to Crawford, and the "cumulative average" (CA) model, due to Wright. The former focuses on the effect of learning unit by unit, while the latter considers the average effect of learning over a number of units. A question often asked is--which model is best? Your author believes that the answer to that question depends largely on your purpose. Here are some things to consider:

- The math for the U model is easier than the CA model math when you are comparing specific units of production
- The math for the U model is easier than the CA model math when you are using segmented learning curves
- The math for the CA model is easier than the U model math when you are cumulating over many units
- It is easier to fit a U curve to historical unit data
- It is easier to fit a CA curve to historical block data
- Both models can be made to have about the same accuracy with due care
- The choice of model is more often governed by company standard practice, habit, or possibly by customer expectation, than by other factors.

II. The Underlying Power Law Equation

As it happens, both the U and the CA models are based on a common mathematical form often called the "power law." This is the exponential equation[1]

$$y = ax^b \qquad \text{Eq(1)}$$

where a and b are constants. If b = 1, this is simply the equation of a straight line passing through the origin, with slope a. If b is positive, the equation will pass through the origin, but if it is negative, it cannot, because that would require a division by zero. In virtually all learning curve situations, for both the U and the CA models, b is negative. It typically is a small negative number with absolute value less than unity, such as −0.25.
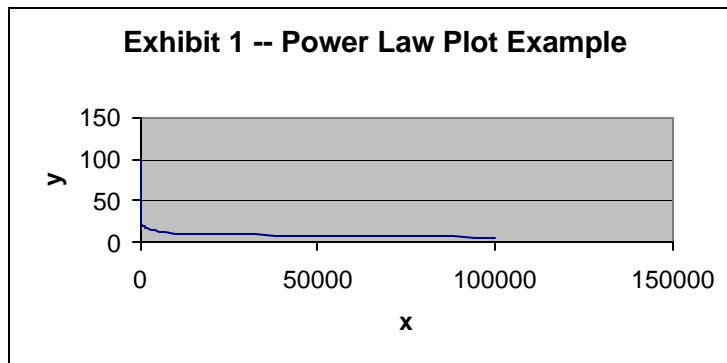
Exhibit 1 -- Power Law Plot Example

Exhibit 1 shows what the power law curve with negative b typically looks like when plotted in ordinary linear coordinates. This particular plot is of the power law $y = 100x^{-0.25}$.

Exhibit 2 -- Power Law Plot Example

---
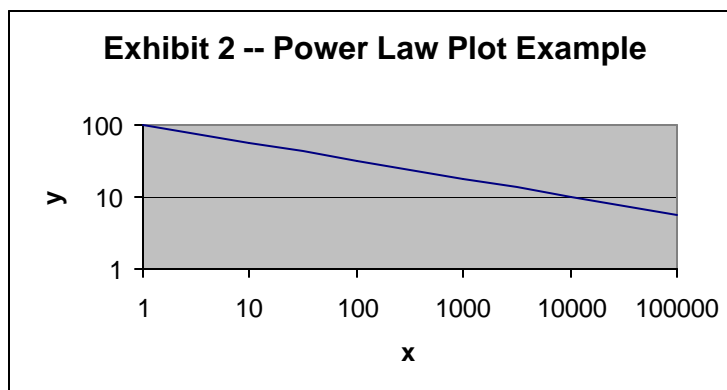
[1] This and all other equations developed in this paper are collected in the appendix.

Plots of learning curves in ordinary linear coordinates usually are not very satisfactory because details tend to get lost. For example, in Exhibit 1, it is not easy to see that for x = 1, y = 100. A usually better way to display learning curves is in a "log-log" plot, where both the x- and the y-axes have logarithmic coordinates.   Exhibit 2 shows the curve of Exhibit 1 plotted in log-log coordinates. In a log-log coordinate system, any power law relationship plots as a straight line.

In learning curve work, the exponent b is called the "natural slope" of the learning curve. Note in Exhibit 2 that because b is negative, the curve goes "downhill" to the right.

III. The U Learning Model

The U (unit) version of the learning curve can be written:

$$H_n = H_1 n^b \qquad \text{Eq(2)}$$

In this equation, H_n is the hours required for the nth unit of production, and H_1 is the hours required for the first unit. b is the "natural slope" of the learning curve, reflecting whether learning proceeds rapidly or slowly. If H_1 and b are known, as well as the unit of production of interest, from this information an estimate of the hours required for unit n can be computed.

Example 1. An estimator believes that the first unit of a product will require 100 labor hours. If, say, b = -0.2, how many hours will the tenth unit require?

Answer: Use the above equation:

$$H_n = H_1 n^b$$

Then,

$$H_{10} = 100(10)^{-0.2} = 63.1 \text{ hours}$$

▪▪▪

Example 2. If the natural slope of the unit curve is b = -0.2, and the 10th unit requires 30 hours, then how many hours will the 50th unit require?

Answer: To solve this problem, it is convenient to recognize a relationship between the hours for any unit, say the nth, and the hours for any other unit, say the mth. Using Eq(2), we can write the ratio:

$$H_n/H_m = (n^b)/(m^b)$$

Hence

$$H_n = H_m (n/m)^b \qquad \text{Eq(3)}$$

with H_1 conveniently canceling out when the ratio is taken. Therefore:

$$H_{50} = 30(50/10)^{-0.2} = 21.7 \text{ hours}$$

▪▪▪

Example 3. For any value of b, what is the relationship between the hours required for unit n, and the hours required for unit 2n (twice as many as n)?

Answer: Making use of Eq(3):

$$H_{2n} = H_n(2n/n)^b = H_n(2^b)$$

Whence:

$$\mathbf{H_{2n}/H_n = 2^b} \qquad\qquad \text{Eq(4)}$$

∙∙∙

Values for the parameter b have simply been provided in the examples so far. The reader may be wondering where these come from. What is the meaning of a "natural slope" such as –0.2 or –0.3? The answer is, it's hard to visualize, so what learning curve practitioners have done is come up with a system for measuring slope, i.e., rate of learning, that is more user friendly. Some would say it could be even friendlier than it is, but by now the system has such wide acceptance and use that to change it would take at least a court order.

The system in virtually universal use measures rate of learning on a scale of 0 to 100, in percentage. Slope expressed this way is called percent (or percentage) slope. You might expect that a 100% slope reflects a furious rate of learning, but quite the contrary. It represents no learning at all. Zero percentage, on the other hand, reflects a theoretically infinite rate of learning, if such a thing can be imagined. But in practice, human operations hardly ever achieve a rate of learning faster than 70% as measured on this scale. So the effective range of industrial learning is essentially 70-100%. Rarely, but it does happen, there is such a thing as "forgetting," and for forgetting percent slopes *exceed* 100%

One advantage of the percentage scale is that one can talk about learning in a way that makes some intuitive sense. For example, "On Project X the electronics assembly operation achieved 90% learning." The disadvantage is that to do actual learning curve calculations, one must first convert percent values to natural slope (b) values using a somewhat clumsy formula, which will be presented shortly.

We will use the letter S to represent percentage slope. Often it will be convenient to express S as a decimal, and the expression for that will be S/100.

Natural slope b is *defined* in terms of S by the formula:

$$\mathbf{b = log(S/100)/log(2)} \qquad\qquad \text{Eq(5)}$$

In this formula, any base of logarithms can be used. After a bit of algebraic manipulation, not shown here, we can arrive at the following equation for S in terms of b. In this equation, logarithms are to base 10.

$$S = 10^{b \log(2)+2}$$ Eq(6)

Example 4. If S = 100% (no learning), what value does b have?

Answer:

$$b = \log(100/100)/\log(2) = 0$$

Note that inserting $b = 0$ into $H_n = H_1 n^b$ results in $H_n = H_1$ for any value of n. This is to be expected if there is no learning.

▪▪

Example 5. If b = -0.2, what value does S have?

Answer:

$$S = 10^{-0.2 \log(2)+2} = 87\%$$

Note: Slopes in the vicinity of 87% (and even steeper) are common in the manufacture of airframes. In the manufacture of electronics, which is typically more complex, 90-95% is more common.

▪▪

While these relationships between S and b are somewhat clumsy, they do have an interesting and useful property. Suppose, for example, that it is somehow determined that a particular production process has a learning slope of 90%. Because the relationship between S and b has been defined as it has been, it is true (as will be shown in a moment) that if the first unit requires 100 hours, then the second will require 90 hours, the fourth will require 90% of 90, or 81, the eighth will require 90% of 81, or 72.9, and so forth. In other words, every time the production quantity is doubled, the hours are reduced to 90% of what they were before doubling. This is true not just for a 90% slope, but also for any percent slope. For example, an 80% slope results in reductions of 20% with every doubling. Moreover, the doubling does not have to begin with the first unit. The phenomenon holds if you start doubling at say, unit 47.

Generalizing, we can write as an hypothesis:

$$H_{2n} = H_n(S/100)$$

or:

$$\mathbf{H_{2n}/H_n = S/100} \hspace{4cm} \text{Eq(7)}$$

To show why this is true requires a bit of math. We present it now.

Recall that in Example 3 we showed that:

$$H_{2n}/H_n = 2^b$$

We have just hypothesized that:

$$H_{2n}/H_n = S/100$$

If these are both true, then:

$$S/100 = 2^b$$

Taking logarithms of both sides of this equation results in:

$$\log(S/100) = b\log(2)$$

Solving for b:

$$b = \log(S/100)/\log(2)$$

But this is precisely the definition of b in terms of S we gave previously in Eq(5).

Example 6. If the learning slope of a certain manufacturing process is 80%, and the first unit requires 1,000 hours, what will the 128th unit require?

Answer: The number 128 results when the number 1 is doubled to 2, 2 is doubled to 4, 4 is doubled to 8, 8 to 16, 16 to 32, 32 to 64, and finally 64 to 128. This is seven doublings. At each doubling, the 1,000 hours for unit 1 is multiplied by 0.8.

$$H_{128} = (0.8)^7(1000) = 209.7 \text{ hours}$$

We can of course obtain the same result using:

$$H_n = H_1 n^b$$

We first find b:

$$b = \log(0.8)/\log(2) = -0.3219$$

Next:

$$H_{128} = 1000(128)^{-0.3219} = 209.7 \text{ hours}$$

∎∎∎

To apply the unit learning curve, it is clearly necessary first to obtain an estimate for one unit in the production sequence. It would seem rational to try to estimate the first unit, and then use that result and an assumed value for S to estimate all of the rest. Sometimes that is done, but experienced estimators often feel uncomfortable trying to estimate unit 1, because experience shows that unit 1's hours are not always typical of the trend that follows. The reason is that unit 1 can be subject to production startup problems that don't get solved until a few units later in the production sequence.

Estimators will frequently estimate what they think a later unit will cost in hours, such as the 10th, or even the 100th, after potential problems have settled down. Then, they will use learning curve theory to calculate "theoretical" first unit hours as a basis for all subsequent calculations. The theoretical hours for unit 1 are often called "T_1." This is identically the same as what we have been calling H_1. We will not use the T_1 notation in this paper, because we prefer to reserve the letter T for total cost of a number of produced items. Another reason for not using this notation is that it (in the author's opinion) contributes to an unfortunate tendency to ignore the distinction between the unit and cumulative average models. This distinction is often glossed over in conversations, and even in formal documentation, to the confusion of all concerned. The theoretical first units in the two models almost never have identical values, when fitted from historical data. When you hear someone say, "The T_1 value was X hours," always remember to ask, "Which model, unit or cumulative average?"

There is some value in being able to predict the hours for any unit given the hours for some other unit. But of even more value is the ability to predict the total hours for a "block" of production. For present purposes, we can define hours for a block of production as the total hours required to produce all units from some unit M to some other unit N, where N>M. This total we designate $T_{M,N}$, defined as:

$$\mathbf{T_{M,N} = H_1[M^b + (M+1)^b + (M+2)^b + \ldots + N^b]} \qquad \text{Eq(8)}$$

There is no concise exact formula for this sum that avoids taking the entire sum, term by term. However, a very good approximation formula has been developed. It is:

$$T_{M,N} = [H_1/(1+b)][(N+0.5)^{1+b} - (M-0.5)^{1+b}] \qquad Eq(9)$$

For readers who want to practice their skills in the integral calculus, this formula is readily evolved from the definite integral

$$\int_{M-.5}^{N+.5} x^b \, dx$$

The accuracy of this approximation formula is quite good. The absolute error is roughly the same (and small) at all production quantities, and the relative accuracy improves, as quantities get larger. The following table illustrates typical accuracy. The table is for a situation where $H_1 = 100$ hours, and $b = -0.2$, equivalent to a slope of 87%. Values in this table are rounded off to the nearest integer.

Qty	Exact	Formula		Qty	Exact	Formula
1	100	101		11	809	810
2	187	188		12	869	871
3	267	269		13	929	931
4	343	345		14	988	990
5	416	417		15	1047	1048
6	485	487		16	1104	1106
7	553	555		17	1161	1162
8	619	621		18	1217	1218
9	684	685		18	1272	1274
10	747	748		20	1327	1329

Example 7. An estimator believes that the 100th unit of production will require 320 hours. What total hours will be required for a production block ranging from unit 1 to unit 80, if the learning curve has a slope of 90%?

Answer: We first need a value for b:

$$b = \log(90/100)/\log(2) = -0.152$$

We can now compute H_1 using Eq(3):

$$H_1 = 320(1/100)^{-0.152} = 644.39 \text{ hours}$$

Also needed will be 1+b:

$$1+b = 1 - 0.152 = 0.848$$

Now we can use the cumulative approximation formula, Eq(9):

$$T_{1,80} = [644.39/0.848][80.5^{0.848} - .5^{0.848}] = 30,973 \text{ hours}$$

Without learning, the total hours would be $(644.39)(80) = 51,551$. Learning results in about a 34% saving in this case. Savings due to learning are typically large, suggesting that improvements in learning are good candidates for investment.

■■■

Learning curves can be used to estimate labor requirements in the factory. To do this, one must assume a production rate, or a scenario of production rates, over time.

Example 8. Consider the following production schedule:

Month	Qty Produced
1	10
2	20
3	30
4	40
5	40
6	40

Assume that $H_1 = 100$ hours, and that each production worker generates an average of 150 effective labor hours per month. Further, assume that the learning curve has a slope of 87% (which results in b = -0.2 and 1+b = 0.8).

The total hours approximation formula [Eq(9)] can tell us the hours consumed in each month. For example, in the first month:

$$T_{1,10} = (100/0.8)(10.5^{0.8} - 0.5^{0.8}) = 748$$

In the second month:

$$T_{11,30} = (100/0.8)(30.5^{0.8} - 10.5^{0.8}) = 1,105$$

Continuing these calculations, we can expand the above table to include a fourth column that shows hours by month. Then, in a fifth column, we can show heads by month by dividing hours by month by 150 (hours per month per head). Here are the results:

Month	Qty Produced	Cum Qty Produced	Hours by Month	Heads by Month
1	10	10	748	5.0
2	20	30	1105	7.4
3	30	60	1404	9.4
4	40	100	1667	11.1
5	40	140	1536	10.2
6	40	180	1450	9.7

∎∎

An occasionally useful way of looking at total hours when the U model is used is the midpoint concept. The formulas given so far deal with virtually any situation where the U model is being used, but you may encounter writings or a situation where the midpoint approach is being used, so it is included in this paper for completeness. The midpoint approach cannot give you results that you can't get from the methods already described, but it does result in a sometimes-useful rule of thumb. It is also useful in certain situation where the U model is fitted to data. This will be seen later.

Consider a production block running from unit M to unit N, inclusive, N>M. If the slope is less than 100%, then the unit of the block that will require the most hours will be unit M. The unit that will require the fewest hours will be unit N.

The number of units in the block is N-M+1, which for convenience we temporarily designate as Q:

$$Q = N-M+1$$

The total hours for the block can be estimated using Eq(9):

$$T_{M,N} = [H_1/(1+b)][(N+.5)^{1+b} - (M-.5)^{1+b}]$$

There must be some unit K between unit M and unit N that has the "average" number of hours in this block, such that:

$$T_{M,N} = H_K Q$$

But we also know that:

$$H_K = H_1 K^b$$

Therefore:

$$T_{M,N} = H_1 K^b Q$$

Hence:

$$K = (T_{M,N}/QH_1)^{1/b}$$

Substituting Eq(9) for $T_{M,N}$ yields the following result for K:

$$\mathbf{K = \{[(N+0.5)^{1+b} - (M-.5)^{1+b}]/[(1+b)(N-M+1)]\}^{1/b}} \qquad \text{Eq(10)}$$

K is usually called the "midpoint unit," although it is most unlikely to ever be half way between unit M and N. While a better name might be "average hours unit" of even "representative unit," the name midpoint unit seems to be well established, so that is what we call it. Since it is hypothetical, the midpoint need not be an integer, as the following example shows.

Example 9. If a production block runs from unit 201 to unit 500 inclusive, with a slope of 75%, what is the midpoint unit? Also, how many hours will be required for this block if H_1 = 450 hours?

Answer: We first need values for b and 1+b:

$$b = \log(75/100)/\log(2) = -0.415$$

$$1+b = 1-0.415 = 0.585$$

Then:

$$K = \{[500.5^{0.585} - 200.5^{0.585}]/[(0.585)(300)]\}^{-1/0.415} = 334.6$$

Now calculate

$$H_{334.6} = 450(334.6)^{-0.415} = 40.32 \text{ hours}$$

The average hours for the block are the hours for the midpoint unit, namely 40.32. The total hours for the block are therefore:

$$T_{201,500} = (40.32)(300) = 12{,}096 \text{ hours}$$

As an exercise, you might want to confirm this using Eq(9).

■■■

Example 10. If the slope is 80%, what is the midpoint of a block running from unit 1 to unit 100? What about a block running from unit 1 to unit 1,000?

Answer: For the block running from unit 1 to unit 100, first, we need b and 1+b:

$$b = \log(80/100)/\log(2) = -0.3219$$

$$1+b = 1\text{-}0.3219 = 0.6781$$

Next:

$$K = \{[100.5^{0.6781} - 0.5^{0.6781}]/[(0.6781)(100)]\}^{-1/0.3219} = 32.3$$

For the lot from 1 to 1,000:

$$K = \{[1000.5^{0.6781} - .5^{0.6781}]/[(0.6781)(1000)]\}^{-1/0.3219} = 304.3$$

■■■

Note in Example 10 that the midpoint was roughly one-third of the way through the lot for both lots. This is roughly true for all lots that begin at unit 1, for a variety of slopes. This gives rise to a rule of thumb that is sometimes used for rough (VERY rough!) estimating. The rule of thumb is:

For a lot beginning with unit 1, the midpoint is about a third of the way through the lot.

What can be done with this rule of thumb? The following example illustrates:

Example 11: A production block of 900 units is contemplated, beginning with unit 1. If unit 1 requires 1,000 hours, and the slope is 87%, use the rule of thumb to estimate how many hours the block will require.

Answer: Assume that the midpoint is approximately at unit 300, one third of the way through the lot. Also, we need:

$$b = \log(85/100)/\log(2) = -0.2$$

Hence:

$$H_{300} = 1000(300)^{-0.2} = 320 \text{ hours}$$

$$T_{1,900} = (320)(900) = 288{,}000 \text{ hours}$$

■■

Example 12. Using Eq(9), compute the "correct" total hours for Example 11 for comparison.

Answer:

$$T_{1,900} = (900/0.8)(900.5^{0.8} - .5^{0.8}) = 259{,}214 \text{ hours}$$

This compares to 288,000 hours obtained by the midpoint rule of thumb method. The error for the midpoint method is about 11% in this case.

■■

Note that the midpoint rule does not apply to blocks that do not begin at unit 1. For these lots, the "midpoint" typically is closer to 40-50% of the way through the block.

For a summary of all the formulas associated with the U model, please see the appendix.

IV. The CA Learning Model

The CA (cumulative average) version of the learning curve can be written:

$$\mathbf{A_n = H_1 n^b} \qquad\qquad \text{Eq(11)}$$

In this equation, A_n is the average hours required for the first n units of production, and H_1 is the hours required for the first unit. As in the U model, b is the "natural slope" of the learning curve, reflecting whether learning proceeds rapidly or slowly. This equation assumes that H_1 and b are known, as well as the unit of production of interest. From this information the average hours for all units from 1 to n can be computed.

Example 13. An estimator believes that the first unit of production will require 500 hours. What will be the average number of hours for units 1 through 150 if b = -0.2?

Answer: Use the equation:

$$A_n = H_1 n^b$$

Then:

$$A_{150} = 500(150)^{-0.2} = 183.5 \text{ hours}$$

▪▪▪

The slope b in the CA model does not have the same meaning as in the U model. However, the model uses the same method of slopes expressed as percentages that is used in the U model, and the conversion formulas are the same, namely:

$$b = \log(S/100)/\log(2)$$

and:

$$S = 10^{b \log(2)+2}$$

It is very easy to compute total hours for a production block using the CA model. For blocks beginning at unit 1 and ending at some unit N, clearly:

$$\mathbf{T_{1,N} = A_N N = H_1 N^{1+b}} \qquad\qquad \text{Eq(12)}$$

Here $T_{1,N}$ is the total hours required for units 1 through N.

Example 14.   An estimator believes that the first unit of a new product will require 150 hours. If the slope is 95%, what will the first 50 units require?

Answer: First we need b and 1+b:

$$b = \log(95/100)/(\log(2) = -0.074$$

$$1+b = 1-0.074 = 0.926$$

Then:

$$T_{1,50} = 150(50)^{0.926} = 5{,}615 \text{ hours}$$

• •

It is not quite as easy, but still not difficult, to compute the total hours for a block beginning at unit M and ending at unit N, N>M. It is easy to see that this result can be obtained from

$$\mathbf{T_{M,N} = T_{1,N} - T_{1,M\text{-}1} = H_1[N^{1+b} - (M\text{-}1)^{1+b}]} \qquad \text{Eq(13)}$$

Example 15.   Production of the first 200 units of a product is nearing its end. Your customer has said he will buy an additional 50 units. There will be no break in production or in learning. The first unit required 75 hours. The first 200 units required 10,400 hours altogether. How many hours will the second block of 50 units require?

Answer: Virtually every learning curve problem requires knowledge of slope either in the form of b or of S. Here, we are given neither b nor S. Looking at the equation:

$$T_{1,N} = H_1 N^{1+b}$$

we see that we are given $T_{1,N}$, H_1, and N. To find b, we take the logarithm of both sides, then solve for b:

$$\log(T_{1,N}) = \log(H_1) + (1+b) \log(N)$$

$$b = [\log(T_{1,N}) - \log(H_1)]/[\log(N)] - 1$$

$$b = [\log(10400) - \log(75)]/[\log(200)] - 1 = -0.06912$$

We don't need to know S to proceed with the calculations, but it could be information we want for other purposes:

$$S = 10^{-0.06912\log(2)+2} = 95.32\%$$

We will also want 1+b:

$$1+b = 1-0.06912 = 0.93088$$

Now that we have 1+b, we can use Eq(13):

$$T_{M,N} = H_1[N^{1+b} - (M-1)^{1+b}]$$

$$T_{201,250} = 75[250^{0.93088} - 200^{0.93088}] = 2,401 \text{ hours}$$

· ·

When using the CA model, there can be situations where we want to know the hours for a particular unit. This is easy in the unit model, but a little more difficult in the CA model. Consider again Eq(13):

$$T_{M,N} = H_1[N^{1+b} - (M-1)^{1+b}]$$

If N happens to equal M+1, the equation must give the hours for unit N. Hence we can write:

$$\mathbf{H_n = H_1[n^{1+b} - (n-1)^{1+b}]} \qquad\qquad \text{Eq(14)}$$

By analogy with Eq(3) of the unit model, we can also write:

$$\mathbf{H_n = H_m[n^{1+b} - (n-1)^{1+b}]/[m^{1+b} - (m-1)^{1+b}]} \qquad \text{Eq(15)}$$

Example 16. In a production situation where the CA model is being used, introduction of new hires to replace retirees seems to be affecting the learning slope. The effects seem to coincide approximately with production of unit 2,200. To structure a new learning curve from unit 2,200 on, the estimator wants to know the unit hours for unit 2,200. The hours for unit 1 were 110, and the slope for the units from 1 to 2,200 was S = 88.5%.

Answer: We first need the b and the 1+b values corresponding to S = 88.5%:

$$b = \log(88.5/100)/\log(2) = -0.17625$$

$$1+b = 1-0.17625 = 0.82375$$

Next, we calculate:

$$H_{2200} = 110[2200^{0.82375} - 2199^{0.82375}] = 23.34 \text{ hours}$$

Note that in any equation (such as the above) where you are subtracting one big number from another big number of similar magnitude, you need to be careful to use considerable precision of the numbers to avoid serious error.

• •

In Example 8, we showed how the U model could be used to estimate manufacturing manpower requirements. The CA model can also be used for this purpose, as the following example illustrates.

Example 17. Consider the following planned production schedule:

Month	Qty Produced	Cumulative Qty Produced
1	20	20
2	40	60
3	60	120
4	60	180
5	60	240
6	40	280

Assume that unit 1 requires 28 hours, and that the slope will be 94%. Assume also that each worker contributes effectively 150 hours per month. Create a monthly manpower plan.

Answer: We will need 1+b:

$$b = \log(.94)/\log(2) = -0.08927$$

$$1+b = 1-0.08927 = 0.91073$$

For the hours in each month, we will use Eq(13):

$$T_{M,N} = H_1[N^{1+b} - (M-1)^{1+b}]$$

For example, in month 1:

$$T_{1,20} = 28[20^{0.91073}] = 428.6$$

In month 2:

$$T_{21,60} = 28[60^{0.91073} - 20^{0.91073}] = 737.1$$

Continuing in this manner, we can populate the table below.

Month	Qty Produced	Cumulative Qty Produced	Hours by Month	Labor Heads by Month
1	20	20	428.6	2.9
2	40	60	737.1	4.9
3	60	120	1025.8	6.8
4	60	180	978.9	6.5
5	60	240	949.6	6.3
6	40	280	621.0	4.1

The rightmost column is the Hours by Month column divided by 150, the contribution of each worker in hours per month.

■■■

For a summary of all the formulas associated with the CA model, please see the appendix.

V. Comparison of the U and CA Models

While the U and CA models are both based on the underlying power law formula, they are quite different in the way they work. As we have already pointed out, the U model looks at the learning effect as a phenomenon affecting each individual unit. Under this model, every unit requires fewer hours than the unit just before it. The CA model, on the other hand, regards learning as a phenomenon affecting the average hours required for a sequence of production units. As the sequence increases in length, the average decreases.

Both models have their proponents, and both have been found useful. The purpose here is not to advocate one or the other, but to make certain graphical comparisons that should help analysts better understand their workings and how they differ.

Let's assume that you are able somehow to estimate H_1 as 50 hours for a particular product. Further, let's assume that you believe the learning slope to be 87% (b = -0.2). You are not sure whether the 87% is based on the U model or on the CA model. You want to compare the behavior of the two models to see what results each would give. There are two elegant ways to make this comparison. One way is to look at what each model says about unit hours; the other is to look at what each model says about cumulative hours.

For the first comparison, we can use the U model formula:

$$H_n = H_1 n^b$$

and the CA model formula:

$$H_n = H_1[n^{1+b} - (n-1)^{1+b}]$$

To compare what the plots of these equations look like for common values of H_1 and b, your author used a side-by-side spreadsheet tabulation for all units from 1 to 100. Exhibit 3 shows the plots that result (please recall that $H_1 = 50$ hours and b = -0.2).

Exhibit 3 -- Comparison of U and CA Models (Units)
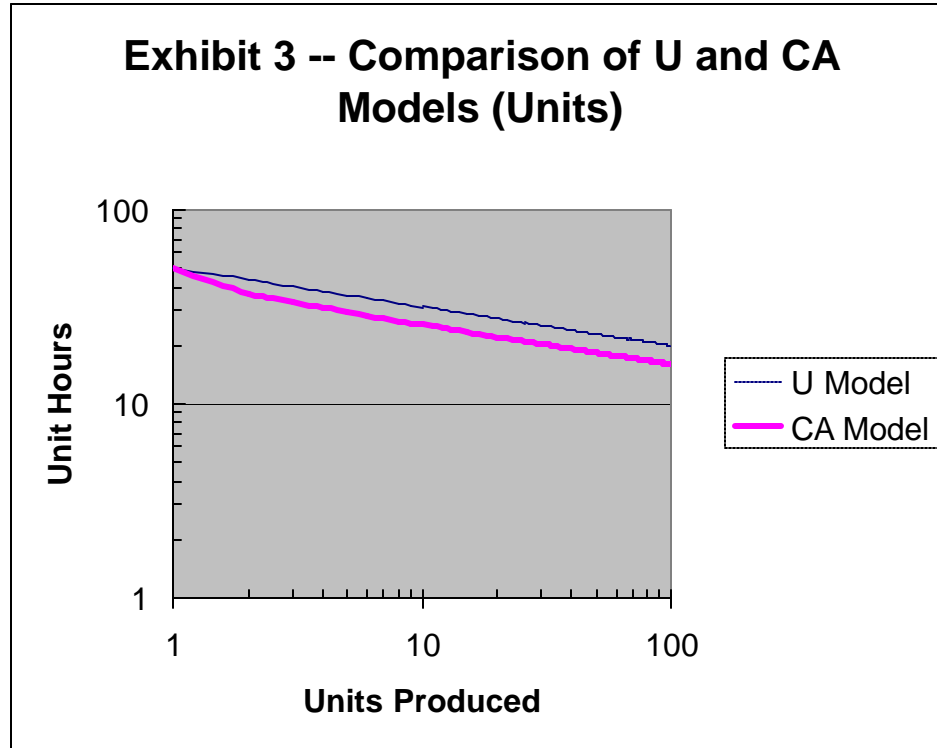
Note that the U model plots as a straight line in log-log coordinates because it is a power law expression. The unit expression of the CA model does not plot as a straight line because it is not a "pure" power law. Both models give the same result for unit 1 because that was assumed, but the CA model gives a lower result thereafter. Note that the two plots become essentially parallel at about units 4 or 5.

This plot makes the point that if someone tells you that a cost analysis was based on a certain value of H_1 and a certain value of b or S, it matters considerably whether they are talking about the U model or the CA model. This point is made even more forcefully when you compare the cumulative plots. The cumulative plots in Exhibit 4 are based on the same H_1 and slope values as Exhibit 3. These plots were made using the equation:

$$T_{M,N} = [H_1/(1+b)][(N+.5)^{1+b} - (M-.5)^{1+b}]$$

with M = 1 for the U model, and the equation:

$$T_{M,N} = H_1[N^{1+b} - (M-1)^{1+b}]$$

with M = 1 for the CA model.

Exhibit 4 -- Comparision of U and CA Models (Cumulative)
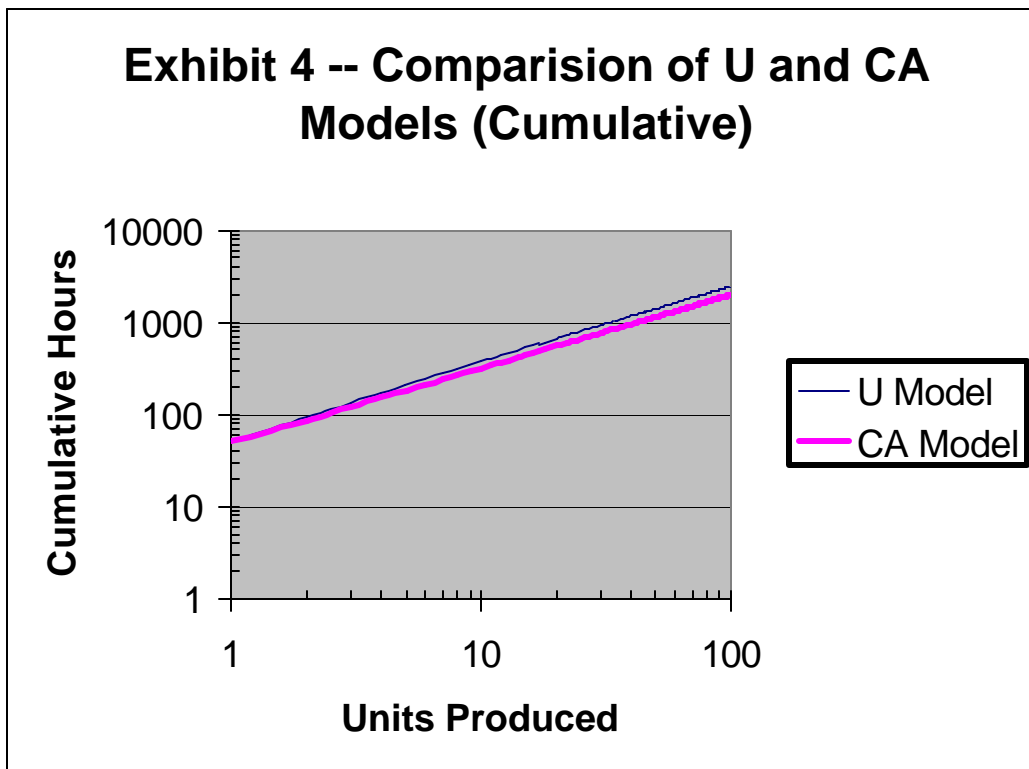
The log-log plot makes the difference appear smaller than it actually is. The value on the U plot at 10 units is about 19% larger than the value on the CA plot. At 100 units, the value on the unit plot is about 24% larger.

It is not easy to see on the plot, but the CA plot is a straight line because it is a power law expression, while the U plot is not a straight line because it is not. However, it is very close to a straight line.

VI. Typical Learning Slopes

The best source of learning slopes is usually production experience, but if experience is not available, it is helpful to have a rough slope in mind when doing a cost analysis. This section presents some slope values that are "in the ballpark" in most situations. Because they are merely approximate, they can be said to be valid for both the U and the CA models.

Operations that are fully automated tend to have slopes of 100%, or a value very close to that. Operations that are entirely manual tend to have slopes in the vicinity of 70%.

If an operation is 75% manual and 25% automated, slopes in the vicinity of 80% are common. If they are 50% manual and 50% automated, expect a slope of about 85%. If they are 25% manual and 75% automated, expect about 90%.

The average slope for the aircraft industry is about 85%. But there are departments in a typical aircraft factory that may depart substantially from that value.

Shipbuilding slopes tend to run between 80 and 85%. Shipbuilding is typically somewhat more labor intensive that aircraft manufacture.

The following typical values assume repetitive operations. They are not valid if operations are sporadic, as in a job shop environment.

Manufacturing Activity	Typical Slope %
Electronics	90-95
Machining	90-95
Electrical	75-85
Welding	88-92

A slope of 93-96% is often applied to raw materials, based on increasing procurement efficiencies, higher yields, and lower scrap rates as manufacturing progresses. A slope in the 80's is typical for purchased parts, with 85% a reasonable average value.

When very large quantities will be built, slopes tend to flatten, because manufacturing planners depend on economies of scale to build better tooling and use more automation. The flattening of slopes for large quantities is typically accompanied by a reduction in first unit hours. This effect has been used to estimate the amount that can be spent on automation. The answer sometimes comes out in favor of automation, but that is not always the case. Many

automobile factories have found automation to be profitable, but there is much less automation, at least so far, in the aircraft industry, because the savings do not justify the high cost of automation equipment.

Slopes also tend to flatten, even for small quantities, when there are many minor changes in product configuration as manufacturing progresses. This often happens for products where each item is somewhat customized. Learning will approach 100% if each item is heavily customized, i.e., becomes "one of a kind."

Slopes tend to be flatter if a project is closely similar to a previous project, the time gap between them is not too large, and many of the same people will be involved. This is sometimes called the "heritage" effect.

Experienced crews tend to have lower first unit costs than inexperienced crews, and since they are already knowledgeable, their learning rate tends to be less. Inexperienced crews tend to have higher first unit costs, and higher learning rates.

A rush to production before all of the bugs are worked out of the product and the manufacturing process can result in low learning early in production, that eventually transitions into a more normal learning rate. This phenomenon is also sometimes associated with a gradual ramp up of production rate. See Section IX for information on modeling these situations.

If you have good data from past manufacturing experience that you believe is applicable to a future project, you should use it to determine learning slope. That subject is dealt with in Section XI.

VII. Error Analysis

You can't ever know for sure how much future error you will have when using a learning curve model. If you could, you could simply correct for it and have no error. But it can be useful to look at error that can possibly result when you use the wrong values in a learning curve calculation. Information about the error that might result could prompt you to make certain prudent decisions that you might not otherwise make. One such decision could be to search for more and better historical data so your learning curve has a more realistic foundation. Another might be to increase a bid amount to decrease the risk of losing money on a project.

In any situation, there are many ways you can be wrong, and often only one way you can be right. The prudent thing to do is to look at the main drivers that can move you away from being right, and try to assess and mitigate their effects. If a given source of error doesn't make much difference, you can pay less attention to that than to one that is a "big swinger."

In learning curves, the main possible sources of error are:

- Error in estimating production total quantity, N
- The possibility that the production work might not follow the learning model you have chosen
- Error in estimating H_1
- Error in estimating S (or equivalently, b).

Here we are going to assume that you can either know N, or characterize the risk from not knowing, and deal with that risk in an appropriate way. As for whether the work follows the model you have chosen, experience and good judgment based on experience are the best guides. That leaves errors in H_1 and S, or equivalently, b.

In both the U and the CA models, H_1 is always a simple multiplier. Therefore if you make an error of x% in estimating H_1, you will also make the same x% error in estimating $T_{1,N}$. If you can tolerate no more than (say) a 10% error in $T_{1,N}$, then the error in H_1 must be no larger than 10%. And that's if there is no error in S!

Because b is an exponent in the learning curve equations, results are very sensitive to errors is b. Consider, for example, the CA model, given by:

$$T_{1,N} = H_1 N^{1+b}$$

If you, in error, chose the value c when you should have chosen the value b in this model, then the ratio of your actual cumulative result to the result you should have obtained will be R, given by:

$$\mathbf{R = N^{1+c}/N^{1+b} = N^{c-b}} \qquad \text{Eq(16)}$$

R can be viewed as a measure of the error due to an incorrect choice of the slope. The above equation makes it clear that this error is a function of both N, and the absolute slope error, c-b. Said another way, for a given c-b the error gets worse the larger the value of N. Another view of the situation is that you need to be much more accurate in your choice of slope for large values of N than for small values.

In practice, the choice of slope generally begins with a choice of S. S is then converted to b for calculations. The usual range of S in industry is from 70% to 100%. Exhibit 5 below shows the corresponding range of b, and plots the relationship.

Exhibit 5 -- b Versus S
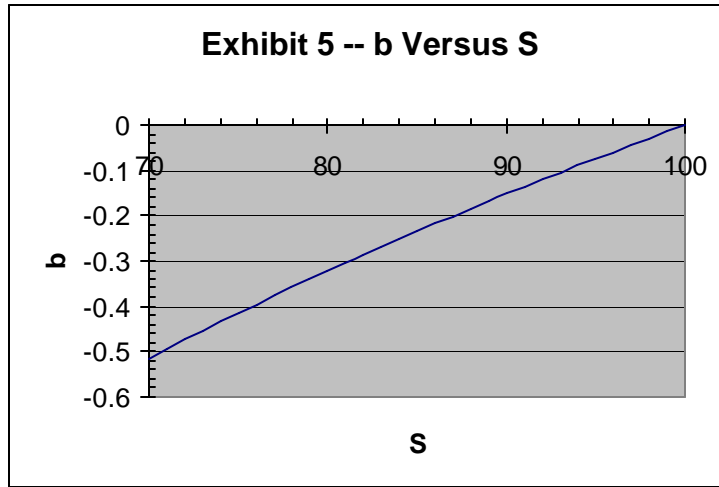
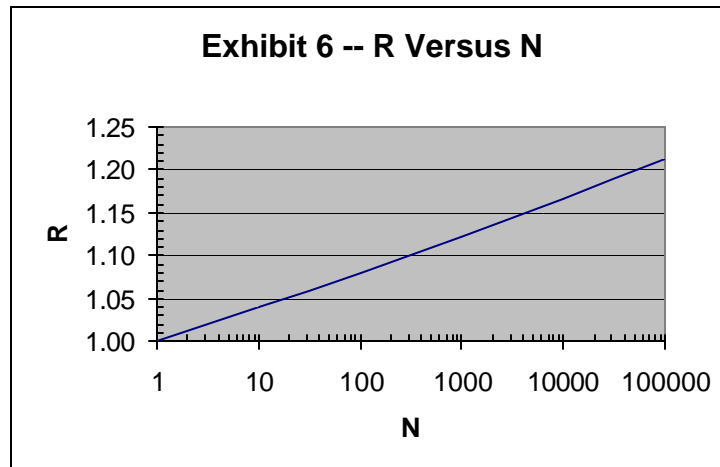As this exhibit illustrates, if you look carefully, each one-percentage point change in S represents approximately a change in b of magnitude about 0.017.

We can use this to express Eq(16) in a form that estimates the error R in terms of the error in S, written as δS.

$$\mathbf{R = N^{0.017^{d}S}} \qquad \text{Eq(17)}$$

(approximately). Plotting R versus N in Exhibit 6 for $\delta S = 1\%$:

Exhibit 6 -- R Versus N

This exhibit shows that at 10 units produced, the estimating error due to a one-percentage point error in S is about 4%. At 100 units produced, it is about 8%. At 1,000 units it is about 12%. At 10,000 units it is about 17%, and at 100,000 it is about 22%. (These estimates are slightly conservative. In practice, the errors will be a bit smaller.)

Example 18. In a certain project, H_1 was accurately estimated at 50 hours. The learning slope was estimated at 87% based on the CA model, but it turned out to be 88%. The production quantity was 1,000 units. According to Exhibit 6, the error should be about 12%. Compute the actual error.

Answer: We first compute the total hours based on an 87% slope, and then we compute the total hours based on an 88% slope. Finally, we compute the percent error.

For an 87% slope, b = -0.2 and 1+b = 0.8.

$$T_{1,1000} = 50(1000)^{0.8} = 12,559 \text{ hours}$$

For an 88% slope, b = -0.18442 and 1+b = 0.81558.

$$T_{1,1000} = 50(1000)^{0.81558} = 13,987 \text{ hours}$$

The actual error is about 11.4%. The approximation in Exhibit 6 is reasonable.

▪▪

Clearly, if one errs in the direction of setting S too low, as in the above example, the result will be an overrun in the hours. On the other hand, if one sets S too high, the result will be an under run. Your author cannot know which you

would prefer to have in a given set of circumstances, but you do, so keep this relationship in mind.

The above analysis of the error effect of slope was done based on the CA model, because of its mathematical simplicity for this purpose. However, it applies with reasonable accuracy to the U model also. The following example illustrates that point.

Example 19. Repeat Example 18 except use the U model.

Answer: For an 87% slope:

$$T_{1,1000} = [50/0.8][1000.5^{0.8} - .5^{0.8}] = 15,670 \text{ hours}$$

For an 88% slope:

$$T_{1,1000} = [50/0.8][1000.5^{0.81558} - .5^{0.81558}] = 17,455 \text{ hours}$$

The percentage error is about 11.4%.

∎∎

The percentage errors due to misestimating H_1 and S are roughly additive. For example, if you estimate H_1 10% too low and the slope S one percentage point too low, your overrun will be about 22% at a quantity of 1,000. On the other hand, if you estimate H_1 10% too high and the slope S one percentage point too low, the errors will partially cancel out. At a quantity of 1,000, the overrun would be about 2%. Estimators should keep this general relationship in mind. The following example illustrates.

Example 20. In Example 18, S was set one percentage point too low at 87%, resulting in about a 12% overrun in a quantity of 1,000. H_1 was set exactly right at 50 hours. Repeat Example 18 with H_1 set 12% too high at 56 hours. Check to see that this approximately cancels the error in setting S too low.

Answer: With H_1 set 12% too high at 56 hours and S set too low at 87%, the error should approximately cancel vis-à-vis using the correct settings of $H_1 = 50$ hours and S = 88%. Previously, in Example 17, we found that with the correct settings $T_{1,1000} = 13,986$ hours. With the incorrect but offsetting errors $H_1 = 56$ hours and S = 87%, we obtain:

$$T_{1,1000} = 56(1000)^{0.8} = 14,067 \text{ hours}$$

This result differs by about one half of one percent from the correct value of 13,986 hours.

▪▪

Here are the important points made in this section:

- Slopes are exponents; serious errors in estimates can result when they are poorly chosen
- If you want to hedge against an overrun, set the percentage slope higher, but if you want to shave costs and be more competitive, set it lower
- The estimating error due to poor choice of S increases significantly as the production quantity increases
- You can (at least to some extent) offset an error in slope by making an "opposite error" in H_1.
- Error increases rapidly as production quantity increases; for quantities in the high thousands, use great care in selecting a learning slope.

A final thought: Many firms put the results of learning curve calculations on factory routing paperwork. If this is done, workers will see it and will know the expectation of the hours to do the work, based on the analyses done by the estimators. If factory morale is good, and the learning rate is not perceived as oppressive, good workers will try to "make the numbers." This can afford some protection against estimating errors, because the learning curve becomes a self-fulfilling prophecy.

VIII. Aggregated Learning Curves.

It is common for a large factory to track learning rates by department. Learning rates typically differ in diverse operations such as foundry, machining, sheet metal fabrication, and final assembly. When these rates are individually tracked, it is possible for each department to have its own learning curve for a given product. Estimators then can create separate estimates for each department based on that department's learning profile.

Contract negotiators and certain other analysts, on the other hand, typically deal more with the "big picture." If the customer wants to change the production quantity, or perhaps make certain other changes, they may not want to or have the time to revisit the learning for each department. They want an aggregated learning curve for the entire factory that is reasonably accurate at any production quantity of the product. Aggregated learning curves can also be useful when learning is applied to both labor and material.

This section develops a simple and robust method for creating an aggregated learning curve. The beginning point of the method is that two sums must be maintained:

- The aggregated H_1 value across all departments must be equal to the sum of the H_1 values for each department
- The aggregated $T_{1,N}$ value across all departments must be the same as the sum of the $T_{1,N}$ values for each department, for the expected production quantity, N.

Mathematically, we represent these conditions by the equations

$$H_{1all} = \sum H_1$$

$$T_{1,Nall} = \sum T_{1,N}$$

Here, it is understood that the summations and the subscript "all" embrace all departments.

If we accept these principles as being reasonable, then for the CA model we can write

$$\mathbf{T_{1,Nall} = H_{1all}N^B} \hspace{3cm} \text{Eq(18)}$$

where B is a composite slope to be determined. To solve for B, we take the logarithm of both sides of Eq(18):

$$\log(T_{1,Nall}) = \log(H_{1all}) + B\log(N)$$

$$\mathbf{B = [\log(T_{1,Nall}) - \log(H_{1all})]/\log(N)} \qquad \text{Eq(19)}$$

Example 21. At a certain company, a project is expected to get underway soon to produce 300 Super Widgets. Three departments will be involved: Machining, Sheet Metal, and Final Assembly. Historically, with similar projects, the learning curves for these departments have had slopes 85%, 87% and 80%, based on the CA model. The first unit hours for these departments for the Super Widget have been estimated at 70, 45, and 25. Develop a composite learning curve for the entire effort.

Answer: We see immediately that

$$H_{1all} = 70 + 45 + 25 = 140 \text{ hours}$$

We next estimate the total hours for each department. For Machining:

$$b = \log(0.85)/\log(2) = -0.23447$$

$$1+b = 1-0.23447 = 0.76553$$

$$T_{1,300} = 70(300)^{0.76553} = 5{,}513 \text{ hours}$$

For Sheet Metal:

$$b = \log(0.87)/\log(2) = -0.2$$

$$1+b = 1-0.2 = 0.8$$

$$T_{1,300} = 45(300)^{0.8} = 4{,}314 \text{ hours}$$

For Final Assembly:

$$b = \log(0.8)/\log(2) = -0.32193$$

$$1+b = 1-0.32193 = 0.67807$$

$$T_{1,300} = 25(300)^{0.67807} = 1{,}196 \text{ hours}$$

Hence:

$$T_{1,300all} = 5{,}513 + 4{,}314 + 1{,}196 = 11{,}023 \text{ hours}$$

Next:

$$B = [\log(11023) - \log(140)]/\log(300) = 0.76547$$

The aggregation model we seek is therefore:

$$T_{1,Nall} = 140N^{0.76547}$$

For N = 300, this model should give 11,023 hours. Let's see if it does.

$$T_{1,300all} = 140(300)^{0.76547} = 11{,}023 \text{ hours}$$

■■

Example 22. You are negotiating the Super Widgets of Example 21 with your prospective customer. Your firm uses a burdened labor rate, including profit, of $80/hour as an average of all departments. Therefore your price for the labor is

$$(80)(11023) = \$881{,}840$$

The Super Widgets are upgrades of the existing Widgets, furnished by your customer, and there are no additional materials costs. Labor is the entire cost. You believe the $80/hour figure covers your risks.

Your customer says he cannot afford to spend this much, and wants to know how many Super Widget upgrades he can get for $600,000. Use the composite formula to answer his question.

Answer: The number of hours your customer can buy for $600,000 is x, where

$$80x = \$600{,}000$$

$$x = 7{,}500 \text{ hours}$$

The production quantity must be N, such that:

$$T_{1,N} = 140(N)^{0.76547} = 7{,}500 \text{ hours}$$

Therefore

$$N = (7500/140)^{1/0.76547} = 181 \text{ Super Widgets}$$

■■

Example 23. Your firm is considering bidding manufacture of an advanced Doobob. Your factory has a single average labor rate of $75/hour, and does not record learning by departments. For products such as the Doobob, your overall learning rate is typically 83.5%, based on the CA model. For materials, you have found that a 95% learning rate closely matches your actual experience, again, based on the CA model. The H_1 hours for the Doobob have been estimated at 3,450, and the H_1 material at $365,200. The planned production quantity is 35. Develop an aggregate model.

Answer: Because materials are in dollars, the model will be developed in dollars rather than in hours. For labor, the H_1 hours are 3,450, and the labor rate is $75/hour, therefore the H_1 labor cost is:

$$\$H_1 = (\$75)(3450) = \$258,750$$

Also for labor:

$$b = \log(0.835)/\log(2) = -0.26015$$

$$1+b = 1-0.26015 = 0.73985$$

$$\$T_{1,35} = \$258,750(35)^{0.73985} = \$3,591,361$$

For material:

$$\$H_1 = \$365,200$$

$$b = \log(0.95)/\log(2) = -0.074$$

$$1+b = 1-0.074 = 0.926$$

$$\$T_{1,35} = \$365,200(35)^{0.926} = \$9,825,118$$

Hence:

$$\$H_{1all} = \$258,750 + \$365,200 = \$623,950$$

$$\$T_{1,35\ all} = \$3,591,337 + \$9,825,118 = \$13,416,479$$

$$B = [\log(13,416,479) - \log(623,950)]/\log(35) = 0.86297$$

The aggregate model is therefore:

$$\$T_{1,Nall} = \$623{,}950N^{0.86297}$$

For N = 35, the model should give $T_{1,35all} = \$13{,}416{,}479$. Let's see if it does.

$$\$T_{1,35all} = \$623{,}950(35)^{0.86297} = \$13{,}416{,}344$$

As a percentage, the error is tiny.

▪▪▪

The development so far has been based exclusively on the CA model. Mathematical difficulties arise when we attempt an analogous process with the U model. The difficulty arises in solving for B in the more complex cumulative equation of the U model. This difficulty is easily avoided by always using the CA model as an aggregate model. This presents no difficulties, even if your component estimates all use the U model. We use an example to demonstrate how this works.

Example 24. You are now a competitor of the firm in Example 23. You are also bidding on the Doobob, but you have a somewhat different design for it. By custom, your firm uses the U learning model. You use an 85% learning rate for labor based on the U model. You also use a 96% learning rate for material based on the U model. Your composite factory labor rate is $72/hour. Your H_1 hours estimate is 4,520 hours, and your H_1 material estimate is $325,000. Develop an aggregate learning model.

Answer: We follow the same two principles used heretofore, namely that the model must be exact at H_1 and at $T_{1,N}$. Again, we put the model in dollar terms rather than in hours, because material is in dollars.

For labor:

$$\$H_1 = (4{,}520)(\$72) = \$325{,}440$$

$$b = \log(0.85)/\log(2) = -0.23447$$

$$1+b = 1 - 0.23447 = 0.76553$$

$$\$T_{1,35} = (\$325{,}440/0.76553)(35.5^{0.76553} - 0.5^{0.76553}) = \$6{,}285{,}072$$

For materials:

$$\$H_1 = \$325{,}000$$

$$b = \log(0.96)/\log(2) = -0.05889$$

$$1+b = 1-0.05889 = 0.94111$$

$$\$T_{1,35} = (\$325{,}000/0.94111)(35.5^{0.94111} - 0.5^{0.94111}) = \$9{,}755{,}378$$

Hence for labor and material combined:

$$\$H_{1all} = \$325{,}440 + \$325{,}000 = \$650{,}440$$

$$\$T_{1,35all} = \$6{,}285{,}072 + \$9{,}755{,}378 = \$16{,}040{,}450$$

B in the aggregate model is therefore:

$$B = [\log(\$16{,}040{,}450) - \log(\$650{,}440)]/\log(35) = 0.90152$$

The aggregate model is therefore:

$$\$T_{1,Nall} = \$650{,}440N^{0.90152}$$

....

The aggregate model we have been discussing exactly matches the sums of the various contributing learning curves at unit one and at unit N, the planned production quantity. At intermediate values, it is not a perfect match, however, it is generally quite accurate. Errors at other quantities usually are on the order of 1% or 2%, sometimes as high as 5% when slopes are far apart. A general analysis of errors is all but impossible. If an aggregated learning curve is useful to you, and you are concerned about errors, you should run a few checks on it before you use it, as in the following example.

Example 25. Consider a situation where there are two departments. Each has an H_1 value of 100 hours, but one has a learning slope of 75%, and the other has a learning slope of 95%, both based on the CA model. The slopes being so different will tend to result in larger errors in the aggregated model than is usually the case.

100 units are to be produced, according to the current plan. Create an aggregated model, and then test it for accuracy at 1, 25, 50, 75, 100, and 110 units.

Answer: Immediately we have:

$$H_{1all} = 100 + 100 = 200 \text{ hours}$$

Then, for the 75% slope:

$$b = \log(0.75)/\log(2) = -0.41504$$

$$1+b = 1-0.41504 = 0.58496$$

$$T_{1,100} = 100(100)^{0.58496} = 1{,}479 \text{ hours}$$

For the 95% slope:

$$b = \log(0.95)/\log(2) = -0.074$$

$$1+b = 1-0.074 = 0.926$$

$$T_{1,100} = 100(100)^{0.926} = 7{,}112 \text{ hours}$$

Therefore:

$$T_{1,100all} = 1{,}479 + 7{,}112 = 8{,}591 \text{ hours}$$

Solving for B:

$$B = [\log(8591) - \log(200)]/\log(100) = 0.81651$$

The aggregate model is therefore:

$$T_{1,100all} = 200(N)^{0.81651}$$

Testing this model at the prescribed quantities:

Quantity	Exact Solution	Aggregated Model	Error of Aggregated Model %
1	200	200	0
25	2,627	2,770	+5.4
50	4,729	4,878	+3.2
75	6,699	6,793	+1.4
100	8,591	8,591	0
110	9,332	9,286	-0.5

．．．

IX. Segmented Learning Curves

Sometimes analysts have need for a learning curve model other than the U and CA models discussed so far. Most of the time, these situations can be handled by means of a segmented learning curve model. A segmented learning curve can be based either on the U or the CA model, but applies it in segments. For example, from unit 1 to unit M, the model may have one slope, while from unit M+1 to unit N, it may have a different slope. Two segments will adequately model many situations, but models with three or more segments are quite feasible.

We will develop two situations to illustrate the modeling process (many other situations are possible). For both of these situations, we will develop both a U and a CA approach.

The first situation is where the analyst believes that after a certain number of units, learning ceases altogether. Imagine that the planned production quantity is N, but the analyst has reason to believe that after the first M units, M<N, learning will cease. This is equivalent to saying that at M units, the production work force reaches a "barrier" of efficiency beyond which it cannot go. It is also equivalent to saying that for units M+1 and higher, the learning slope is 100%, and the hours required per unit are the same as for unit M.

The second situation is where the analyst believes that in the early stages of production, learning will be slow, but at some unit M, learning will achieve a more normal pace. This situation is sometimes associated with a rush to production before all of the bugs have been worked out in the design and in the production processes. It is also frequently associated with situations where production is ramped up gradually.

We begin with two examples of cessation of learning, the first example using the U model and the second using the CA model.

Example 26. Your company will soon begin production of a small aircraft. Based on past experience, the learning slope will be 84% from unit 1 to unit 150 based on the U model. Thereafter it will be 100% (no learning). A total of 200 units are planned. The first unit hours (H_1) will be 1,500. No learning is expected for materials. Estimate the total hours required for all 200 units.

Answer:

$$b = \log(0.84)/\log(2) = \text{-}0.25154$$

$$1+b = 1\text{-}0.25154 = 0.74846$$

For the first 150 units, using Eq(9):

$$T_{1,150} = (1500/0.74846)(150.5^{0.74846} - 0.5^{0.74846}) = 84{,}259 \text{ hours}$$

Now we need the hours for unit 150, because all units past 150 will have that number of hours.

$$H_{150} = 1500(150)^{-0.25154} = 425 \text{ hours}$$

$$T_{151,200} = (50)(425) = 21{,}250 \text{ hours}$$

Therefore:

$$T_{1,200} = 84{,}259 + 21{,}250 = 105{,}509 \text{ hours}$$

•••

Example 27. Your company plans to build an electronic consumer item. The planned production quantity is 10,000. Based on past experience, learning will likely stop at unit 5,000. The slope up to that point will be 92%, based on the CA model. The H_1 hours are expected to be 13.5. Estimate the total hours.

Answer:

$$b = \log(0.92)/\log(2) = -0.12029$$

$$1+b = 1-0.12029 = 0.87971$$

For the first 5,000 units:

$$T_{1,5000} = 13.5(5000)^{0.87971} = 24{,}230 \text{ hours}$$

We need the hours for unit 5,000, because all subsequent units have the same hours.

$$H_{5000} = 13.5(5000^{0.87971} - 4999^{0.87971}) = 4.263 \text{ hours}$$

$$T_{5001,10000} = (5000)(4.263) = 21{,}315 \text{ hours}$$

$$T_{1,10000} = 24{,}230 + 21{,}315 = 45{,}545 \text{ hours}$$

•••

We now look at the situation where learning is initially slow, then improves after M units have been built. Exhibit 7 illustrates the situation as it might appear in log-log coordinates.

Exhibit 7 – Example Segmented Learning Curve with Initial Slow Learning
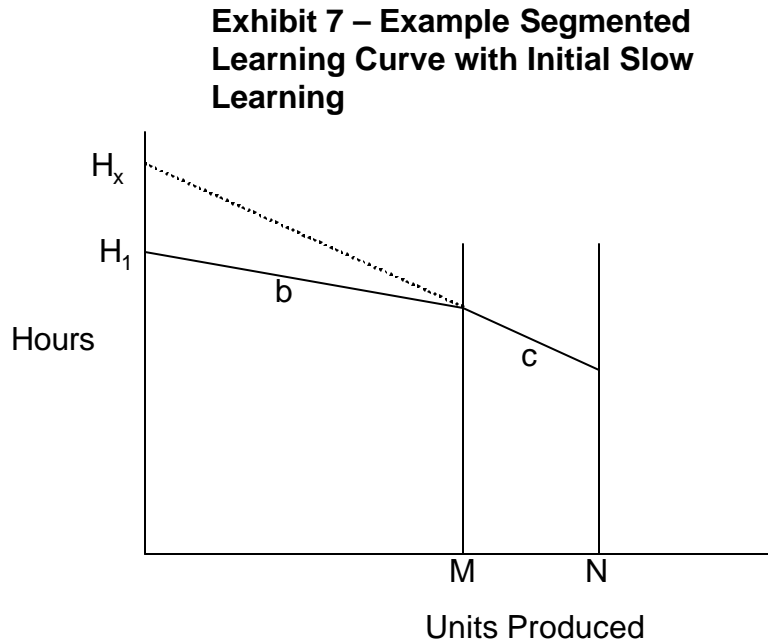
In Exhibit 7, beginning at unit 1, the unit hours decline along a slope b. This continues until the transition at unit M, where the slope changes to c. Slope c continues until unit N, the end of production. We define H_x in the exhibit above as the hypothetical first unit hours of a learning curve that matches the hours of unit M but has slope c. We also call it the projected first unit hours on slope c.

We now describe a method for conveniently and quickly working with segmented learning curves. If we are working with the U model, we can compute the hours from unit 1 to unit M using Eq(9):

$$T_{1,M} = [H_1 / (1+b)][(M+0.5)^{1+b} - 0.5^{1+b}]$$

We can compute the hours for unit M as follows:

$$H_M = H_1 M^b$$

But also:

$$H_M = H_x M^c$$

Hence:

$$H_x = H_1 M^{b-c}$$

Now we can write:

$$T_{M+1,N} = [H_x/(1+c)][(N+0.5)^{1+c} - (M-0.5)^{1+c}]$$

Finally:

$$T_{1,N} = T_{1,M} + T_{M+1,N}$$

This method is easily extended to more than two segments. For the additional segments, we would need to compute H_y from H_x, H_z from H_y, etc., using the above process.

Example 28. Your company is considering immediately beginning production of a product that has just completed it final design review, but no production planning has been done. The first unit estimate at the current state of preparedness is 600 hours. A total of 100 units will be built. Although you would normally expect a learning slope of 87%, you believe that is not likely to be achieved in the present situation. You think it more likely that the initial slope will be 95%, and that this will last for the first 20 units, until all the bugs are worked out. Both slopes are based on the U model. Use the segmented method explained above to build a segmented learning curve to estimate this situation.

Answer: We first determine the natural slopes:

$$b = \log(0.95)/\log(2) = -0.074 \text{ (initial slope)}$$

$$1+b = 1-0.074 = 0.926$$

$$c = \log(0.87)/\log(2) = -0.2 \text{ (final slope)}$$

$$1+c = 1-0.2 = 0.8$$

Next, we determine the cumulative hours for the first 20 units, at slope b:

$$T_{1,20} = (600/0.926)(20.5^{0.926} - .5^{0.926}) = 10{,}282 \text{ hours}$$

The projected H_x value is:

$$H_x = 600(20)^{-0.074-(-0.2)} = 875 \text{ hours}$$

Hence:

$$T_{21,100} = (875/0.8)(100.5^{0.8} - 20.5^{0.8}) = 31{,}462 \text{ hours}$$

Finally:

$$T_{1,100} = 10{,}281 + 31{,}462 = 41{,}743 \text{ hours}$$

. . . .

We now develop this method using the CA model. The hours from unit 1 to unit M are:

$$T_{1,M} = H_1 M^b$$

The hours for unit M are:

$$H_M = H_1[M^{1+b} - (M-1)^{1+b}]$$

But also:

$$H_M = H_x[M^{1+c} - (M-1)^{1+c}]$$

Hence:

$$H_x = H_1[M^{1+b} - (M-1)^{1+b}]/[M^{1+c} - (M-1)^{1+c}]$$

Now we can write:

$$T_{M+1,N} = H_x[N^{1+c} - (M-1)^{1+c}]$$

Finally:

$$T_{1,N} + T_{1,M} + T_{M+1,N}$$

An example will solidify the method for the CA model.

Example 29. On its last major project, your company experienced slow learning on the first 100 units of the product. The learning rate was only 98%, based on the CA model. But after 100 units, the rate accelerated quickly to a more normal 90%. You are expecting similar behavior on a new project just beginning to get ready for production. On this new project, H_1 is estimated at 650 hours, and 800

units will be built. Using a segmented learning curve, estimate total hours for all 800 units.

Answer: We need the natural slope of both learning curves. For the 98% curve:

$$b = \log(0.98)/\log(2) = -0.02915$$

$$1+b = 1-0.02915 = 0.97085$$

For the 90% curve:

$$c = \log(0.9)/\log(2) = -0.152$$

$$1+c = 1-0.152 = 0.848$$

The hours from unit 1 to unit 100 are:

$$T_{1,100} = 650(100)^{0.97085} = 56{,}835 \text{ hours}$$

The projected H_x is given by:

$$H_x = 650[100^{0.97085} - 99^{0.97085}]/[100^{0.848} - 99^{0.848}] = 1{,}309.5 \text{ hours}$$

Hence:

$$T_{M+1,N} = 1309.5(800^{0.848} - 101^{0.848}) = 313{,}670 \text{ hours}$$

Finally:

$$T_{1,800} = 56{,}835 + 313{,}670 = 370{,}505 \text{ hours}$$

■■■

X. Interruption of Production

Quite frequently, production is interrupted for various reasons. When this happens, it is almost certain that some learning will be "lost." A loss of learning means that there will be a discontinuity in the learning curve if and when production resumes. It means also that the first and subsequent units produced after the interruption will require more hours than if production had continued uninterrupted.

Manufacturers sometimes claim in negotiations that an interruption in learning warrants a return to the top of the learning curve. While that might be true for a five-year interruption, it rings a bit hollow for a six-month interruption. And it almost certainly is invalid in most cases for a one-month interruption.

What is needed is a rational and supportable way to estimate the learning effects of production breaks. The method we develop and demonstrate here is similar to a method developed by George Andelohr and presented in his 1969 article "What Production Breaks Cost," published in *Industrial Engineering* magazine.

We begin by noting that by interruption of production, we mean a total cessation, however short that may be. However, we also note that the main consideration is not the length of the time interval per se, but the events that happen in that time interval that are of most consequence. Depending on circumstances, an interruption of a few days could have more impact than a delay of weeks. But on the other hand, a long delay provides more opportunity for "bad things" to happen.

A partial design change while production is ongoing is not quite the same as an interruption of production. Typically, the part that is changed is treated as a new product that starts at the top of the learning curve, while the part that is unchanged goes on down the original curve. The exception is if the change requires the entire production operation to shut down. Then, interruption methods would apply.

Andelohr hypothesizes that there are five areas of learning and that these areas may suffer different losses of learning in a given interruption scenario. The overall effective loss of learning is the loss summed across the five areas. It might be reasonable and valid for the reader to define the five areas somewhat differently than Andelohr does, or even to use a number of learning areas other than five. Your author believes that the main consideration is the recognition that learning can be lost in more that one way, and that each of the ways should be considered in a given situation. The five areas of learning suggested by Andelohr are:

- **Personnel learning** – This area considers that the job-peculiar acquired skills of the people doing the work may be lost. Evidently, there are two ways this can happen. One is that people are laid off or transferred to other work and are not available when production resumes. Another is gradual atrophy of the skills due to the passage of time. Even if people from the prior crew are able to return when production resumes, it is reasonable to expect that it will take them some time to get back into the rhythm of the work.
- **Supervisory learning** – Supervisors also can be reassigned or leave the company. They too can need time to get back into the rhythm of the work.
- **Continuity of production** -- This area relates to the physical configuration of the production operation. If all workstations remain intact and the work in process inventory remains in place during the interruption, then this area will have little effect. But if workstations are modified or relocated, or if all work in process inventory has been eliminated, this factor could result in a major loss of learning.
- **Methods** – This area has to do with the methods documents that describe how the work is to be done. If they remain on file and available, and are not revised, then the effect of this area is small. Otherwise, it could cause significant loss of learning.
- **Special tooling** – In a production interruption of any length, it is not unusual for tooling to be warehoused, and sometimes misplaced. Also, an interruption might trigger tooling changes, for example a shift from soft (short term or temporary) to hard (long term or permanent) tooling.

In Andelohr's approach, each of the areas where learning can be lost is assigned a percentage weight, with all of the weights summing to 100%. He suggests that if there is any doubt about how to assign the weights, equal weight should be given to each area. Thus if there are five areas, each would be assigned a weight of 20%. The subsequent analysis is best explained by an example, which we now produce.

Example 30. Production is temporarily stopped in your factory after 20 units of a planned 50 units because of an unexpected funding problem that your customer has experienced. Additional production to complete the remaining 30 units is expected to begin in about three months, at which time the customer is expected to have the funding needed. The theoretical first unit hours are 3,285, and the slope has been 83%, as planned, based on the U model. Your original bid was based on uninterrupted production, and you expect your customer to pay for the added costs due to the interruption.

Because production is expected to resume, efforts have been made to keep the production capability as intact as possible under the circumstances. However,

certain disruptions have been unavoidable. As negotiations with the customer near, you want to be ready with a convincing rationale for the cost of lost learning. You already know that the customer's initial negotiating position is that you should be able to continue down the old learning curve as though there had been no loss of learning. You want to persuade your customer that loss of learning is real, and that he should pay for it.

Answer: You begin by constructing the following table, using Andelohr's scheme for describing elements that contribute to learning:

Elements of Learning	Weight Assigned	Learning Loss
Personnel Learning	40%	2%
Supervisor Learning	20%	20%
Continuity of Production	40%	20%
Methods	0%	0%
Special Tooling	0%	0%

You justify the above learning loss entries as follows:

Personnel learning: You expect that 80% of the trained personnel will be available when production resumes, and you assert that in three months they will have retained 75% of their learning. The estimated learning loss is:

Weight	Personnel Lost	Learning Lost	Weighted Loss
40% x	20% x	25% =	2%

Supervisory learning: Unfortunately, it has been necessary to reassign all supervisors to other tasks. New supervisors will be assigned when production resumes. The learning loss calculation is:

Weight	Personnel Lost	Learning Lost	Weighted Loss
20% x	100% x	100% =	20%

Continuity of production: Unfortunately, half of the workstations have had to be moved. You equate this to 50% of the 40% weight, or 20%

Methods: All methods have been carefully preserved. This area is not expected to cause any loss of learning.

Special tooling: Hard tooling was used in this project from the outset. All tooling is being maintained at the workstations. This area is not expected to cause any loss of learning.

Summed across all five areas, the loss of learning forecast is 42%. But 42% of what? How much learning has there been? We need to be explicit about what we mean by learning, in a quantitative sense. We define it this way: Provided production has been uninterrupted, cumulative learning at unit n in the production sequence is the difference in unit hours between H_1 and H_n. Mathematically expressed:

$$\mathbf{L_n = H_1 - H_n} \hspace{3cm} \text{Eq(20)}$$

In the scenario we have defined, H_1 = 3,285 hours. We will eventually need both b and 1+b:

$$b = \log(0.83)/\log(2) = -0.26882$$

$$1+b = 0.73118$$

Production stopped at unit 20, so we need the hours at that point. We use Eq(2):

$$H_n = H_1 n^b$$

$$H_{20} = 3285(20^{-0.26882}) = 1,468 \text{ hours}$$

Therefore:

$$L_{20} = 3,285 - 1,468 = 1,817 \text{ hours}$$

Of this amount, we say that 42% has been lost. Lost learning is therefore:

$$(42\%)(1,817) = 763 \text{ hours}$$

These are NOT the lost hours you will ask your customer to pay for. You will want him to pay for much more, as will be seen. The reason is that the loss of learning affects every future unit produced, not just the first unit of the new production sequence.

Our approach will be as follows:

Step 1. We calculate what the hours for unit 21 would have been without the interruption.
Step 2. We add the 763 hours of lost learning to that amount. This gives us a new value for H_{21}; call it H_{21+}.

Step 3. Some earlier unit, call it unit x, 1<x<21, will have the same number of hours as H_{21+}. We start the learning curve over with this unit (x is seldom an integer).

Step 4. We compute the cumulative hours from unit x to unit 30+x. These will be the total hours we want the customer to pay for, for the last 30 units.

Step 5. We compare the hours in Step 4 to the hours that would have been required for units 21 through 50 had there not been an interruption. This allows us to identify the cost (in hours) of the interruption.

We proceed.

Step1: Without the interruption, the hours for unit 21 would have been:

$$H_{21} = 3285(21)^{-0.26882} = 1{,}449 \text{ hours}$$

Step 2: We add the 763 hours for lost learning to that amount:

$$H_{21+} = 1{,}449 + 763 = 2{,}212 \text{ hours}$$

Step 3: We need to find an earlier unit, x, that requires 2,212 hours.

$$H_x = 2{,}212 = H_1x^b = 3285x^{-0.26882}$$

Solving for x yields:

$$x = 4.35416$$

For the last 30 units, learning begins again at "unit" 4.35416, and continues for 30 units, that is, to unit 33.35416.

Step 4: We compute the cumulative hours from unit 4.35416 to unit 33.35416. These will be the adjusted hours we expect to spend for the last 30 units. We use Eq(9):

$$T_{4.35416,33.35416} = (3285/0.73118)(33.885416^{0.73118} - 3.85416^{0.73118}) = 46{,}963 \text{ hours}$$

Step 5: The hours we would have spent on the last 30 units sans the interruption are:

$$T_{21,50} = (3285/0.73118)(50.5^{0.73118} - 20.5^{0.73118}) = 38{,}162 \text{ hours}$$

The 46,963 hours obtained in the previous step is an increase of 8,801 hours over the original plan. As this example shows, breaks in production can be expensive.

The approach in the above example is purely analytical. Sometimes people use graphs to make these estimates, but that can introduce a fair amount of error. However, it can be instructive to see what a graph of interrupted production looks like as produced by a spreadsheet. Exhibit 8 below illustrates.

The method illustrated in this example can also be done using the CA model. One need only replace the U model equations with their CA model equivalents. The mathematical effort is about the same for either model.

Exhibit 8 -- Example Plot of Interrupted Learning

Learning starts again at unit determined by amount of learning lost

Production break

Hours

Units Produced
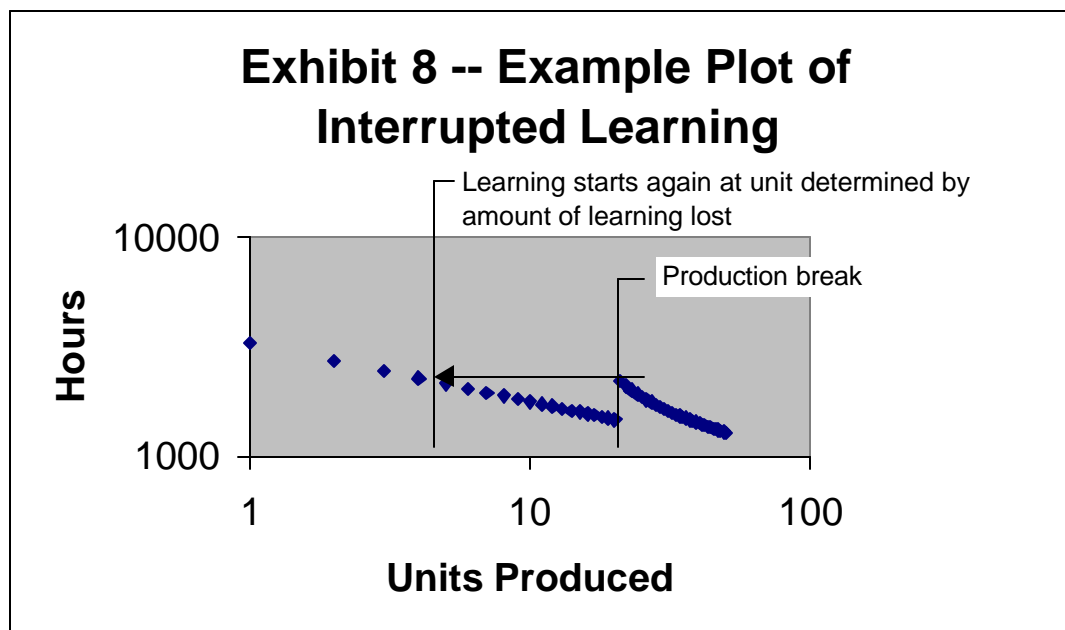
XI. Fitting Learning Curves to Production Data

In new production situations where there is no prior experience, it is difficult to know what slope to assign to a learning curve. Sometimes, in these situations, one can do no better than an educated guess based on general industry experience (see, for example, Section VI). But when one has data, one should make intelligent use of it. Unless there is some compelling reason to believe otherwise, what happened in the last project is the best guide to what will happen in the next project.

Data typically comes in one of two forms. One we call unit data, the other we call block data. Unit data is data that have been collected for each unit of production. For example:

Unit	Hours
1	200
2	185
3	176
4	165
5	150
Etc.	

Block data are data that have been collected for a range of production units. For example:

Units	Hours
1-50	55,260
51-150	98,320
151-200	37,360

Ideally, we would like to be able to always work with unit data, because it contains the most information. Unfortunately, block data is all we are able to get much of the time.

What information do we need to extract from historical data for possible use in future learning curve applications? The information we would most like to have is the slope. The theoretical first unit hours may be irrelevant if we are estimating them from scratch for the new project. However, there are situations where we would like both the slope and the theoretical first unit hours. Therefore, the methods we will develop will always strive to get a good estimate of both.

As you might expect, treatment of data is different for the U and CA models. Typically, if you fit the same data to both models, both the theoretical first unit hours and the slope will be somewhat (but not hugely) different between the two models. This is to be expected, given the fundamentally different nature of the models.

We begin with a simple technique that works for either the U or the CA model when unit data are available, although the implementation is slightly different for the two models. The next example illustrates the technique for the U model; the subsequent example illustrates it for the CA model.

Example 31. Consider the hypothetical unit data shown below. We will fit it to the U model.

Unit	Hours	Unit	Hours	Unit	Hours
1	48.3	11	30.1	21	26.5
2	42.0	12	29.9	22	26.3
3	37.9	13	29.6	23	25.9
4	36.2	14	28.8	24	25.9
5	33.0	15	28.6	25	25.6
6	32.2	16	28.1	26	25.5
7	30.9	17	27.8	27	25.1
8	31.8	18	27.5	28	25.0
9	31.1	19	27.1	28	24.6
10	30.8	20	26.8	30	24.6

Answer: Recall Eq(7):

$$H_{2n}/H_n = S/100$$

This equation says that the ratio of hours at unit $2n$ to hours at unit n is equal to the percentage slope expressed as a decimal. Let's make a table of doublings and calculate the implied value of $S/100$ in a third column. We will average the $S/100$ values at the bottom of the table.

Unit	Hours	S/100
1	48.3	-
2	42.0	0.86957
4	36.2	0.86190
8	31.8	0.87845
16	28.1	0.88365
Average		0.87339

We repeat this for doublings starting at unit 3:

Unit	Hours	S/100
3	37.9	-
6	32.2	0.84960
12	29.9	0.92857
24	25.9	0.86622
Average		0.88146

We repeat again for doublings starting at unit 5:

Unit	Hours	S/100
5	33.0	-
10	30.8	0.9333
20	26.8	0.87013
Average		0.90173

We could continue this process for doublings starting at unit 6, 7, etc., but for the sake of brevity we will not. Usually the accuracy of the fit does not improve much after three or four of these doubling analyses.

We have collected three averages. What we want next is a grand average, and we will do it as a weighted average, with the weights being the number of S/100 values computed in each doubling series.

Weighted average = [(4)(0.87339) + (3)(0.88146) + (2)(0.90173)]/9 = 0.88238

We infer that the percentage slope is 88.238% to a very close approximation.

We could jump to the conclusion that the first unit hours are 48.3, but that value may not be the best fit to this data. Let's use an approach that is likely to give us a much better value. We begin by summing the hours across all 30 units in the

data set we are working with. The result is 893.5 hours. Next, we find the values of b and 1+b corresponding to the value of S we have found.

$$b = \log(0.88238)/\log(2) = -0.18053$$

$$1+b = 0.81947$$

Consider now Eq(9):

$$T_{M,N} = [H_1/(1+b)][(N+.5)^{1+b} - (M-.5)^{1+b}]$$

We know all of the values in this equation except H_1, so we solve for H_1:

$$893.5 = (H_1/0.81947)(30.5^{0.81947} - 0.5^{0.81947})$$

$$H_1 = 46.08 \text{ hours}$$

We have fitted the U model equation:

$$H_n = 46.08n^{-0.18053}$$

to our data set. As an exercise, you might want to verify that the average error over the 30 data points is about 0.02 hours.

∎∎∎

In the next example, we will use a similar method to fit a CA model to the same data set.

Example 32. Recall Eq(12) for the CA model:

$$T_{1,N} = H_1 N^{1+b}$$

We can also write:

$$T_{1,2N} = H_1(2N)^{1+b}$$

We take the ratio of these two equations:

$$T_{1,2N}/T_{1,N} = 2^{1+b}$$

This equation says that the ratio of the *cumulative* hours at any unit 2N to the cumulative hours at unit N is equal to 2^{1+b}. If we convert our data table to a cumulative data table, we can follow a process similar to that in the previous

example to find b. From b, we can find S. Here is the cumulative version of the data set from the previous example:

Unit	Cumulative Hours	Unit	Cumulative Hours	Unit	Cumulative Hours
1	48.3	11	384.3	21	665.0
2	90.3	12	414.2	22	691.3
3	128.2	13	443.8	23	717.2
4	164.4	14	472.6	24	743.1
5	197.4	15	501.2	25	768.7
6	229.6	16	529.3	26	794.2
7	260.5	17	557.1	27	819.3
8	292.3	18	584.6	28	844.3
9	323.4	19	611.7	28	868.9
10	354.2	20	638.5	30	893.5

We now do doubling routines similar to those in the previous example. The first one starts at unit 1.

Unit	Cumulative Hours	2^{1+b}
1	48.3	-
2	90.3	1.86957
4	164.4	1.82060
8	292.3	1.77798
16	529.3	1.81081
Average		1.81974

We repeat this for doublings starting at unit 3:

Unit	Cumulative Hours	2^{1+b}
3	128.2	-
6	229.6	1.79095
12	414.2	1.80401
24	743.1	1.79406
Average		1.79634

We repeat again for doublings starting at unit 5:

Unit	Cumulative Hours	2^{1+b}
5	197.4	-
10	354.2	1.79433
20	638.5	1.80265
Average		1.79849

The weighted average of the averages is:

Weighted average $= [(4)(1.81974) + (3)(1.79634) + (2)(1.79849)]/9 = 1.80722$

Hence we find that:

$$2^{1+b} = 1.80722$$

to a close approximation. Taking logarithms of both sides:

$$1+b = \log(1.80722)/\log(2) = 0.85377$$

$$b = -0.14623$$

Using Eq(6):

$$S = 10^{-0.14623 \log(2) + 2} = 90.36\%$$

If we need H_1, we can use Eq(12) as follows:

$$893.5 = H_1(30)^{0.85377}$$

$$H_1 = 48.97 \text{ hours}$$

As previously noted, it should not be a surprise that the U and CA models give somewhat different H_1 and S values when data are fitted. The models work in different ways.

The fitted CA model is:

$$A_n = 48.97n^{-0.14623}$$

Cumulatively, the fitted equation is:

$$T_{1,N} = 48.97N^{0.85377}$$

···

The average error on a *cumulative* basis over all 30 points is about 4 hours.

. . . .

We next consider the situation where only block data are available. If there is only one block available, there is insufficient information to fit a learning curve using either the U or the CA model. There must be at least two blocks. The case where there are just two blocks is fairly easy to fit using simple algebraic methods combined with a bit of trial and error in the case of the U model. We illustrate it for the U model in an example.

Example 33. Two blocks of production have been completed. There was no loss of learning between the blocks. These were the results. Fit the U model to this data.

Block Units	Block Hours
1-50	2,812
51-150	4,026

Answer: The total hours for both blocks are 6,838 hours. We can write:

$$T_{1,50} = [H_1/(1+b)][50.5^{1+b} - 0.5^{1+b}] = 2,812 \text{ hours}$$

and also:

$$T_{1,150} = [H_1/(1+b)][150.5^{1+b} - 0.5^{1+b}] = 6,838 \text{ hours}$$

These two equations can in theory be solved for the two unknowns b and H_1. However, they are complex and highly nonlinear, so a direct solution is difficult. We take note of the fact that the term 0.5^{1+b} is always a small number, typically in the range 0.5 to 0.7. So to a first approximation, we can neglect it and write the following ratio:

$$T_{1,50}/T_{1,150} = (50.5/150.5)^{1+b} = 0.41123$$

Taking logarithms of both sides and solving for $1+b$ yields:

$$1+b = 0.81375$$

Also:

$$b = -0.18625$$

$$S = 10^{-0.18625\log(2)+2} = 87.89\%$$

This result is an approximation because we neglected the term 0.5^{1+b}. Using the above approximate solution as a starting point, we proceed by trial and error to find that to a very close approximation, $1+b = 0.8002$ and $S = 87.067\%$. If we need H_1, we can use:

$$T_{1,150} = (H_1/0.8002)(150.5^{0.8002} - 0.5^{0.8002}) = 6{,}838 \text{ hours}$$

Solving for H_1, we find that $H_1 = 100.0$ hours.

∙∙∙

Fitting a CA model to two block data is much easier. We do this in the next example.

Example 34. Let's fit a CA model to the same block data we used in the previous example. We repeat the data here for convenience.

Block Units	Block Hours
1-50	2,812
51-150	4,026

Answer: The total hours for both blocks are 6,838 hours. We can write:

$$T_{1,50} = H_1(50)^{1+b} = 2{,}812 \text{ hours}$$

and also:

$$T_{1,150} = H_1(150)^{1+b} = 6{,}838 \text{ hours}$$

Taking the ratio of these two equations:

$$T_{1,50}/T_{1,150} = (50/150)^{1+b} = 0.41123$$

We take the logarithm of both sides and solve for $1+b$:

$$1+b = 0.80883$$

Hence:

$$b = -0.19117$$

$$S = 10^{-0.19117\log(2)+2} = 87.59\%$$

If we need H_1 we can write:

$$T_{1,150} = H_1(150)^{0.80883} = 6{,}838 \text{ hours}$$

Solving for H_1:

$$H_1 = 118.8 \text{ hours}$$

....

When we have more than two blocks, the best recourse for most analysts is generally linear regression, also called ordinary least squares.[2] Linear regression is a method for fitting linear equations of the form $y = ax + b$ to a set of x,y data pairs. The fit is optimum, in the sense that the sum of the squared distances from the fit line to the data is minimized. We will use it here to fit to block data, but it can also be used to fit to unit data. We will not delve into the theory behind linear regression. There are many excellent textbooks that deal with the subject in detail. The solutions obtained in the following examples were obtained using the regression feature of an MS Excel spreadsheet.

Example 35. In this example, we will assume that three blocks of production have been completed with no loss of learning between blocks. We want to fit a U model to these data. The data are as follows.

Block Units	Block Hours	Average Unit Hours
1-25	13,412	536.48
26-50	8,805	352.20
51-85	10,250	292.86
Total	32,467	

The column on the right is the block hours divided by the number of units in the block.

Eq(2) defines the basic U model:

$$H_n = H_1 n^b$$

---

[2] It is well known that linear regression introduces some bias into learning curve fits, however this effect is typically small. If you have access to more sophisticated optimization methods, this bias can be avoided. Since many analysts do not have access to such methods, but do have access to ordinary least squares (typically via their spreadsheets, or even pocket calculators), we present the linear regression approach in this paper.

We take the logarithm of both sides of this equation:

$$\log(H_n) = \log(H_1) + b\log(n)$$

We make the following substitutions in this equation:

$$y = \log(H_n)$$
$$a = \log(H_1)$$
$$x = \log(n)$$

The result is:

$$y = a + bx$$

This linear transformation of Eq(2) can be fitted to the data using linear regression. We first, however, must do some data manipulation.

We need an "average" unit number to go with the average block hours in the rightmost column in the table above. A way to do that was presented in Section III, called the midpoint method. Unfortunately, to find the midpoint, we must first know the slope, and that is what we are trying to find. Surprisingly, as shall be seen, the way out of this dilemma is to simply guess at the slope! Let's arbitrarily guess at 85%, the midpoint of the range of slopes that typically occur in practice (70%-100%).

For this assumed slope:

$$b = \log(0.85)/\log(2) = -0.23447$$

$$1+b = 0.76553$$

Recall Eq(10), the midpoint formula:

$$K = \{[(N+0.5)^{1+b} - (M-0.5)^{1+b}]/[(1+b)(N-M+1)]\}^{1/b}$$

The midpoints, by block, using the assumed slope, are:

Block Units	Midpoint Unit K w/ 85% Slope
1-25	9.427
26-50	37.16
51-85	67.08

Now we can tabulate units versus average hours:

Units (Midpoints)	Hours
9.427	536.48
37.16	352.20
67.08	292.86

Next, we take the logarithms of the numbers in the above table.

x	y
0.9743	2.7296
1.5701	2.5468
1.8266	2.4667

Doing linear regression on the numbers in this table results in:

$$y = 3.0300 - 0.30817x$$

This gives a new tentative value of the slope, namely –0.30817, equivalent to $S = 80.77\%$. The tentative value of H_1 is the antilog of 3.0300, namely 1,072 hours.

Note that we made a guess at slope and used that guess to compute midpoints. Then we did regression using the midpoints and came up with a new slope. By repeating this process, we can get an even better guess of both slope and first unit hours. We use the new value of slope to recalculate the midpoints, and we repeat the process.

Block Units	Midpoint Unit K w/ 80.77% Slope
1-25	9.0254
26-50	37.0790
51-85	67.0020

Units (Midpoints)	Hours
9.0254	536.48
37.0790	352.20
67.0020	292.86

Next, we take the logarithms of the numbers in this table.

x	y
0.95547	2.7296
1.56913	2.5468
1.82609	2.4667

Doing linear regression on the numbers in this table results in:

$$y = 3.01179 - 0.30122x$$

This gives a new tentative value of the slope, namely –0.30122, equivalent to S = 81.16%. The tentative value of H_1 is the antilog of 3.01179, namely 1,042 hours.

We are homing in on both slope and first unit hours. We will iterate one more time to see if there is much change.

Block Units	Midpoint Unit K w/ 81.16% Slope
1-25	9.0518
26-50	37.084
51-85	67.007

Units (Midpoints)	Hours
9.0518	536.48
37.084	352.20
67.007	292.86

Next, we take the logarithms of the numbers in this table.

x	y
0.95673	2.7296
1.56919	2.5468
1.82612	2.4667

The regression result is:

$$y = 3.01868 - 0.30168x$$

We will accept as final that b = -0.30168, equivalent to S = 81.13%. Also, we accept as final H_1 = antilog(3.01868) = 1,044 hours.

▪▪

We now show how a CA model can be fitted to block data. This is much easier than fitting block data to the U model. We give an example to demonstrate the process.

Example 36. Let's use the same block data we used in the previous example. We repeat it here for convenience.

Block Units	Block Hours	Average Unit Hours
1-25	13,412	536.48
26-50	8,805	352.20
51-85	10,250	292.86
Total	32,467	

Answer: We will make use of Eq(12):

$$T_{1,N} = A_N N = H_1 N^{1+b}$$

We transform the above table to a table of cumulative hours:

Unit	Cumulative Hours
25	13,412
50	22,217
85	32,467

We next transform Eq(12) into linear form by taking the logarithm of both sides:

$$\log(T_{1,N}) = \log(H_1) + (1+b)\log(N)$$

We make the following substitutions:

$$y = \log(T_{1,N})$$
$$a = \log(H_1)$$
$$x = \log(N)$$

The result is:

$$y = a + (1+b)x$$

We next take the logarithms of the data (base 10):

x	y
1.39794	4.12749
1.69897	4.34669
1.92942	4.5114

The regression result is:

$$y = 3.117692 + 0.722703x$$

We find that $b = 0.722703 - 1 = -0.2773$, equivalent to $S = 82.51\%$. Also, $H_1 = $ antilog(3.117692) = 1,311 hours.

■■

XII. Tradeoff Analysis with Learning Curves

Learning curves are a powerful analytical tool in many situations where tradeoffs are made between competing production processes, materials, product designs, or project options. There is virtually no limit to the variety of such situations, so this paper cannot hope to capture all of them. What can be done, and your author believes he has, is present some examples that are rich enough in variety to give readers at least a starting point in most problems that arise in practice.

Example 37. A tradeoff so common it has become a cliché is the tradeoff between machining and casting a metal part. Machining characteristically has a low front-end cost—some minor tooling and setup costs—but a relatively high first unit cost, and usually a relatively steep learning curve. Casting, on the other hand, typically has a high front-end cost—molds must be made—and a relatively low first unit cost. The learning curve tends to be shallower than for machining. The labor rates may also be different for the two operations.

To make the example explicit, assume that the following information has been developed for the machined version of the part, based on the CA model.

Front-end cost	$2,500
Front-end cost	$2,500
First unit hours (H_1)	85
Labor rate	80 $/hour
Learning slope	87% (b= -0.2;1+b = 0.8)

Assume further that the following information has been developed for the cast version of the part.

Front-end cost	$15,500
First unit hours (H_1)	18
Labor rate	65 $/hour
Learning slope	95% (b= -0.074;1+b = 0.826)

Suppose that the production quantity has not yet been specified. We are interested in determining the production quantity at which each process has the same total cost, i.e., the "breakeven" quantity. It appears that below the breakeven quantity we should choose machining, and above that quantity, we should choose casting. In this example, we will use a plot to locate the breakeven quantity. The plot is of the two equations:

$$\$TC_{machined} = \$2500 + \$80(85n^{0.8})$$

XII. Tradeoff Analysis with Learning Curves

Learning curves are a powerful analytical tool in many situations where tradeoffs are made between competing production processes, materials, product designs, or project options. There is virtually no limit to the variety of such situations, so this paper cannot hope to capture all of them. What can be done, and your author believes he has, is present some examples that are rich enough in variety to give readers at least a starting point in most problems that arise in practice.

Example 37. A tradeoff so common it has become a cliché is the tradeoff between machining and casting a metal part. Machining characteristically has a low front-end cost—some minor tooling and setup costs—but a relatively high first unit cost, and usually a relatively steep learning curve. Casting, on the other hand, typically has a high front-end cost—molds must be made—and a relatively low first unit cost. The learning curve tends to be shallower than for machining. The labor rates may also be different for the two operations.

To make the example explicit, assume that the following information has been developed for the machined version of the part, based on the CA model.

Front-end cost	$2,500
First unit hours (H_1)	85
Labor rate	80 $/hour
Learning slope	87% (b= -0.2;1+b = 0.8)

Assume further that the following information has been developed for the cast version of the part.

Front-end cost	$15,500
First unit hours (H_1)	18
Labor rate	65 $/hour
Learning slope	95% (b= -0.074;1+b = 0.826)

Suppose that the production quantity has not yet been specified. We are interested in determining the production quantity at which each process has the same total cost, i.e., the "breakeven" quantity. It appears that below the breakeven quantity we should choose machining, and above that quantity, we should choose casting. In this example, we will use a plot to locate the breakeven quantity. The plot is of the two equations:

$$\$TC_{machined} = \$2500 + \$80(85n^{0.8})$$

$$\$TC_{cast} = \$15500 + \$65(18n^{0.826})$$

In these equations, n is the unknown quantity at which the costs are equal. By plotting the two curves for many values of n, we can identify this quantity. The plot is shown in Exhibit 9 below.
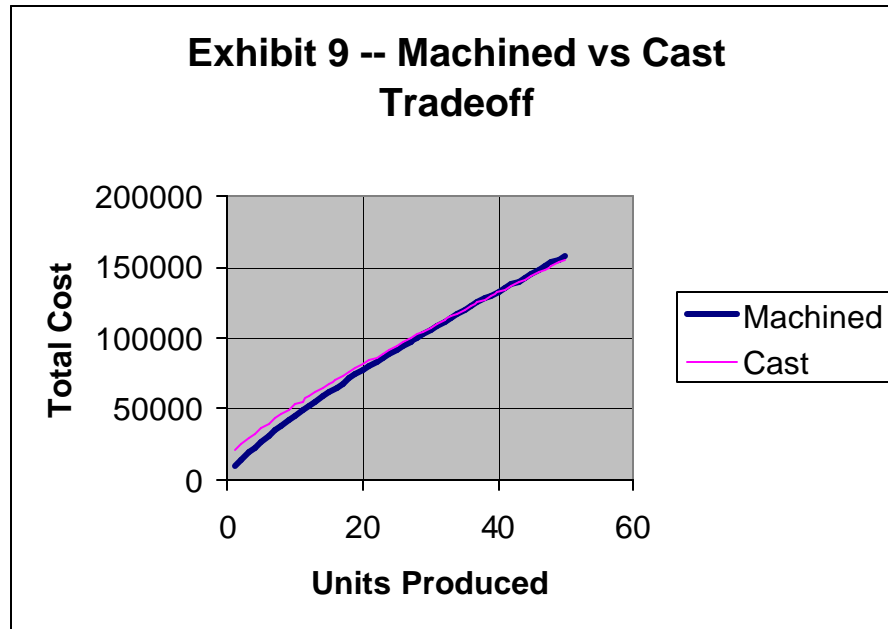
In this particular case, for quantities of 37 or more, castings are cheaper. (But note: This does not necessarily mean that even prototypes should use castings. Once a commitment is made to use a casting, a design change could mean that the mold has to be scrapped. The prudent course may be to use machined parts for the first few units until there is confidence that the design has stabilized.)

Example 38. You are trying to get a product to market as soon as possible to beat your competitors. You have two competing designs for the product, A and B. You need to quickly choose between them.

Based on market studies, you believe the product must be priced at not more than $9,500. If you can price at that level or below, you believe you can sell 15,000 units. You expect a profit of not less than 10%, or $950 average per unit. Therefore your total cost to produce and distribute must not exceed $8,550 average per unit. You have estimated a distribution cost of $1,850 per unit, for either Design A or Design B. Your cost of production therefore must not exceed $6,700 average per unit.

You define time to market as the time it will take to get the first 1,000 units ready to ship. Front-end costs of either design will be amortized over the full 15,000

units. Your decision rules for selecting which design to take to market are, choose the one that can get to market first, unless it violates the target cost. But if only one design does not violate the target cost, go with that one. If both designs violate the target cost, go with neither.

Once the chosen design goes to production, you will have a team of 80 workers who will each devote an effective 35 work hours per week to building the product. You will therefore have 80 x 35 = 2,800 productive hours per week. Your factory burdened labor rate is $75/hour.

You have developed the following information for Design A:

Front-end cost	$1,550,000
Front-end time	12 weeks
First unit hours (H_1; in $)	$8,250
Labor learning slope (CA)	90% (b = -0.152; 1+b = 0.848)
First unit material cost	$6,565
Material learning slope (CA)	96% (b = -0.05889; 1+b = 0.9411)

You have developed the following information for Design B:

Front-end cost	$1,830,000
Front-end time	15 weeks
First unit hours (H_1; in $)	$7,350
Labor learning slope (CA)	92% (b = -0.12029; 1+b = 0.8797)
First unit material cost	$7,485
Material learning slope (CA)	96% (b = -0.05889; 1+b = 0.9411)

Determine which design, if any, to pursue.

Step 1: Compute the average total cost of units over the entire 15,000 units.
Step 2: Compute the time to market for each product.
Step 3: Based on the decision rule previously discussed, which of the designs, if any, will be produced?

Answer:

Step 1: For Design A:

$$TC_A = \$1,550,000 + \$8,250(15,000)^{0.848} + \$6,565(15,000)^{0.9411} = \$86,135,811$$

The average total cost per unit for Design A is:

$$(\$86,135,811)/(15,000) = \$5,742$$

For Design B:

$$TC_B = \$1,830,000 + \$7,350(15,000)^{0.8797} + \$7,485(15,000)^{0.9411} = \$100,228,483$$

The average total cost per unit for Design B is:

$$(\$100,228,483)/(15,000) = \$6,682$$

Costs for Designs A and B are both under the $6,700 target. The choice between them will depend on which can be first to market.

Step 2: For Design A:

You have determined that "ready for market" means that the first 1,000 units are ready to ship. The first unit labor cost for Design A is $8,250. Since the labor rate is $75/hour, the first unit hours are $(\$8,250)/(75\$/hour) = 110$ hours. The hours required to build the first 1,000 units are given by:

$$T_{1,1000} = 110(1000)^{0.848} = 38,494 \text{ hours}$$

As already noted, the factory of 80 people generates 2,800 productive hours per week. Therefore it will require $(38,494)/(2,800) = 13.7$ weeks to produce the first 1,000 units. To this must be added the front-end time of 12 weeks, for a total of $13.7 + 12 = 25.7$ weeks.

For Design B:

The first unit labor cost for Design B is $7,350. Since the labor rate is $75/hour, the first unit hours are $(\$7,350)(75\$/hour) = 98$ hours. The hours required to build the first 1,000 units are given by:

$$T_{1,1000} = 98(1000)^{0.8797} = 42,690 \text{ hours}$$

As already noted, the factory of 80 people generates 2,800 productive hours per week. Therefore it will require $(42,690)/(2,800) = 15.2$ weeks to produce the first 1,000 units. To this must be added the front-end time of 15 weeks, for a total of $15.2 + 15 = 30.2$ weeks.

Step 3: Both designs meet the target cost, but Design A can be in the market 4.5 weeks sooner than Design B. Incidentally, Design A also has a lower total cost by $940 per unit. The best choice is clearly Design A.

■■■

The next example uses the segmented learning technique discussed in Section IX. The idea behind this example is that when there is a rush to get into production, learning may be relatively slow, increasing production costs. But if more time and money are spent on getting ready for production, normal learning can be achieved from the very beginning of production. There are many aspects to such a tradeoff that learning curves cannot help decide, such as the advantage to starting production quickly in order to get to market first. But, as the following example illustrates, learning curves can help trade off the production cost issues.

Example 39. Your company has been planning to go into production of its new Gizmo in three months, but it fears that by waiting three months until all planning has been done and all of the manufacturing processes have been thoroughly proven, it may lose competitive advantage in the market. So it is considering starting production now. You have been asked to do an analysis of the relative costs. The planned production quantity is 1,000 units

After careful analysis and consultation with your manufacturing experts, you conclude the following:

Option 1: Begin production in three months

First unit hours (H_1)	30
Learning slope (U model)	87% (b = -0.2; 1+b = 0.8)

Option 2: Begin production now

First unit hours (H_1)	35
Initial learning slope (U model)	95% (b = -0.074; 1+b = 0.926)
Transition quantity	100 units
Final learning slope (U model)	87% (b = -0.2; 1+b = 0.8)

Answer: Option 1 is easily and quickly analyzed:

$$T_{1,1000} = (30/0.8)(1000.5^{0.8} - 0.5^{0.8}) = 9{,}402 \text{ hours}$$

In option 2, for the first 100 units we have:

$$T_{1,100} = (35/0.926)(100.5^{0.926} - 0.5^{0.926}) = 2{,}681 \text{ hours}$$

The projected first unit hours H_x (see Section IX) for the second learning segment are given by:

$$H_x = 35(100)^{-0.074-(-0.2)} = 62.5 \text{ hours}$$

Now we can write:

$$T_{101,1000} = (62.5)(1000.5^{0.8} - 100.5^{0.8}) = 13{,}207 \text{ hours}$$

Finally:

$$T_{1,1000} = 2{,}681 + 13{,}207 = 15{,}888 \text{ hours}$$

This is more than 169% of the hours that result from waiting three months. However, there may be some offsets that make this penalty less severe than it seems. For example, by getting to market first, you may be able to charge a much higher price.

XIII. Local Rate of Learning

In log-log coordinates, a learning curve plots as a straight line, whether it is the U model or the CA model. But actual learning is not a straight line. It is steepest at the beginning of production, and thereafter continuously becomes less steep. This accords with the intuitive notion that when you are learning something new, you make the most progress when you first take up the subject, but later you have less and less to learn, so learning slows down.

Sometimes it is of interest to know the local rate of learning. (We will explain what that means in a moment.) The differential calculus is an excellent tool for this purpose. If your calculus skills are rusty or never existed, you may not understand parts of the subsequent discussion, but that won't keep you from using the results.

Consider first the U model, as given by Eq(2):

$$H_n = H_1 n^b$$

Let us write this as:

$$H_x = H_1 x^b$$

Here, the intent is to transform an essentially discrete relationship into a continuous one. This will introduce some error, but generally the error will be small.

If we differentiate (a calculus term, in case you didn't know) this relationship with respect to x, we can write the result approximately as:

$$\mathbf{DH_x = bH_1 x^{\,b\text{-}1} Dx} \qquad\qquad \text{Eq(21)}$$

This says that for a small change in x, Δx, the change in H_x is equal to $bH_1 x^{b-1}\Delta x$. Let's do an example to illustrate the use of this result.

Example 40. You have just finished unit 80 of a 300 unit production lot. You know that the unit slope is 87%, b = -0.2. You want to know how much the unit hours should drop between units 80 and 81. Let $\Delta x = 1$, and assume that $H_1 = 370$ hours.

Answer: We will get a more accurate answer if we let x be midway between unit 80 and unit 81.

$$\Delta H_{80.5} = -0.2(370)(80.5)^{-1.2} = -0.38 \text{ hours}$$

We can check this by using Eq(2) to compute both H_{80} and H_{81} and noting the difference:

$$H_{80} = 370(80)^{-0.2} = 154.02$$

$$H_{81} = 370(81)^{-0.2} = 153.64$$

The difference is –0.38 hours.

∎∎

We can do the same thing with the CA model. We begin with Eq(14):

$$H_n = H_1[n^{1+b} - (n-1)^{1+b}]$$

Replace n with x and differentiate to get, approximately:

$$\textbf{DH}_x = [(1+b)\textbf{H}_1][x^b - (x-1)^b]\textbf{Dx} \qquad \text{Eq(22)}$$

We illustrate this equation with an example.

Example 41. Let's use the numbers from the previous example, but assume the CA model rather than the U model. Again, we get a more accurate answer if we let x be midway between 80 and 81.

Answer:

$$\Delta H_{80.5} = [(0.8)(370)][80.5^{-0.2} - 79.5^{-0.2}][1] = -0.31 \text{ hours}$$

As an exercise, you might want to check this using Eq(14) twice, once for $x = 80$ and once for $x = 81$, and taking the difference.

∎∎

Appendix – Collected Formulas

This appendix collects in one place all of the 22 learning curve formulas developed in this paper. Throughout the paper, the formulas are sequentially numbered: Eq(1), Eq(2), etc., as they are developed. They are presented in the same order here, together with brief explanations of their meaning and significance.

$$y = ax^b \qquad\qquad \text{Eq(1)}$$

This is the power law model that underlies both the U and the CA models. With rare exceptions, in learning curve applications b is a negative number with absolute value less that one. b is called the natural slope of the learning curve.

$$H_n = H_1 n^b \qquad\qquad \text{Eq(2)}$$

This is the basic equation of the U model. H_n is the number of hours for unit n; H_1 is the number of units for unit 1.

$$H_n = H_m (n/m)^b \qquad\qquad \text{Eq(3)}$$

This equation of the U model allows an estimate for the hours of some unit n, if the hours are known for any other unit m.

$$H_{2n}/H_n = 2^b \qquad\qquad \text{Eq(4)}$$

This U model equation says that the ratio of the hours required for a production quantity 2n (two time n) to the hours required for unit n is equal to 2^b.

$$b = \log(S/100)/\log(2) \qquad\qquad \text{Eq(5)}$$

This equation relates the natural slope of a learning curve b to the percentage slope S. It is valid for both the U and the CA models. Logarithms may be to any base.

$$S = 10^{b \log(2)+2} \qquad\qquad \text{Eq(6)}$$

This is the inverse of Eq(5). It gives the percentage slope S in terms of the natural slope b. It is valid for both the U and the CA models. The logarithm is to base 10.

$$H_{2n}/H_n = S/100 \qquad\qquad \text{Eq(7)}$$

This, like Eq(4), is a ratio of hours between a production quantity 2n (twice n) and n in the U model. It is a consequence of the relationship between S and b.

$$T_{M,N} = H_1[M^b + (M+1)^b + (M+2)^b + ... + N^b] \qquad \text{Eq(8)}$$

This is the exact cumulative relationship for the U model. $T_{M,N}$ is the total hours required for units M through N, inclusive, N>M.

$$T_{M,N} = [H_1/(1+b)][(N+0.5)^{1+b} - (M-0.5)^{1+b}] \qquad \text{Eq(9)}$$

This is a close approximation to Eq(8), frequently used to avoid having to sum all of the terms of Eq(8).

$$K = \{[(N+0.5)^{1+b} - (M-0.5)^{1+b}]/[(1+b)(N-M+1)]\}^{1/b} \qquad \text{Eq(10)}$$

In the U model, K is the so-called midpoint or average unit between units M and N, N>M.

$$A_n = H_1 n^b \qquad \text{Eq(11)}$$

This is the basic equation of the CA model. A_n is the average hours for the first n units produced. H_1 is the hours for the first unit, and b is the natural slope. In the CA model, b plays a different role than it does in the U model.

$$T_{1,N} = A_N N = H_1 N^{1+b} \qquad \text{Eq(12)}$$

This equation expresses the total hours for the first N units in the CA model.

$$T_{M,N} = T_{1,N} - T_{1,M-1} = H_1[N^{1+b} - (M-1)^{1+b}] \qquad \text{Eq(13)}$$

This equation expresses the total hours for all units M to N inclusive in the CA model, N>M.

$$H_n = H_1[n^{1+b} - (n-1)^{1+b}] \qquad \text{Eq(14)}$$

In the CA model, the hours for unit n in terms of unit 1 is given by this equation.

$$H_n = H_m[n^{1+b} - (n-1)^{1+b}]/[m^{1+b} - (m-1)^{1+b}] \qquad \text{Eq(15)}$$

This equation gives the hours for any unit in the CA model in terms of the known hours for any other unit. It is the equivalent of Eq(3) for the U model.

$$R = N^{1+c}/N^{1+b} = N^{c-b} \qquad \text{Eq(16)}$$

R is a measure of the error when natural slope c is erroneously chosen when b should have been chosen. R is the ratio of your actual cumulative result based on the realized slope c to the hours predicted by erroneously using slope b in calculations. R is a function of production quantity N and the slope difference c-b. While this equation is derived from the CA model, it also works well for the U model.

$$R = N^{0.017dS} \qquad\qquad Eq(17)$$

This is an alternate view of the ratio R of Eq(15). In Eq(16), R was presented in terms of N, b, and c. Here, the difference c-b is approximately represented in terms of the error in percent slope, δS. δS should be expressed in percentage points, for example, if the correct slope is 88% and 86% was used in computations, then $\delta S = 2$.

$$T_{1,Nall} = H_{1all}N^B \qquad\qquad Eq(18)$$

This is an approximate aggregate cumulative learning curve formula in power law format. $T_{1,Nall}$ is the sum of all contributing hours from all departments for all production unit from 1 to N. H_{1all} is the sum of all first unit hours from all departments. The next equation shows how to calculate the exponent B. This equation is exact at unit 1 and at unit N. The error is usually small, especially in the vicinity of N.

$$B = [\log(T_{1,Nall}) - \log(H_{1all})]/\log(N) \qquad\qquad Eq(19)$$

This equation permits computation of B for use in Eq(18). It results from solving that equation for B.

$$L_n = H_1 - H_n \qquad\qquad Eq(20)$$

This equation defines learning at unit n as the difference between the theoretical hours for unit 1, and the theoretical hours for unit n.

$$DH_x = bH_1x^{b-1}Dx \qquad\qquad Eq(21)$$

This equation gives the approximate difference in unit hours in the U model between a unit "x" and another unit Δx different from x. The formula is not recommended for x larger than 2 or 3.

$$DH_x = [(1+b)][x^b - (x-1)^b]Dx \qquad\qquad Eq(22)$$

This equation gives the approximate difference in unit hours in the CA model between a unit "x" and another unit Δx different from x. The formula is not recommended for x larger than 2 or 3.

If you have questions about this paper, you may contact the author, Evin Stump, at:

Voice: 310-414-3222 x628
e-mail: estump@galorath.com