Projecting Future Costs with Improvement Curves: Perils and Pitfalls

Brent M. Johnstone

Abstract: Improvement curves are one of the most common projection tools used by cost estimators. Their use is surrounded however by perils and pitfalls. Common errors include: the fallacy of "straight edge and graph paper" projection, the dangers of recovery slopes, failure to understand how development and production environments differ, and the dangers of using learning curve slopes to measure production line efficiency. This paper examines these potential pitfalls and proposes ways to avoid them.

Introduction

Improvement curves are one of the most common tools that cost estimators use to project future costs. Unlike a ladder or power tool bought at the hardware store, improvement curves do not come with warning labels. Perhaps they should: the consequences of misusing them can be quite significant. The stakes of a bad cost estimate can be high – millions or even billions of dollars in funding or profits may depend on decisions estimators make. Consider this paper in some sense a warning label: it identifies the perils and pitfalls of improvement curves and looks at common errors in projecting future costs based on the author's experience in the military aircraft industry.

This paper will examine five potential perils:

- the peril of straight-line projection
- failure to account for the impacts of development versus production
- the dangers of recovery slopes
- carelessness about designating the first unit
- dangers of using learning curve slopes to measure production line efficiency

Peril: The Straight-Line Projection

A common method of projection using the learning curve is to regress historical data, calculate the curve slope, then assume that same slope to project the cost of future work. "You are on an 83% learning curve," the analyst announces as if he is stating an inviolable law of nature. "You should be on the same slope for future lots." Proof that this slope is valid for future projection is typically buttressed by a statement of the regression line's R^2 – the higher the R^2 the "better" the model and the more certain the future projection. This can be called the "straight edge and graph paper" school of estimating – projecting the future is no more difficult than drawing a best fit line on log-log paper and projecting that line through the number of units being estimated.

What could be wrong with this? Empirical studies have demonstrated that this is in fact is not a reliable method to project future costs. Dutton (1984) cautioned:

> In general, the empirical findings caution against simplistic uses of either industry experience curves or a firm's own progress curves. Predicting future progress rates from past historical patterns has proved unreliable.

Similarly, Fox, et al. (2008) cited:

> Even with both an excellent fit to historical data (as measured by metrics like R^2), and meeting almost all of the theoretical requirements of cost improvement, there is no

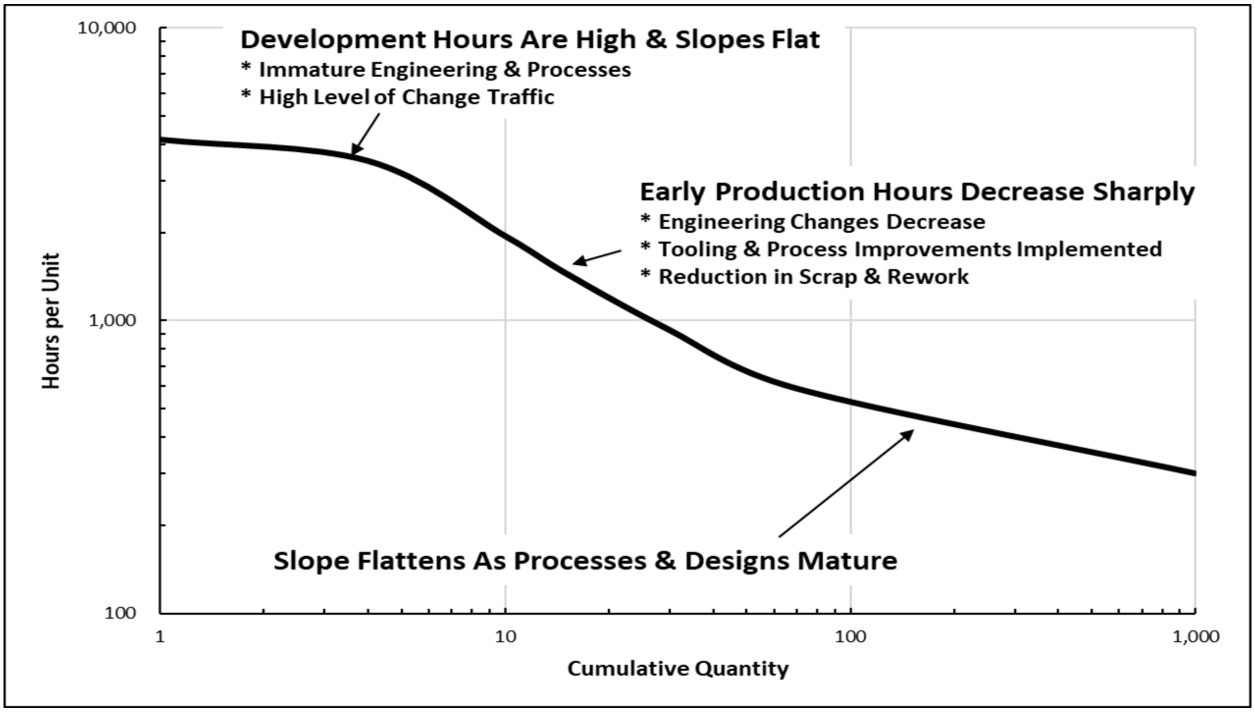

Figure 1. Profile of the S-Curve (Notional)

> guarantee of accurate prediction of future costs….[E]ven projections based on producing an almost identical product over all lots, in a single facility, with large lot sizes, and no production break or design changes, do not necessarily yield reliable forecasts of labor hours.

Continuing, Fox, et al. writes:

> Out-of-sample forecasting using early lots to predict later lots has shown that, even under optimal conditions, labor improvement curve analyses have error rates of about +/- 25 percent.

The primary reason for this failure is that the learning curve is frequently not a straight line in log-log space over the product life cycle. The initial learning curve studies (Wright, 1936; Crawford, 1944) understood improvement curves as straight-line logarithmic functions. Within a few years, however, observers began to see improvement curves not as straight lines in a log-log space, but curvilinear functions that exhibited an "S" shape based on product and process maturity (Carr, 1946; Stanford Research Institute, 1949; Asher, 1956; Cochrane, 1960; Cochrane, 1968).

The S-shaped improvement curve as commonly drawn is composed of three stages, captured graphically in Figure 1 (Carr, 1949; Cochrane, 1960; Cochrane, 1968).

The first stage, typically in the product development phase, shows high hours per unit and relatively flat improvement curve slopes. The limited degree of improvement is caused by an evolving engineering design and immature manufacturing processes. Part shortages disrupt the continuity of production. Scrap and rework is high, and there are typically a high number of engineering changes.

In the second stage, typically during early production, the hours per unit decrease sharply along a relatively steep improvement curve. The production rate increases significantly from the relatively low delivery rates of the development phase. Engineering changes decrease sharply, while improvements in tooling and manufacturing processes are implemented. Manufacturing scrap and rework also decreases at a faster rate. Shortages decrease as the supply chain begins efficiently feeding the production line.

In the third stage, production rates continue to increase to their maximum build rate. Manufacturing processes, tooling and engineering designs mature. Consequently, the pace of production improvements slow and the learning curve slope flattens in response. (Boone, 2021)

The easiest way to understand the changing curve slope over time is to understand the definition of the learning curve itself. A Northrop publication from the 1960's defines the learning curve as "the rate at which management identifies and solves problems in relation to design, methods, shortage of parts, inspection and shop education." (Jones, 2001) Logically, the rate at which problems are solved will change over time – the "low hanging fruit" with the fastest payoffs will be picked first, leaving the more intractable and difficult problems to be solved later, or maybe not at all.

What is the significance for our estimator? If he does not consider where he is in the product life cycle but blindly continues the historical slope, he may significantly overstate or understate future hours. (Reference Figure 2.) If his history is from the initial development stage, he may miss the steepening which typically occurs in the early production stage and overstate his estimate. If his history is from the early production stage, he may miss the flattening that occurs as product designs and manufacturing processes mature and understate his estimate. This does not mean that the analyst should never project a historical slope forward. Suppose the program has reached full production and its engineering and manufacturing processes are mature. In such a case it might be appropriate to project the next production lot by continuing the historical slope. But these decisions cannot be made carelessly

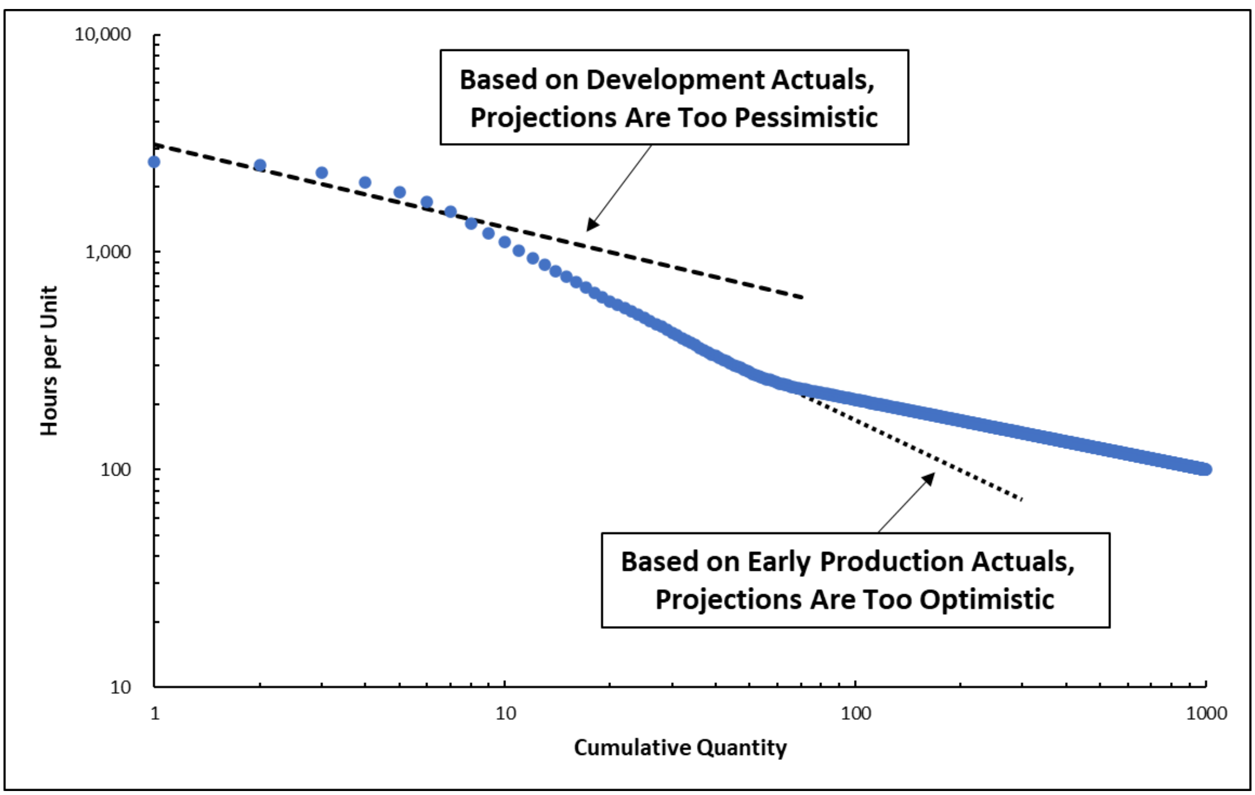

Figure 2. S-Curve & Impact on Projections (Notional)

without understanding where historical experience falls along the product life cycle.

But what about those sterling best fit statistics our analyst quoted earlier? True, many writers on learning curves recommend using best fit statistics as a criterion for choosing a particular learning curve slope and theory -- but as *a* criterion, not the *sole* criterion. Nussbaum and Mislick (2015) introduce numerous factors which should be considered in determining learning curve slopes including the nature and quality of production tooling; supplier competence and experience; expected number of design changes; length of part lead times; similarity of the product to other systems; and historical experience across the product lifecycle. These factors can and do change over the course of a production program.

Using R^2 blindly to justify continuing a straight-line projection – on the basis that past is prologue – recalls the metaphor of driving a car by only looking through the rear-view mirror. Schumeli (2010) distinguishes sharply between explanatory models and predictive power. R^2 is a statistic which explains the historical association between the variables of a model. It can make no justifiable claim about the future. As Schumeli notes, models which do a good job of explaining *observed* behavior may do a poor job of predicting *future* behavior.

Continuing on this theme, Schumeli writes:

> Researchers report R^2-type values and statistical significance of overall F-type statistics to indicate the level of explanatory power. …A common misconception in various scientific fields is that predictive power can be inferred from explanatory power. However, the two are different and should be assessed separately. … Measures such as R^2 and F would indicate the level of association, but not causation. …In general, measures computed from the data to which the model was fitted tend to be overoptimistic in terms of predictive accuracy: "Testing the procedure on the data that gave it birth is almost certain to overestimate performance." (Mosteller and Tukey, 1977)

Regardless of the historical R^2, if a regression model ignores product and manufacturing maturity and their associated cost impacts, it will not do a good job of predicting the future.

Solution: Using Multiple-Leg Curves Prevents "Straight Edge" Fallacy

Without actual cost history, analogous program data combined with analysis of the programmatic factors previously referenced by Nussbaum and Mislick can be used to derive the projected learning curve slopes and breakpoints to project a S-shaped improvement curve. My earlier paper on improvement curves (Johnstone, 2015) suggests a methodology for early production when there are limited actual cost history. In this instance, let us assume there is sufficient historical data on the program in question, and a change in slopes can be inferred from a visual inspection of the data.

There are several learning curve models which allow an S-shaped improvement curve to be derived (Miller, 1971; Jones, 2001). This paper suggests a discontinuous regression model which can be easily built from historical data.

We start from our familiar improvement curve model:

$$y = \alpha_1 x^{\beta_1} \qquad (1)$$

Where:

y = Manufacturing hours per unit

x = Cumulative units built to date

α_1 = Y-intercept, equal to theoretical first unit (TFU) hours

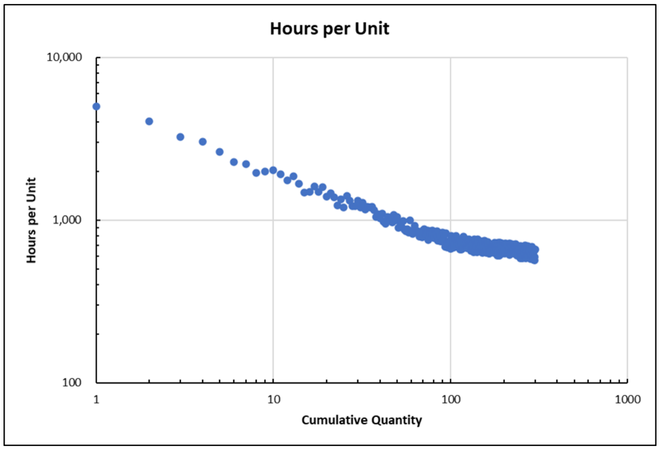

Figure 3. Notional Data Set

β_1 = Rate of learning, such that 2^β equals learning curve slope

After hours per unit and cumulative quantities are converted to natural logarithms, this yields the following linear form:

$$\ln y = \ln \alpha_1 + \beta_1 \ln x \qquad (2)$$

Kennedy (1992) outlines a method for using dummy variables to capture a change in the intercept and slope coefficients between two periods. To create a two-leg segmented learning curve, we introduce breakpoint unit *T*. Based on our *a priori* selection for *T*, data is separated into pre-break period 1 ($x < T$) and post-break period 2 ($x \geq T$). In addition, dummy variable *D* is created such that *D* is zero for period 1, and one for period 2. Product dummy variable *Dx* is also created such that *Dx* takes the value *x* in period 2 but is 0 otherwise. This creates the regression equation:

$$\ln y = \ln(\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \ln Dx \qquad (3)$$

Equation (3) represents two separate cases. Where $x < T$, variables *D* and *Dx* are 0 and equation (3) reduces back to our standard improvement curve equation (2). But where $x \geq T$ and *D* takes the value of one, different intercept and slope values are introduced such that:

$$\ln y = \ln \alpha_1 + \ln \alpha_2 D + \beta_1 \ln x + \beta_2 \ln Dx \qquad (4)$$

Where:

y = Manufacturing hours per unit (HPU)

α_1 = Y-intercept for leg #1, equal to theoretical first unit hours for leg #1

α_2 = Intercept adjustment for leg #2, such that $\alpha_1 + \alpha_2$ equals the Y-intercept for leg #2

β_1 = Rate of learning for leg #1, such that 2^β equals learning curve slope #1

β_2 = Rate of learning for leg #2, such that $2^{(\beta 1 - \beta 2)}$ equals learning curve for leg #2

To demonstrate how such a curve can be built, a notional data set was constructed as follows. Based on a visual inspection of Figure 3, unit 101 was chosen as the breakpoint *T*. To illustrate further, a table of selected units (Table 1) is displayed to show *D*, *x* and *Dx.* Finally, a sample output from Microsoft Excel (Figure 4) is shown after selecting the natural logarithm of hours per unit as the dependent variable *y* and regressing *D*, *x* and *Dx* as independent variables.

We may interpret the results as follows: For units 1-100, hours per unit are calculated using a theoretical first unit (TFU) of 5,192 hours and a slope of 74.3%. For units 101-300, hours per unit are calculated using a TFU of 1,497 hours (calculated as $e^{(8.555 - 1.244)}$ or $\alpha_1+\alpha_2$) and a slope of 89.9% (calculated as $2^{(-0.428 + 0.274)}$ or $2^{(\beta 1 + \beta 2)}$). This equation also has a high R^2 of 0.97 – which significantly fits the historical data better than an equivalent single slope learning curve (R^2 = 0.925).

It can be argued that the R^2 of the best fit of a discontinuous line will always show some improvement, however miniscule, over the best fit of a single line. To test the statistical significance of the parameter values for period 1 (pre-break) and period 2 (post-break), a Chow test can be performed in the format suggested by Kennedy (2002). By comparing the sum of squared errors (SSE) of the regressions for a

Table 1. Notional Data Table (Partial).

					Dependent Variable	Independent Variables		
Unit	HPU	LN(Unit)	T	LN(T)	LN(HPU)	LN(x)	D	LN(Dx)
1	5,020	-	101	4.62	8.52	-	-	-
2	4,065	0.69	101	4.62	8.31	0.69	-	-
3	3,248	1.10	101	4.62	8.09	1.10	-	-
4	3,038	1.39	101	4.62	8.02	1.39	-	-
5	2,628	1.61	101	4.62	7.87	1.61	-	-
6	2,272	1.79	101	4.62	7.73	1.79	-	-
7	2,216	1.95	101	4.62	7.70	1.95	-	-
8	1,949	2.08	101	4.62	7.58	2.08	-	-
9	2,001	2.20	101	4.62	7.60	2.20	-	-
10	2,030	2.30	101	4.62	7.62	2.30	-	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	682	4.60	101	4.62	6.53	4.60	-	-
100	668	4.61	101	4.62	6.50	4.61	-	-
101	798	4.62	101	4.62	6.68	4.62	1	4.62
102	677	4.62	101	4.62	6.52	4.62	1	4.62
103	724	4.63	101	4.62	6.59	4.63	1	4.63
104	692	4.64	101	4.62	6.54	4.64	1	4.64
105	680	4.65	101	4.62	6.52	4.65	1	4.65
106	746	4.66	101	4.62	6.61	4.66	1	4.66
107	799	4.67	101	4.62	6.68	4.67	1	4.67
108	724	4.68	101	4.62	6.59	4.68	1	4.68
109	763	4.69	101	4.62	6.64	4.69	1	4.69
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

single slope as opposed to a multiple leg slope, we can create an F-statistic of the form:

$$\frac{[SSE(combined) - SSE(separated)]/K}{SSE\ (separated)\ /\ (T_1 + T_2 - 2K)} \quad (5)$$

The "combined" dataset represents the residuals for the single leg curve, while the "separated" dataset represents the residuals for the two-leg curve with its dummy variables representing pre- and post-break datasets, where K represents the number of parameters (including the intercept) of the combined dataset, T_1 the number of observations in period 1, and T_2 is the number of observations in period 2.

The resulting test-statistic derived from equation (5) can be evaluated against a F-table at the desired level of error with K and T_1+T_2-2K degrees of freedoms. Our null hypothesis – that $\alpha_1 - \alpha_2 = 0$ and $\beta_1 - \beta_2 = 0$ – would conclude there is no significant structural break in the hours data to justify a two-leg curve. The alternate hypothesis – that $\alpha_1 - \alpha_2 \neq 0$ and $\beta_1 - \beta_2 \neq 0$ – would conclude just the opposite: there is a significant structural break in the data beginning at unit T. In our notional example (calculations not shown here but available upon request), we can reject the null hypothesis with a 99.9% confidence, concluding that our data does indeed show a break in the learning curve slope.

As noted above, a high R^2 – even one buttressed by a sufficiently high F-statistic for the Chow test -- does not guarantee the accuracy of the forecasts made from this equation. But this two-leg model is more in line with the theoretical expectations set by the S-curve as well as historical experience, and therefore more likely to give us a better projection of future costs.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.985
R Square	0.971
Adjusted R Square	0.970
Standard Error	0.058
Observations	300

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	32.352	10.784	3,249.084	0.000
Residual	296	0.982	0.003		
Total	299	33.335			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Natural log - Intercept (ln α1)	8.555	0.023	365.375	-	8.509	8.601	8.509	8.601
Natural log - Beta-1 (ln β1)	(0.428)	0.006	(68.580)	0.000	(0.440)	(0.416)	(0.440)	(0.416)
Natural log - Alpha-2 (ln α2)	(1.244)	0.074	(16.877)	0.000	(1.389)	(1.099)	(1.389)	(1.099)
Natural log - Beta-2 (ln β2)	0.274	0.015	18.706	0.000	0.245	0.303	0.245	0.303

Results:

5,192	TFU for Leg #1
74.3%	Slope for Leg #1
1,497	TFU for Leg # 2
89.9%	Slope for Leg #2

Figure 4. Microsoft Excel Output

Peril: Development Slopes – Ignorance Is Not Bliss

One of the conclusions of the S-curve theory is that improvement slopes in a development phase of a program will be relatively flat. This is because so many programmatic issues conspire together to prevent rapid improvement in costs. These include very high number of engineering changes, late parts usually due to late engineering release, tooling which requires rework, engineering errors, and the realization that manufacturing processes and part flows that work on the drawing board don't necessarily work on the shop floor. It's hard to overstate the chaos of the start-up of a manufacturing line. It's a recurring theme on many programs that parts are installed in one station only to be ripped out and replaced just a few stations further down the line because of engineering changes.

Yet in much of the learning curve literature this tends to be glossed over. In many surveys the data from the development units is either excluded, or data limitations prevent an analysis of development slopes. For example, in RAND's 2001 study of military fighter aircraft (F-14, F-15, F-16, F-18, AV-8B), Engineering and Manufacturing Development (EMD) data is included as a single aircraft lot, not as individual units. This prevents any analysis of a unique EMD slope. RAND's conclusion that the average improvement slope for manufacturing is 80% therefore tells us little about the shape of the improvement curve in the development phase itself (Younossi, 2001). Similar issues plague other industry-wide studies (Resetar, 1991; Hess, 1987; Levinson, 1966). However, insight into individual unit cost data is often only found in company-proprietary datasets. Only at the individual unit level does the slower rate of improvement for the development phase becomes apparent.

So why does this matter? Because decisions which are made about the slope of the initial units can be critical to establishing the eventual production cost.

Let us take a simple example. An estimator establishes the cost of a 300-unit program using an S-curve profile. For the ten-unit development phase, he projects using an 86% slope. When production begins at unit 11, the slope steepens to a 72% slope which is maintained until T-101, at which point it flattens to 82%. When the estimate starts running through the company approval cycle, however, the program manager objects.

For one thing, the program manager doesn't like the idea of a three-leg curve. Shouldn't a learning curve be a single line? Moreover, a relatively flat development slope might appear uncompetitive to the source selection committee. The discussion goes on for several minutes, until the program manager suggests that the program use the same T-1 and T-300 costs as originally proposed but simply draw a single slope in log-log space between those two points. The program manager recognizes that the development phase will be initially understated, but it is only for ten units, after all, and it might put the company in a better competitive position.

Figure 5 illustrates the program manager's solution. Unfortunately, this solution does not just put the development cost estimate at risk. It also significantly understates the cost of the first three production lots.

While it is true that the gap between the two approaches begins to close at Lot 3, the damage has been done. If the analyst's original estimate was right, the first two production lots will overrun by 21% and 15% respectively. This could lead to adverse publicity, and the perception that the program is unable or unwilling to control its costs. It could also lead to a substantially degraded financial position for the company. If the original estimate is wrong – and history says the odds it will be too low are far better than being too high – then the damage to the program and to the company will be even greater.

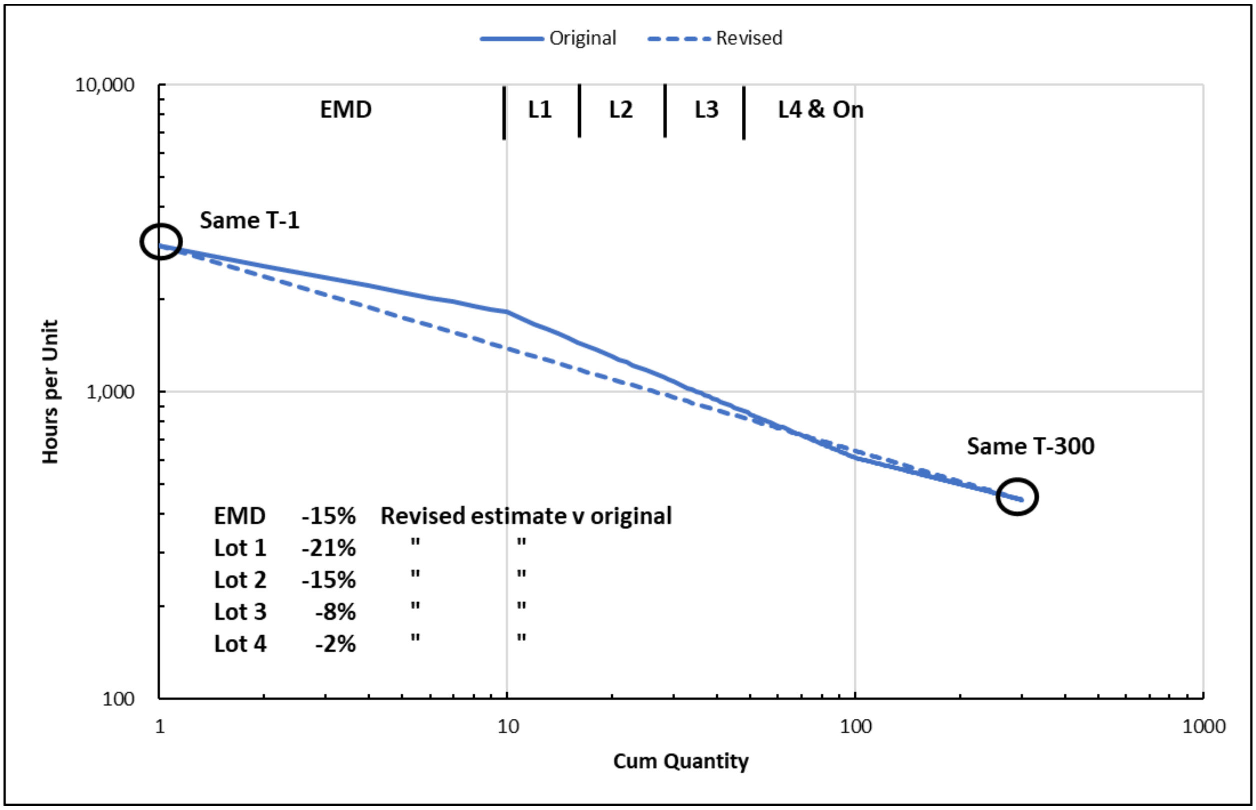

Figure 5. Impact of Flatter Development Slope on Performance (Notional)

Solution: Recognize That Choices Matter

The solution here is relatively simple – be aware that choices about learning curve slopes during the development phase impact the estimate. These choices can be consequential whether the program follows a straight-line logarithmic function or an S-curve pattern.

One word on the assumption that development programs always have relatively flat improvement slopes: It is true that you can sometimes find development programs which have a steep improvement curve. In the author's experience, such an occurrence is typically due to an unusually high first unit cost, which in turn is driven by programmatic issues. Programs that push the manufacturing state of the art by introducing new or radical processes often show high first unit costs as companies struggle to implement these on the shop floor. This poor performance is typically followed by rapid cost improvement as issues are worked through. The Convair B-58 program, built in the 1950's and 1960's, provides an example. Not only was the B-58 the first supersonic bomber, but it introduced the first widespread use of honeycomb bonded structure (Hess, 1987). Issues with the fabrication of the panels and their subsequent installation led to a high first unit cost but a rapid movement down the learning curve for follow-on units (Large, 1974). These examples are the exception, however, and not the rule.

Peril: The "Slippery Slope" – Extraordinary Impacts and Recovery Slopes

One of the most vexing situations for an estimator are those cases where there are sharp increases in unit cost over time -- but the increases are expected to be mitigated over time. These can be divided roughly into two camps: (a) "expected" disruptions, such as major engineering changes,

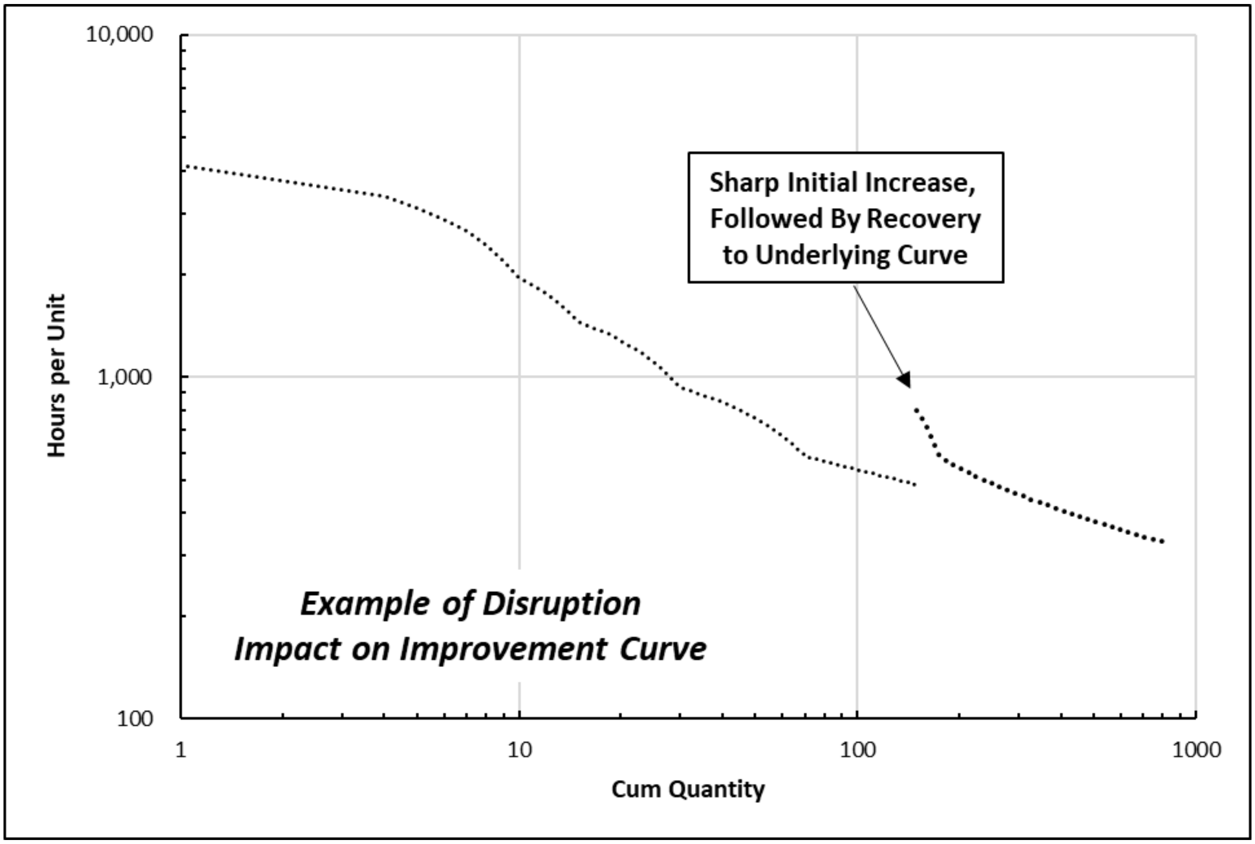

Figure 6. Disruption Example (Notional)

production breaks or work transfers between sites and (b) "unexpected" disruptions caused by unforeseeable circumstances. An example of an unexpected disruption would be a critical load part shortage which creates significant behind schedule and out of station costs. Both types of disruptions appear similarly on graphs of historical costs. Figure 6 shows an example of this kind of behavior, with a sharp initial increase in cost and an eventual asymptotic recovery to the underlying curve.

Of course, *ex ante* we do not have the advantage of how and when this recovery will occur. Herein is our estimating dilemma. How might we deal with this issue?

This is best illustrated by an example. At unit 150 a severe part shortage produces a substantial behind schedule position with workarounds and significant out of station work. This situation (See Figure 7) is expected to end at some point but no one can say with confidence when.

There are two often-taken approaches to this. The first is to simply ignore these units and project the cost as if these impacts had never occurred (See Figure 8). This is often justified by a claim that it represents where the company "should be" performing had the extraordinary impact not occurred. Whether the extraordinary impact is anticipated (e.g., driven by an engineering change or a production break) or unexpected (e.g., driven by part shortages or schedule problems), this procedure is never justified. Assuming away these type of cost increases may seem like a viable approach to the cost estimator. It is never one to the shop floor managers and directors who cannot deal with the world as we wish it was, but as it is. This approach often creates an insurmountable gap between current performance and what the analyst thinks the values "should be."

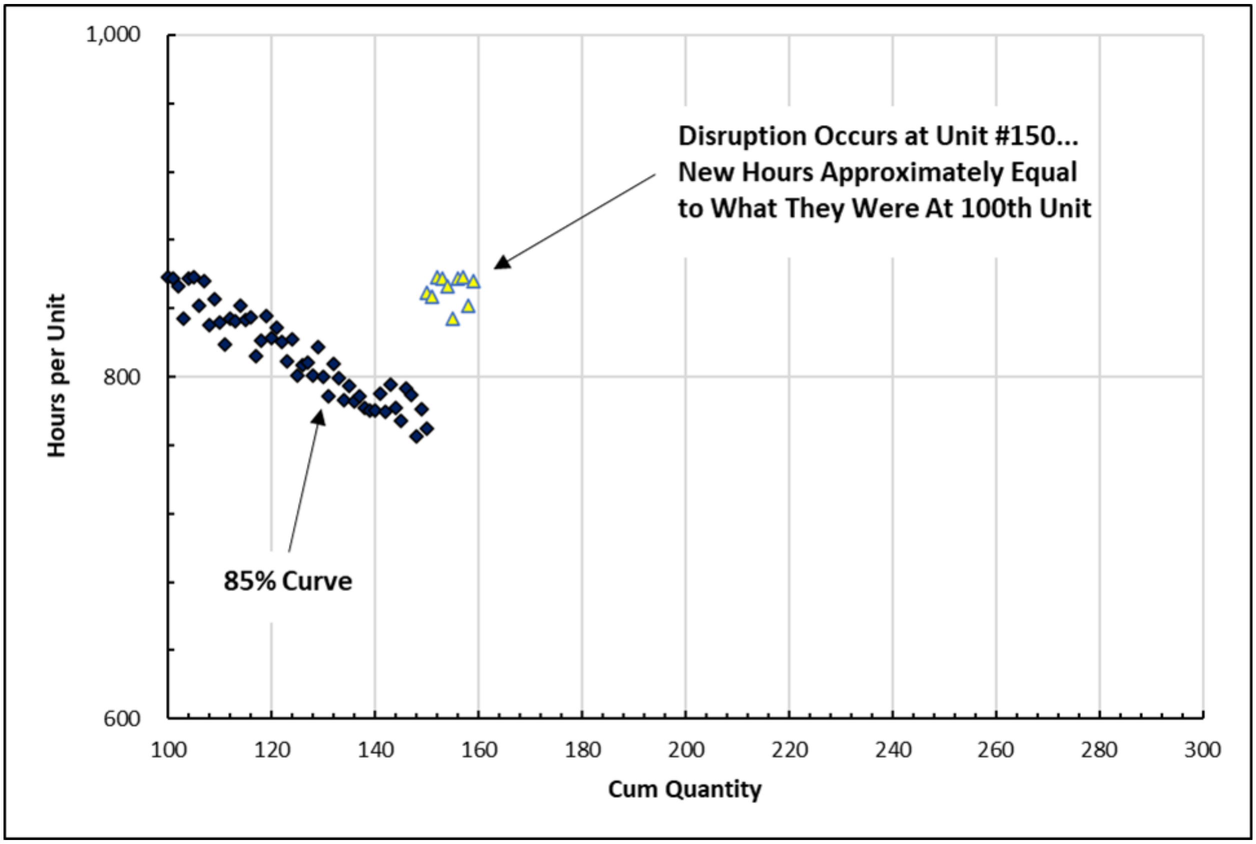

Figure 7. Illustration of Disruption (Notional)

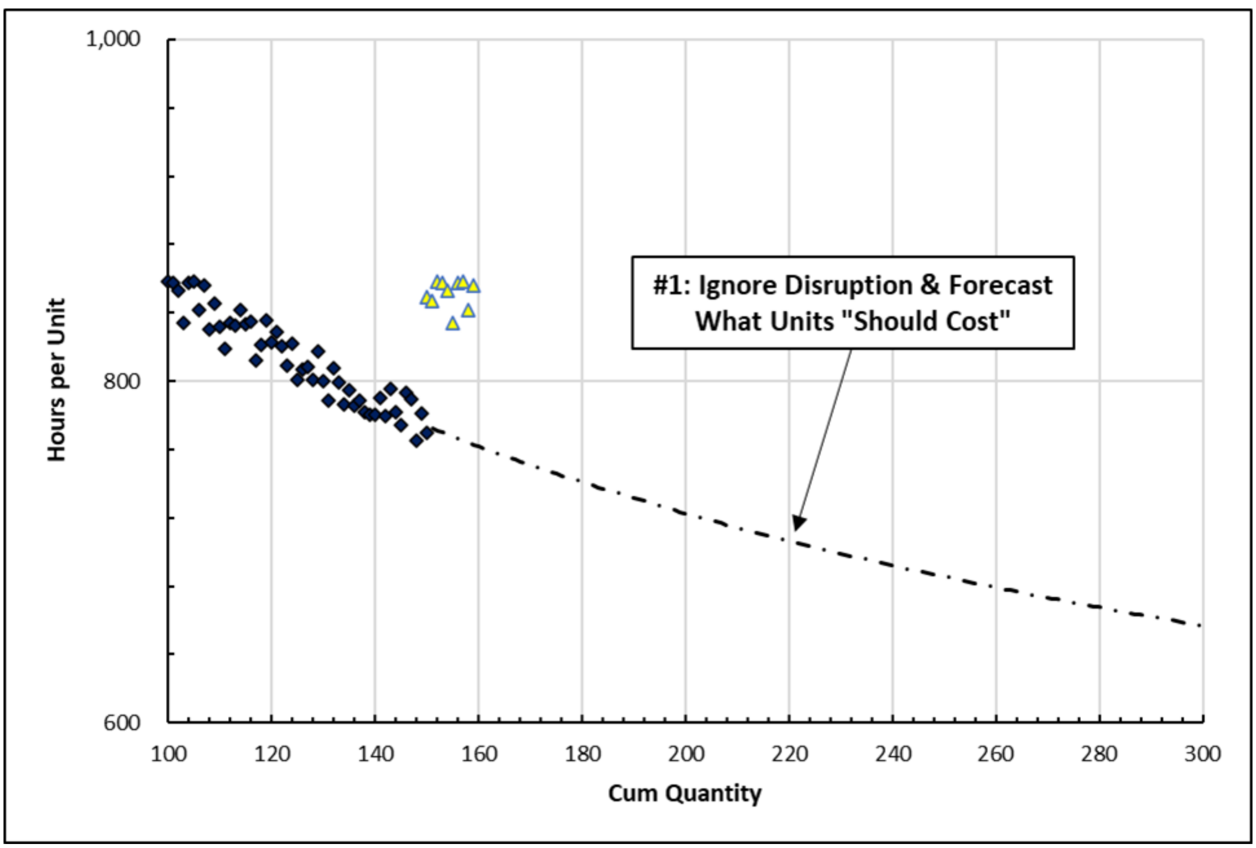

Figure 8. Recovery Curve – Doing Nothing

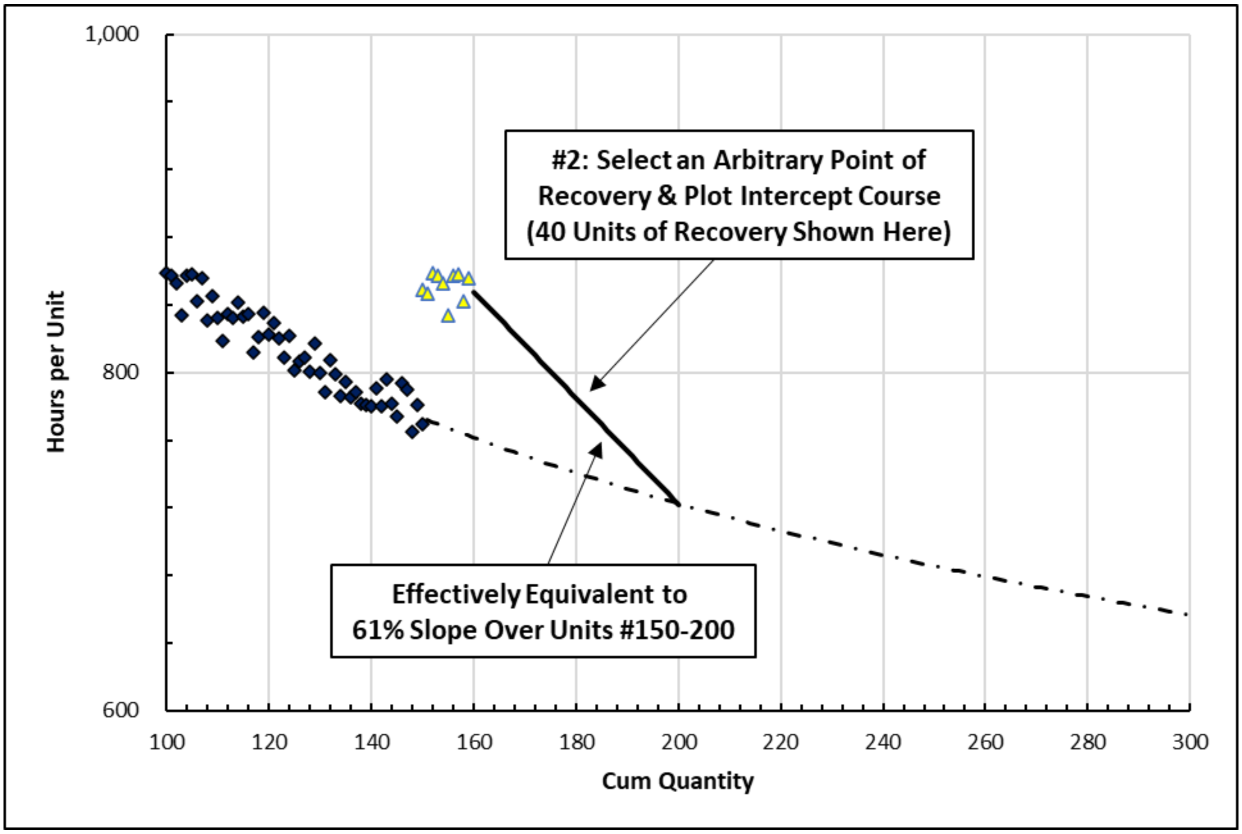

Figure 9. Recovery Curve - Point of Recovery

Unfortunately for the shop floor, that gap cannot simply be wished away.

The second approach is to create a "recovery slope" which accepts the cost increases but returns unit cost near to what it would have been had the extraordinary impact not occurred. This is clearly a more realistic approach than the first. But how quickly should we forecast recovery?

Frequently, an arbitrary number of units is chosen, and the recovery is then forecast over that number of units (See Figure 9). Sometimes, the choice of units is based on historical analogies. Sometimes it is based on a point in time when the manufacturing schedule recovers to the baseline. Sometimes it is simply picked out of the air. All these have problems. Our historical analogies may not be apt, or we may not have the data. Cost improvements usually lag schedule improvements, especially since schedule improvements are often made by increasing manpower or overtime or both. Bottom line, it is very easy to make unrealistic shop projections which cannot be achieved.

Solution: Calculating Learning Setback and Projecting Forward

A more reasonable approach is to take the break-in point of the disruption and set back the unit position on the learning curve (Fowlkes, 1963). For example, prior to the part shortage we were on an 85% slope. The first unit to feel the impact of the shortage represents approximately 850 hours per unit – equivalent to position 100 on that same 85% slope. To forecast the recovery, we regress on the learning curve back to unit 100 and forecast future units on the same pattern as established in the past, i.e., the next five units are equivalent to the cost of units 101 thru 105, etc. on an 85% improvement slope.

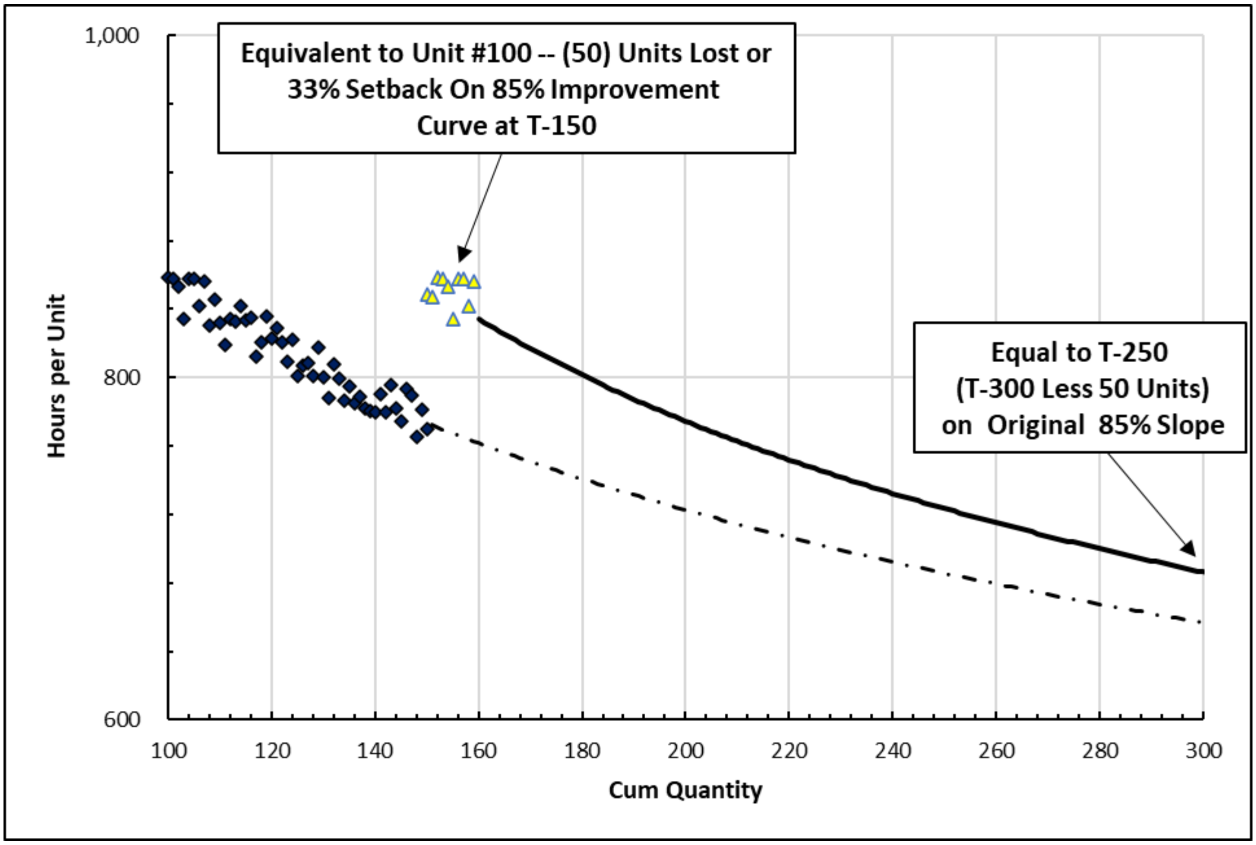

Figure 10. Recovery Curve – Setback

This produces a "scallop" in the overall improvement curve (See Figure 10). True to the learning curve pattern, the most improvement is seen in the initial units after the disruption, with the unit-to-unit decreases slowing as we move farther away from the initial disruption. In this case, we recover asymptotically to the old cost curve – that is, we never achieve the same hours per unit we would have anticipated had the disruption not occurred. But we come closer and closer to it until eventually the difference between the two becomes marginal.

The use of setback in the learning curve is widely accepted for production breaks and engineering changes (Anderlohr, 1968; DCAA, 1994; Smith, 1986). But there is sometimes resistance to using it in other scenarios.

This resistance is largely based around the idea that learning – and the loss of learning – exclusively centers around the operator on the shop floor. In the case of engineering changes (the operator must learn a new way of building the part) and production breaks (there is a significant turnover on the floor with employees receiving new assignments), there is clearly an impact to the body of knowledge the shop floor operator has accumulated. But our common use of the term "learning curve" often misleads us into believing that cost improvement only results from repetitive operations by the mechanics. It is more accurately called out as a "cost improvement curve."

In his paper on production breaks, Anderlohr defined five elements of learning: (1) operator learning, (2) supervisory learning, (3) tooling, (4) continuity of production and (5) manufacturing methods (Anderlohr, 1969). Yet the improvement that comes from the repetition of tasks by shop personnel accounts for slightly more than 20% of the total cost improvement. The rest of "learning"

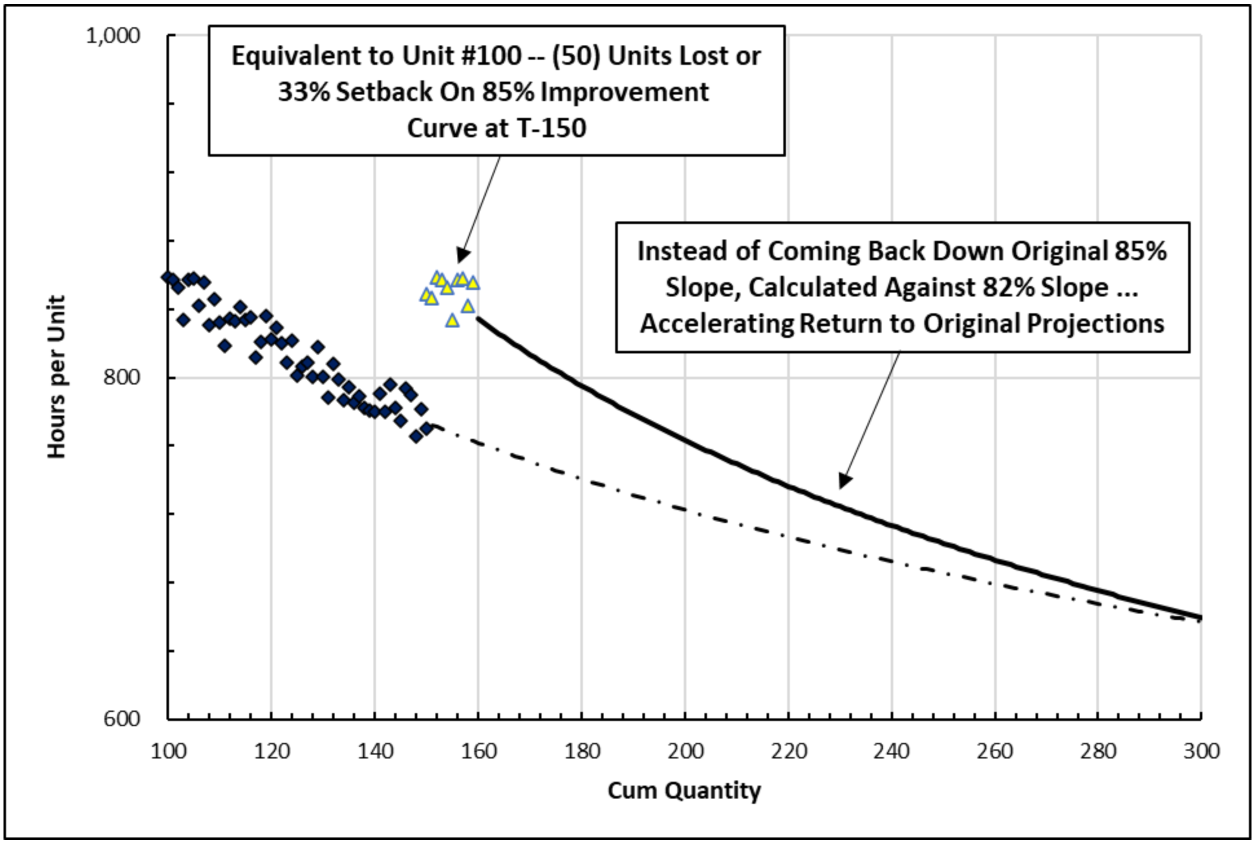

Figure 11. Recovery Slope – Accelerated Setback

comes from other sources. (Jones, 2001). If we can adjust our position on the improvement curve for negative impacts to operator or supervisor learning, surely it is legitimate to adjust it for negative impacts to the other three areas as well? Viewed from this perspective, it should be plain that, for example, an interruption to the supply chain due to late parts – which in turn creates part shortages, workarounds and behind schedule conditions – represents a retrograde to the existing improvement curve and can be fairly represented by a setback on the learning curve.

In the author's experience, this produces the most realistic and reasonable recovery slope and the one most achievable by the shop floor. But there are cases where a more aggressive approach may seem appropriate. Smith (1986) makes a common argument: "The firm is reexperiencing, not experiencing; they are going down a cost improvement curve they have been over before and should be better equipped to solve the problems the second time around so some method of accelerating recovery…may be useful". We can modify the setback methodology shown above and assume that we forecast the new units not on the same pattern seen in the past – the 85% improvement slope – but a slightly more aggressive one. In this case, an 82% slope has been used (See Figure 11).

This allows us to completely return to the hours per unit projected on the old cost curve. (In fact, had we continued the projection another ten or twenty units, the recovery slope would fall underneath the old cost curve, giving us a lower per unit value.) The more aggressive the slope assumption, the faster the interception point will be achieved. However, we can easily fall in the same trap as the earlier case where we selected an arbitrary number of units and drew a line to intercept the old cost curve. If 82% was an appropriate slope, why not 80%? Why not 78%? Why not 75%? It is easy to rationalize the answer

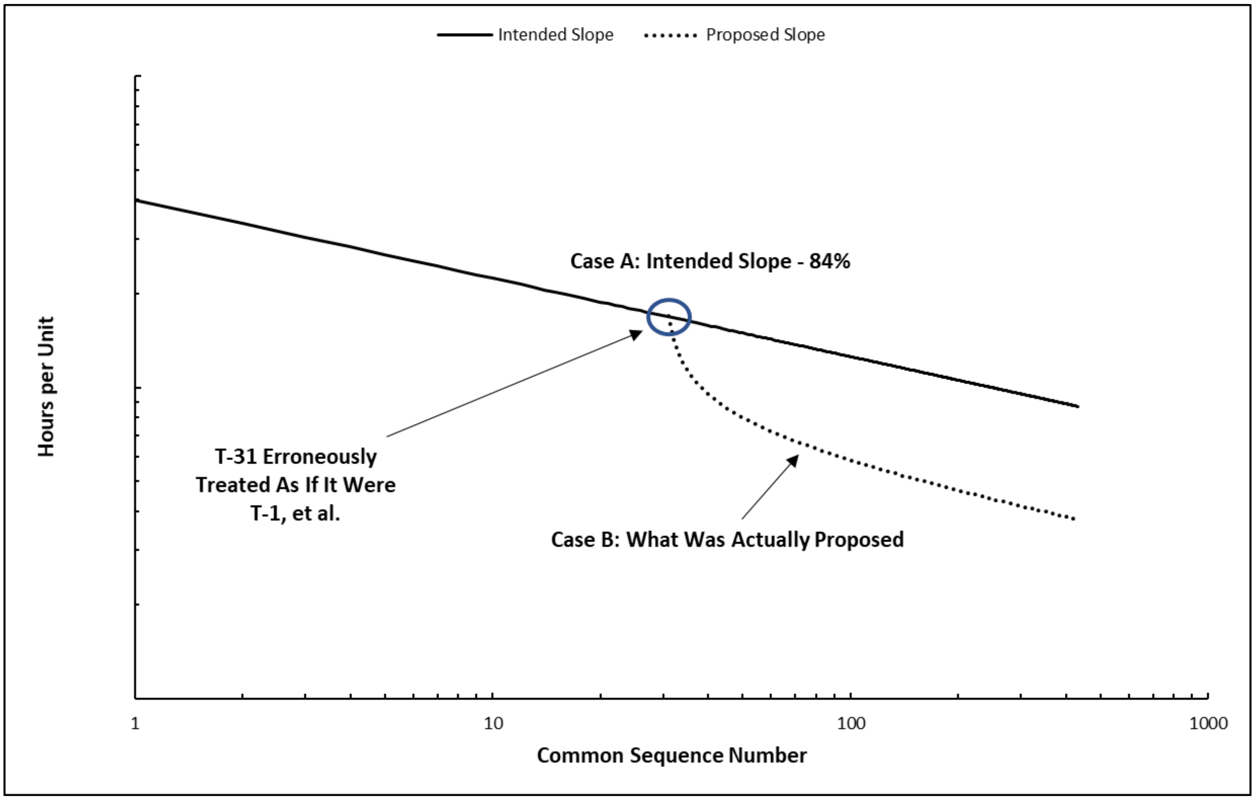

Figure 12. Illustration of Misidentified First Unit (Notional)

we (and our management) want to hear. Cochrane (1968) suggests a methodology for calculating an accelerated recovery, but it too requires an arbitrary choice of an acceleration factor which might be difficult to justify. The best guide would seem to be prior experience, but often we cannot find an analogy which exactly correlates to the situation we are estimating.

In short, projecting recovery slopes from disruptions is fraught with potential risks and the greatest care must be taken with doing so. When calculating a recovery slope, it is always best to review your assumptions and projections with the shop floor to make sure what you have mapped out can in fact be realized.

Peril: Being Careless When Establishing the First Unit

This example is drawn from an actual proposal. Values referenced below are notional.

A small aircraft pylon used to carry mission equipment required subassembly work. For the first thirty units, the Special Projects organization produced it on an 84% learning curve. At unit 31, the task was transferred from Special Projects to the regular Production department, who would produce the next order for 400 units.

The cost analyst (fortunately not the author!) proposed that the first Production unit would have the same hours per unit as the last unit produced by Special Projects. He also proposed the same 84% learning curve slope going forward. However, for projection purposes, he treated the first Production unit as unit one on the learning curve. The estimator apparently believed he was setting the unit costs back on the learning curve. But while he reset the cumulative unit count, he did not adjust the hours at unit #31 to something higher.

Figure 12 shows the consequences. Case A represents what the Production department expected to see when the contract was awarded.

Case B represents what they found in the estimate – an estimate that was approximately half of what they expected!

By treating T-31 and subsequent units as if we were restarting the learning curve back at unit 1, we have restarted the 16% cost reduction that occurs every time the number of units doubles. This significantly accelerates the rate of learning – which was not the intention of the estimator. GAO (2020) refers to this as "disjoint theory" (treating the first production unit as T-1 and restarting the curve) as opposed to "sequential theory" (treating the first production unit as the last development unit plus one). There may be times where disjoint theory is appropriate, but in this case the analyst simply did not realize what had been done.

Fortunately, Production was able to mitigate the impact by holding a series of lean events and substantially restructuring the production process – as it turned out, there were significant inefficiencies in the existing production process which were subsequently eliminated. However, this happy accident cannot be counted on in the future to save an estimator from his mistakes.

Solution: Take Care and Graph, Graph, Graph!

Fortunately, the solutions are relatively simple. As a rule, analysts should always graph their learning curve results – preferably in both a log-log and an arithmetic space. Graphing the actual cost history as well as the projected hours per unit would have quickly surfaced the problem. In addition, examine your takeoff point for projections and its position on the curve carefully. Seemingly insignificant decisions can have profound impacts on the numbers.

Peril: Steep Curves = Efficiency?

The author has heard proposal evaluators frequently assert that a flat learning curve is proof of manufacturing inefficiency. Its counterpart is often asserted as well: a steep learning curve proves the efficiency of a manufacturing operation. In fact, the slope of a learning curve by itself does not prove that a factory is efficient or inefficient. A hypothetical example will demonstrate this.

Company A assembles widgets; it has demonstrated an 80% learning curve over 1,000 units. Company B builds a similar but not identical product and demonstrates a 90% learning curve over the same range. There has been no transfer of manufacturing knowledge or personnel between the two companies. Which company is more efficient?

Many cost estimators would immediately answer Company A since it has the steeper learning curve. But this ignores the reasons why Company A had such a steep learning curve. This is quickly demonstrated by comparing the performance of the two companies on an hours per pound basis (See Figure 13). This shows Company A's high first unit cost, exceeding 40 hours per pound. Upon investigation, it turns out this high T-1 was driven by late engineering release, inadequate tooling, late material, and the oversizing of shop floor crews to recover manufacturing schedule.

Company B on the other hand had its engineering released on time, which allowed its tooling program to build high quality tools and deliver them to the floor when needed. On-time engineering allowed the supply base to deliver its parts on time, which in turn allowed Production to size its crews efficiently and still maintain the production schedule. Its first unit cost was almost half of Company A's.

Both companies ended the 1,000th unit at the same hours per pound. But over the course of those thousand units, it took Company A almost 25% more hours to produce its product.

A steep learning curve can demonstrate a strong dedication to lower costs and continuous improvement. It can also indicate the necessity to recover from poor performance and

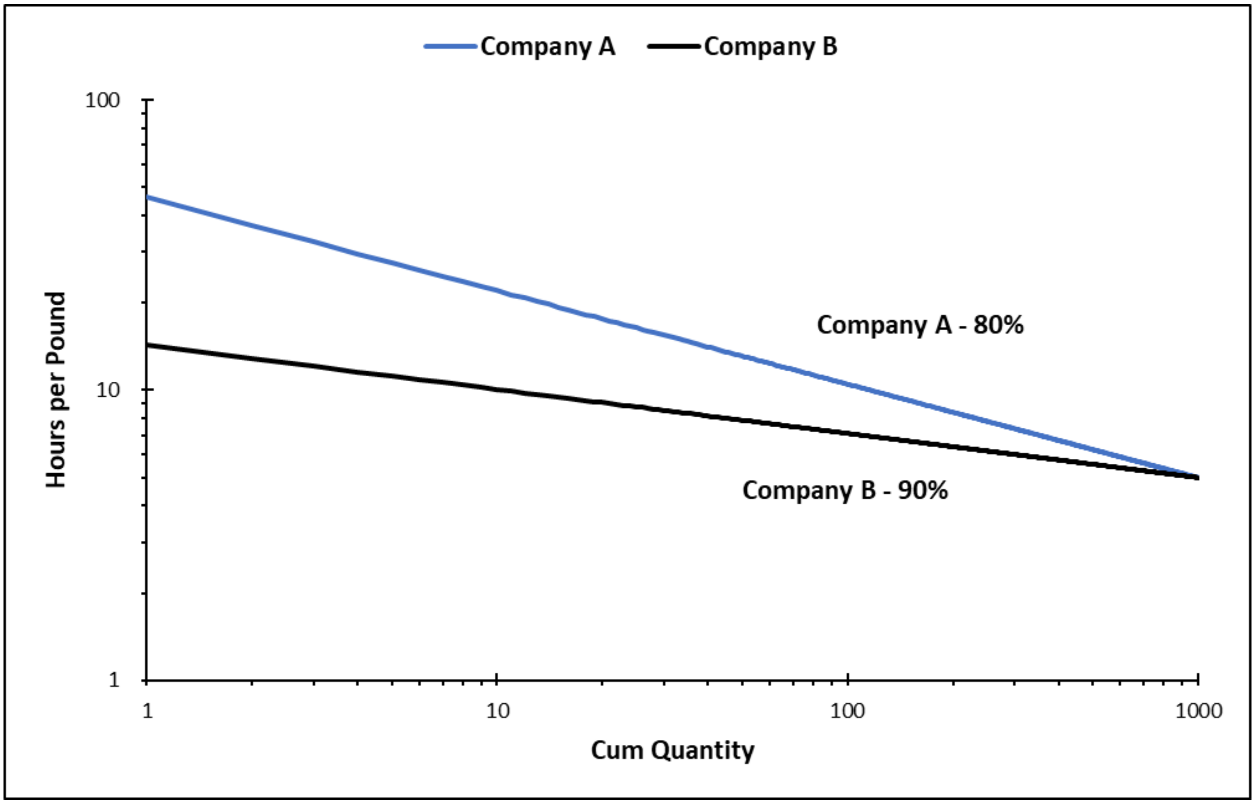

Figure 13

mismanagement on the earliest units. "[T]he more room there is for improvement," noted Fowlkes (1963), "the more improvement there is to be expected." Without further investigation, it cannot be determined from the numerical value of a learning curve slope alone which of these two cases is true.

Solution: Avoiding the Facile Conclusion

There is a widespread perception among cost estimators that relatively flat learning curves are a symptom of production inefficiency, and -- by implication – that relatively steep slopes are proof of manufacturing efficiency. In fact, as our example demonstrates, just the opposite may be the case. The learning curve slope alone cannot tell us if a manufacturing operation is efficient or not. Further analysis and understanding behind the underlying trends are necessary. Unfortunately, there is no easy way out.

Conclusions:

The quantity of books, articles, and academic research published about learning curves is astonishing. Literally hundreds of publications have been released since T. P. Wright's original 1936 article. Authors have suggested a variety of models and approaches: Wright's cumulative average model, Crawford's unit curve model, the Stanford B-curve, DeJong's incompressibility model and Cochran's S-curve are only a few examples. (Wright, 1936; Stanford Research Institute, 1949; DeJong, 1957; Cochran, 1960; Cochran, 1968) And yet within this wealth of material, there are relatively little guidance on what *not* to do. A new driver should not be handed a key to the sports car in the driveway before being previously schooled on speed limits and stop signs.

In the introduction, this paper was offered as a warning label to be attached to the improvement curve. Each of the five situations outlined represents a potential pitfall which can entangle the cost estimator and transform the learning curve from a useful tool to a danger to himself and others. The negative consequences of a bad estimate – on company profits and government funding – can be severe.

Unfortunately, most learning curve training rarely addresses these issues. It is content to show the basic calculations for Wright and Crawford curves, offer some advice on midpoint calculation and show a methodology for dealing with major engineering changes or production breaks. But it rarely goes much beyond these areas. It simply assumes estimators will find out about those other matters "soon enough." They will – but they might take someone else down with them in the process.

Cautionary tales rarely make compelling reading. After all, who among us actually reads the warning labels attached to the products we buy? But in this case, questioning long-held premises or putting in an extra half hour of analysis may yield unexpected benefits. By obeying the speed limits of estimating, our hypothetical driver and his sports car in the driveway might make it back home in one piece.

References

Anderlohr, George (1969, September). "What Production Breaks Cost," *Industrial Engineering*, 34-36.

Asher, H. (1956). *Cost-Quantity Relationships in the Airframe Industry*. Santa Monica, California: RAND Corporation.

Boone, E., Elshaw, J., Koschnick, C., Ritschel, J., Badiru, A. (2021). "A Learning Curve Model Accounting for the Flattening Effect in Production Cycles." *Defense Acquisition Research Journal*, 28(96), 72–97.

Carr, G.W. (1946, April). "Peacetime Cost Estimating Requires New Learning Curves." *Aviation*, *45*, 76-77.

Cochran, E.B. (1960, July-August). "New Concepts of the Learning Curve." *The Journal of Industrial Engineering*, 317-327.

Cochran, E.B. (1968). *Planning Production Costs: Using the Improvement Curve*. San Francisco, CA: Chandler Publishing Company.

Crawford, J. R. (1944). *Learning Curve, Ship Curve, Ratios, Related Data*. Burbank, CA: Lockheed Aircraft Corporation.

DCAA Contract Audit Manual (1996). DCAAM 7640.1, vol. 2. Washington, DC: Government Printing Office.

DeJong, J. R. (1957). "The Effects of Increasing Skill on Cycle Time and Its Consequences for Time Standards." *Ergonomics*, 1(1), 51-60.

Dutton, J., Thomas, A. (1984, April). "Treating Progress Functions As a Managerial Opportunity." *The Academy of Management Review*, 235-247.

Fowlkes, Tommie F. (1963) *Aircraft Cost Curves*. Fort Worth, TX: General Dynamics Convair Division.

Fox, B., Brancato, K., Alkire, B. (2008). *Guidelines and Metrics for Assessing Space System Cost Estimates*. Santa Monica, CA: RAND.

Government Accountability Office (GAO) (2020). *Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Program Costs*. GAO-20-195G. Washington, DC: United States Government Accountability Office.

Hess, R. W., Romanoff, H.P. (1987). *Aircraft Airframe Cost Estimating Relationships: Bombers and Transports*. Santa Monica, CA: RAND.

Johnstone, Brent M. (2015) *Improvement Curves: An Early Production Methodology*. Paper presented at the 2015 International Cost Estimating and Analysis Association (ICEAA) Professional Development and Training Workshop, San Diego, CA. Retrieved from http://www.iceaaonline.com/ready/wp-content/uploads/2015/06/PA03-Paper-Johnstone-Improvment-Curves.pdf.

Jones, Alan R. (2001). "Case Study: Applying Learning Curves in Aircraft Production – Procedures and Experiences." *Maynard's Industrial Engineering Handbook*, 5th edition. New York, NY: McGraw-Hill.

Kennedy, Peter (1992). *A Guide to Econometrics*. Cambridge, MA: The MIT Press.

Large, Joseph P.; Hoffmayer, Karl; Kontrovich, Frank (1974). *Production Rate and Production Cost.* Santa Monica, CA: RAND.

Levinson, G. S.; Barro, S. M. (1966). *Cost Estimating Relationships for Aircraft Airframes*. Santa Monica, CA: RAND.

Miller, F. D. (1971, July). "The Cubic Learning Curve – A New Way to Estimate Production Costs." *Manufacturing Engineering & Management*, 14-15.

Nussbaum, Daniel A; Mislick, Gregory K. (2015). *Cost Estimation: Methods and Tools*. Hoboken, NJ: John Wiley & Sons, Inc.

Resetar, Susan A; Rogers, J. Curt; Hess, Ronald W. (1991). *Advanced Airframe Structural Materials: A Primer and Cost Estimating Methodology*. Santa Monica, CA: RAND.

Shumeli, Galit (2010). "To Explain or to Predict?" *Statistical Science*, 25(3), 289-310.

Smith, Larry L. (1986). *Cost Improvement Analysis*. Dayton, OH: Air Force Institute of Technology.

Stanford Research Institute (1949). *An Improved Rational and Mathematical Explanation of the Progress Curve in Airframe Production*. Stanford, CA: Stanford Research Institute.

Wright, T.P. (1936, February). "Factors Affecting the Cost of Airplanes." *Journal of the Aeronautical Sciences*, 3, 122-128.

Younossi, Obaid; Kennedy, Michael; Graser, John C. (2001). *Military Airframe Costs: The Effects of Advanced Materials and Manufacturing Processes*. Santa Monica, CA: RAND.

***Brent M. Johnstone** is a Technical Fellow and production air vehicle cost estimator at Lockheed Martin Aeronautics Company in Fort Worth, Texas. He has 34 years' experience in the military aircraft industry, including 31 years as a cost estimator. He has worked on the F-16 program and since 1997 has been the lead Production Operations cost estimator for the F-35 program. He has a Master of Science from Texas A&M University and a Bachelor of Arts from the University of Texas at Austin.*

The International Cost Estimating and Analysis Association is a 501(c)(6) international non-profit organization dedicated to advancing, encouraging, promoting and enhancing the profession of cost estimating and analysis, through the use of parametrics and other data-driven techniques.

www.iceaaonline.com

Submissions:

Prior to writing or sending your manuscripts to us, please reference the JCAP submission guidelines found at

www.iceaaonline.com/publications/jcap-submission

Kindly send your submissions and/or any correspondence to JCAP.Editor@gmail.com

International Cost Estimating & Analysis Association

4115 Annandale Road, Suite 306 | Annandale, VA 22003

703-642-3090 | iceaa@iceaaonline.org