Chapter Five

AIRFRAME COST-ESTIMATING METHODOLOGY

In this chapter, we discuss how the survey results presented in Chapter Four should be applied to CER estimation.[1] We then estimate CERs for recurring labor hours in fighter production. These estimates are based on the MACDAR database that was developed in a series of research projects by RAND, Tecolote Research, Incorporated, and Science Applications International Corporation (SAIC) and sponsored by the Air Force Cost Analysis Agency (AFCAA). We have also included a short discussion on nonrecurring CERs in Appendix D.

APPLICATION OF SURVEY COST RATIOS TO CERs

A primary objective of estimating cost ratios of the sort shown in Table 4.9 is to facilitate accurate estimation of CERs for labor hours required to build new aircraft. Many CERs, including those we estimate in this report, are for labor hours used to produce the whole airframe. In this case, *the cost ratios of Table 4.9 must be adjusted before they are valid for use with the CERs.* This is because, as the ground rules given in Chapter Four state, these cost ratios apply only to labor used for all structural fabrication and assembly up through the airframe group level (wing, fuselage, and empennage) and do not

[1]CERs are equations in which cost is the left-hand-side (dependent) variable. On the right-hand side of the equation are explanatory (independent) variables such as cumulative and annual production and aircraft weight. With a data set of such variables, regression or some other statistical method can be used to estimate the parameters of the right side. Cost is usually measured as either dollars or labor hours; all CERs estimated in this report use labor hours.

apply to final assembly and checkout or to any subsystem installation. In a modern fighter aircraft, the latter two categories account for roughly 35 percent of recurring manufacturing labor, 13 percent of recurring and nonrecurring tooling, 40 percent of recurring quality assurance, and 57 percent of nonrecurring and recurring engineering.[2] These values must be accounted for in applying the cost ratios to CERs for all airframe labor.

Resetar, Rogers, and Hess (1991) proposed the following method for applying the cost ratios. Suppose there are M material types and L labor categories. Let C_l^m be the cost ratio associated with material m ($m = 1, \ldots, M$) and labor category l ($l = 1, \ldots, L$) (e.g., from Table 4.9, $C_{rec.\ manu.}^{aluminum}$ is 0.9). Let S_m be the share of material m in the airframe, with $\sum_{m=1}^{M} S_m = 1$.[3] Let σ_l be the share of category l labor hours associated with structural fabrication and assembly through the airframe group level. RRH then defined a weighted material cost factor for each labor category l, $(WMCF)_l$, as

$$(WMCF)_l = \sigma_l \sum_{m=1}^{M} S_m C_l^m + (1 - \sigma_l) \tag{5.1}$$

RRH proposed that $(WMCF)_l$ be used as a multiplicative factor on CERs.

---

[2]See Resetar, Rogers, and Hess (1991), p. 71 and below.

[3]Since the list of materials in Table 4.9 is obviously not exhaustive, an "other" category is needed. In recent practice, the cost ratios for carbon-epoxy have been used for "other," since it encompasses many nonmetallic composites, including fiberglass. However, one should look at the specific material composition of any aircraft to be analyzed before making such a decision.

We use a modified cost ratio, γ_l^m, which is the ratio appropriate to apply to all airframe hours, not just group-level (fuselage, wing, empennage) hours. We define

$$\gamma_l^m = \sigma_l C_l^m + (1 - \sigma_l) \quad (5.2)$$

Then we note

$$\begin{aligned}(WMCF)_l &= \sigma_l \sum_{m=1}^{M} S_m C_l^m + (1 - \sigma_l) \\ &= \sum_{m=1}^{M} S_m \gamma_l^m\end{aligned} \quad (5.3)$$

The γ_l^m can thus be applied directly to calculate the multiplicative factors for all-airframe CERs. They are a good intuitive measure of the total airframe-hour penalty associated with using nonaluminum material rather than the penalty associated with just one part of hours.

Table 5.1 shows our estimates of the γ_l^m, the cost ratios appropriate for applying at the all-airframe labor level for the late 1990s based on our survey. They are related to the C_l^m of Table 4.9 by Equation 5.2. Values of σ assumed were 65 percent for recurring manufacturing labor, 87 percent for recurring and nonrecurring tooling, 60 percent for recurring quality assurance, and 43 percent for nonrecurring and recurring engineering.

The reader will note that this section has assumed that material type does not affect hours required for final assembly, checkout, and subsystem installation. We in fact make this assumption throughout the study. We have seen no data or report indicating that this assumption is wrong. In our discussions with industry and other experts, we raised this issue, and no one argued that the assumption was misleading. However, future research in this area should certainly be done with an open mind, and new information and data sets that emerge should be analyzed to test the continuing appropriateness of the assumption.

Table 5.1

Late 1990s Cost Ratios, All-Airframe Labor Basis[a]

Material	Non-recurring Engineering	Non-recurring Tooling	Recurring Engi-neering	Recurring Tooling	Recurring Manufac-turing	Recurring Quality Assurance
Aluminum	1.00	0.97	0.97	0.94	0.94	0.95
Aluminum-lithium	1.00	1.09	1.00	1.06	1.00	1.04
Titanium	1.00	1.38	1.04	1.38	1.40	1.18
Steel	1.02	1.07	1.03	1.22	1.18	1.12
Carbon-epoxy	1.14	1.33	1.29	1.54	1.38	1.62
Carbon-BMI	1.16	1.42	1.32	1.67	1.46	1.65
Carbon-thermo-plastic	1.14	1.59	1.26	1.75	1.50	1.71

[a]Late 1980s aluminum = 1.0.

STATISTICAL ANALYSIS OF THE RECURRING COSTS OF RECENT FIGHTERS

The MACDAR Database

MACDAR contains annual recurring labor hour data for the aircraft and years shown in Table 5.2. It should be noted that the "number of aircraft" figure is somewhat misleading with regard to the size of the database; the cost data are accumulated by lot (on a yearly basis), showing, for each year, total production, total recurring labor hours associated with that production, and other annual data. In addition to the data on production years shown in Table 5.2, for each aircraft type there are data on the engineering and manufacturing development (EMD) lot,[4] which is generally produced over a period of more

[4]This activity was called full-scale development when the aircraft in Table 5.2 were developed.

than a year. The number of observations for this analysis in a statistical sense is the total number of lots, which is 64.

Table 5.3 shows the material distribution of each of the MACDAR aircraft.

"Stylized Facts"

Before proceeding to the statistical analysis, we present some interesting summary statistics derived from the MACDAR database. We begin with the share of the different types of recurring labor. Following standard DoD reporting practice, MACDAR identifies the same four types of recurring labor discussed in the last chapter: engineering, tooling, manufacturing, and quality assurance. Table 5.4 shows average shares of each in total recurring labor at four stages of production: EMD, first production lot, all lots, and last two lots. *It is critical to note that because of differences in production patterns among these programs, these figures have no formal statistical validity. They are literally descriptors of the data that are meant for general context only.*

Table 5.2

Aircraft in the MACDAR Database

Aircraft	Production Years (number of lots)	Number of Aircraft
AV-8B	82–86 (5)	195
F-14	71–78 (8)	439
F-15	73–91 (19)	1241
F-16	77–90 (14)	2357
F/A-18	79–91 (13)	1196

Table 5.3

MACDAR Aircraft Material Distribution (percentage of airframe structure weight)

Material	F/A-18	AV-8B	F-16	F-15	F-14
Aluminum	51	41	75	50	39
Steel	12	10	11	2	21
Titanium	14	6	1	34	30
Composites	11	25	2	2	2
Other	12	17	11	12	8

Table 5.4

Percentage of Recurring Labor at Different Stages of Production

Stage of Production	Engineering	Tooling	Manufacturing	Quality Assurance
EMD	24	22	45	9
First production lot	21	16	55	8
All production lots	15	13	62	10
Last two production lots	12	11	66	11

Our later regression estimates should be used to make formal statistical assessments. These results are literally intended to communicate to the reader "what the data look like"; hence the summary characterization "stylized facts." These are the arithmetic averages of each aircraft's proportions.

With that caveat in mind, we see interesting patterns in the data, none of which is surprising. Manufacturing labor dominates, constituting almost two-thirds of the program total. The shares of engineering and tooling fall as production proceeds; the share of manufacturing rises; and that of quality assurance is stable.

Another interesting set of stylized facts is the distribution of manufacturing labor across subcategories. MACDAR includes five such categories: fuselage, wing, empennage, subsystem installation, and integration. We categorize the first three into "structural" labor and the last two into "nonstructural" labor. ("Structural" is equivalent to "airframe group level" as defined in Chapter Four; "integration" is equivalent to "final assembly and checkout" in Chapter Four.) On average across aircraft types in MACDAR, structural labor is 65 percent of manufacturing labor and nonstructural 35 percent. Of total manufacturing labor, 45 percent is fuselage, 15 percent wing, 5 percent tail, 17 percent subsystem installation, and 18 percent integration. Another interesting set of average figures from the data set pertains to the relations of the various kinds of weight of the aircraft. (Weight definitions are given in Appendix B.) On average, structural weight is roughly three-quarters of airframe unit weight (AUW), and AUW is about two-thirds of empty weight. Thus, structural weight is approximately one-half of empty weight. Once again, these figures are just interesting descriptors of the data set and should not be construed as having formal statistical validity.

Statistical Results

We begin with regression equations in which labor hours are a function of several explanatory variables. There are four equations, one for each recurring labor category. The equations combine all the MACDAR aircraft observations referred to in Table 5.2, so they are of the "time-series cross-section" or "panel" form, with 64 total observations, one for each aircraft-lot combination (each year's production is defined as a lot, and the EMD data observation for each aircraft is another lot). The dependent variable is hours per aircraft divided by AUW. The independent variables include cumulative production to reflect learning effects and lot size (annual production except for EMD lots) to reflect rate effects. Weight is included to reflect the well-known effect in which hours per pound fall with pounds. There is a dummy variable for the EMD lot, both to reflect actual differences in the manufacturing of EMD aircraft[5] versus the aircraft built during production and to reflect the fact that the EMD aircraft were produced in a period of over a year, so the EMD coefficient also includes some rate effects.

Another dummy variable (called *MODEL*) represents significant model change. What constitutes a significant model change is obviously a judgment call. In these estimates, we assume that there were no such changes in the AV-8B or F-14 and that the following were significant model changes: F-15A/B to C/D and then to E; F-16A/B to C/D and then Block 30/32 to 40/42; and F/A-18A/B to C/D. If an aircraft was on its first model, the value of *MODEL* is zero; if on its second model (resulting from the first significant model change) the

[5]Not all EMD aircraft are production configuration, and the tooling used during development is different from that used in production.

value is one; and if on its third model (e.g., F-15E or F-16 Block 40/42 or higher) the value is two.[6]

We also include the variable $(WMCF)_l$ as defined above to represent the effect of material mix on labor hours. For this estimation, $(WMCF)_l$ values based on the MACDAR database were used. MACDAR's data were based on the cost ratios of Table 4.1—i.e., those reported in Resetar, Rogers, and Hess for the mid-1980s based on their industry survey.

The specific functional form we use is

$$\frac{H_{l,j}/Q_j}{W_j} = \alpha_l (WMCF)_l (Q_j)^{\rho_l} (W_j)^{\eta_l} (C_j)^{\beta_l} \exp\left[\omega_l (EMD)_j + \theta_l (MODEL)_j\right] \quad (5.4)$$

This represents four regression equations, one for each of the four recurring labor categories (indexed by l). The index j represents production lots and goes from 1 to 64. The variables are defined as follows:

$H_{l,j}$ = recurring hours of type l used in the production of lot j

Q_j = number of aircraft produced in lot j in units

W_j = average AUW for aircraft produced in lot j in pounds

$(WMCF)_l$ = weighted material cost factor associated with labor category l as defined above

[6]This definition of model change was proposed by SAIC analysts working on the MACDAR database and should be credited to them. We also did regressions in which two separate dummy variables were included, one for first-model change and one for second. The difference between the coefficients was not statistically significant, and the values were very close to each other. Thus, the specification of equal cost impact of each model change is supported by the data. Readers should be cautious in applying the model change variable to future programs in that future model changes would have to be comparable to the ones in this data set for the results to be applicable.

C_j = cumulative aircraft production for lot j, defined as the level of cumulative production, in units[7]

$(EMD)_j$ = a dummy variable whose value is unity if this is an EMD lot and zero otherwise

$(MODEL)_j$ = a dummy variable representing significant model changes, defined above

Estimated values of the coefficients are given in Table 5.5. The implied learning slopes, weight-sizing slopes, and rate slopes are given in Table 5.6. All the learning and weight coefficients are statistically significant at the 95 percent confidence level; only tooling shows a statistically significant rate effect at that confidence level. However, the rate coefficients for engineering, manufacturing, and quality assurance are marginally significant between the 73 and 87 percent confidence levels. (Owing to the proprietary nature of individual observations in MACDAR, we cannot show the residuals from the regressions.)

---

[7]Since a lot has more than one unit, lot cumulative production is ambiguous. Given that we are using unit learning theory, the correct cumulative production figure for each lot is

$$C = \left(\sum_{i=F}^{L} \frac{i^{\beta}}{(L-F+1)} \right)^{1/\beta},$$

where L and F are the first and last units of the lot and β is the natural logarithm of the learning slope divided by the natural logarithm of two. Since calculation of this depends on β but β is estimated from the data, an iterative procedure is used. First β is estimated using an estimated lot midpoint as the measure of cumulative production; then C is calculated and β is reestimated. In all results reported here, the second estimate of β was very close to the first and was in fact equal to two significant digits.

Table 5.5

Coefficient Estimates for Equation 5.4[a]

Coefficient	Coefficient Symbol	Engineering	Tooling	Manufac-turing	Quality Assurance
Constant term	α	4113	1502	295	3235
C	β	–0.49 (–5.1)	–0.37 (–4.9)	–0.33 (–11.7)	–0.23 (–4.1)
W	η	–0.53 (–3.1)	–0.41 (–3.2)	–0.24 (–5.3)	–0.72 (–7.3)
Q	ρ	–0.19 (–1.5)	–0.41 (–4.0)	–0.04 (–1.1)	–0.08 (–1.1)
EMD	ω	–0.51 (–1.6)	–0.36 (–1.6)	0.02 (0.2)	0.35 (1.7)
MODEL	θ	0.30 (2.2)	0.09 (0.9)	0.23 (5.7)	0.21 (2.6)
Standard error of estimate	SEE	0.43	0.34	0.12	0.30
Coefficient of determination	R^2	0.78	0.88	0.96	0.78

[a]The t-statistics are in parentheses.

Table 5.6

MACDAR Learning, Weight-Sizing, and Rate Slopes (in percentages)

Category	Learning	Weight Sizing	Rate
Engineering	71[a]	69[a]	88
Tooling	77[a]	75[a]	75[a]
Manufacturing	80[a]	85[a]	97
Quality assurance	85[a]	61[a]	95

[a]Statistically significant at the 95 percent confidence level.

EMD effects, while only marginally significant at best, are large in three categories. First, there is a large negative effect in engineering and tooling. This is probably because recurring engineering hours are used largely to *improve* production activity and to make minor system *changes*. Neither of these will occur much in EMD, when the system and its production mode are *initially* being developed using nonrecurring engineering hours. Similarly, much of recurring tooling during production is to refurbish tools, which are just being built during EMD. The large positive effect on QA hours probably reflects the working out of QA procedures on the shop floor during EMD and really represents a high learning rate between EMD and regular production.

The coefficients on the variable *MODEL* imply that each new model introduction causes a 30 percent increase in engineering, a 9 percent increase in tooling, a 23 percent increase in manufacturing, and a 21 percent increase in quality assurance. All but the tooling coefficient are significant at the 95 percent confidence level; the tooling coefficient is significant at the 63 percent level. In this model's functional form, after the one-time jump in labor associated with each model change, hours continue to follow a learning pattern with the same slope as before.

The $(WMCF)_l$ factors used in these estimates are shown in Table 5.7. As noted above, these were taken from the MACDAR database, which was itself based on the cost ratios of Table 4.1 from the earlier RRH study—i.e., on estimates for the late 1980s. The MACDAR database actually includes estimated values not of $(WMCF)_l$ but instead of

$$\sum_{m=1}^{M} S_m C_l^m \tag{5.5}$$

We used Equation 5.1 to convert these numbers to $(WMCF)_l$.

The reader will note that in Equation 5.4, the variable $(WMCF)_l$ does not have an exponent—or, put more precisely, an exponent of unity is imposed. We did this because the $(WMCF)_l$ factors were meant to modify CERs multiplicatively and were in fact constructed *specifically* for this use. Put another way, the $(WMCF)_l$ factors answer the question "What is the ratio of the cost of using material *X* to the cost of using aluminum," which implies an exponent of unity when modifying a CER equation. Therefore, our use of them here corresponds to the manner in which they were constructed and intended to be used.

Table 5.7

$(WMCF)_l$ Factors and σ Values Used for the Regression Results Shown in Table 5.5

$(WMCF)_l$ Factors

Aircraft	Engineering	Tooling	Manufacturing	Quality Assurance
AV-8B	1.14	1.41	1.24	1.30
F-14	1.07	1.34	1.15	1.17
F-15	1.08	1.35	1.20	1.18
F-16	1.02	1.08	1.04	1.07
F-18	1.09	1.31	1.16	1.20

σ Values

Aircraft	Engineering	Tooling	Manufacturing	Quality Assurance
AV-8B	0.43	0.87	0.66	0.60
F-14	0.50	0.94	0.66	0.60
F-15	0.43	0.87	0.74	0.60
F-16	0.19	0.69	0.45	0.60
F-18	0.43	0.87	0.66	0.60

Some cost analysts have argued that rather than impose an exponent of unity on the $(WMCF)_l$ factor in Equation 5.4, one should estimate a value for that exponent from the data, as is the case with other exponents in the equation. If such a procedure finds an exponent different from one, two explanations are possible. One is that the construction of the $(WMCF)_l$ factors was *systematically* flawed so that they are in fact related to cost nonproportionately. This would imply that estimating an exponent gives the best results. (Note that simple random errors in estimating the $(WMCF)_l$ factors would not lead to a nonunity estimate of the exponent; the flaws would have to follow a highly specific pattern to lead to a true nonunity coefficient.) The second explanation is that the true exponent on $(WMCF)_l$ is in fact unity, but some other variables that affect cost and are correlated with $(WMCF)_l$ are left out of Equation 5.4. If this is the case, imposing an exponent of unity provides better results.

We did estimate values for coefficients on the $(WMCF)_l$ factors to explore this issue. Table 5.8 shows the results, giving the point estimate of the exponent and the lower and upper bounds of the 95 percent confidence interval.

Table 5.8

Point Estimates and 95 Percent Confidence Interval Bounds for the Exponent on *(WMCF)$_l$*

Equation	Lower Bound	Point Estimate	Upper Bound
Engineering	–13.32	–9.70	–6.40
Tooling	–0.47	0.80	2.06
Manufacturing	0.29	0.94	1.59
Quality Assurance	0.27	1.97	3.67

We see that the exponent in the manufacturing equation is estimated to be very close to one. The exponents in the tooling and quality assurance equations are estimated imprecisely (i.e., they have wide confidence intervals) but do include unity in the confidence interval. Since these equations do not imply that a value of unity is statistically rejected, we accept unity as the value. The exponent in the engineering equation not only has the wrong sign but also has a highly implausible absolute value (a change of $[WMCF]_l$ from 1 to 1.5 changes hours by a factor of 50). We infer that this must be due to other variables that are left out of the equation and thus also accept the value of unity for this equation. In summary, we do not find evidence of systematic misestimation of the $(WMCF)_l$ that would lead us to reject functional form 5.4, in which the $(WMCF)_l$ values are entered multiplicatively as they were designed to be.

Some cost analysts have pointed out that in the manufacturing equation, if only observations for the *first* model of each aircraft are included, then an estimated exponent on $(WMCF)_{manu}$ is significantly different from one (i.e., these would be observations for which the variable $[MODEL]_j$ has a value of zero). We also explored this issue, and the results are given in Table 5.9, which shows the point estimate of the exponent on $(WMCF)_{manu}$ as well as the associated confidence interval for four samples. The first two, in the first column, are for the full sample and for just the original models. Full sample results are, of course, the same as those in Table 5.8. The bottom left box confirms the first sentence in this paragraph. If the manufacturing equation is estimated on observations only for the original models, the estimated coefficient on $(WMCF)_{manu}$ is significantly greater than one.

Table 5.9

Point Estimates and 95 Percent Confidence Interval Bounds for the Exponent on $(WMCF)_{manu}$ in Four Samples[a]

Sample	All Aircraft	Excluding F-16
All models	0.94 (0.3, 1.6)	0.65 (–0.6, 1.9)
Original aircraft model only (observations for which $[MODEL]_j = 0$)	2.03 (1.5, 2.6)	0.22 (–1.8, 2.3)

[a]The first entry in each cell is the point estimate; confidence interval bounds are in parentheses below.

However, the second column of Table 5.9 puts this in a different perspective. This column shows results for the same two samples but *excludes the F-16*. For these samples, although the coefficient on $(WMCF)_{manu}$ is estimated imprecisely, the value of unity is not rejected. We interpret these results as in fact a "missing variable" phenomenon. The F-16A/B was deliberately developed as a *low-cost* fighter, with extreme attention paid to affordability (see Lorell and Levaux, 1998, Chapter Five). It also happens to have the lowest $(WMCF)_{manu}$ value in the sample. Thus, "extreme attention to affordability" is a programmatic variable that is left out of Equation 5.4, and it happens to be correlated with low $(WMCF)_{manu}$. In the "original model only" equation, the coefficient on $(WMCF)_{manu}$ is therefore estimated at a high value owing to the correlation between it and the left-out variable. We base this interpretation on the results that are derived when the F-16 is not included in the sample at all.

Based on all this, we conclude that the best statistical procedure is indeed to use Equation 5.4 as is, applying the $(WMCF)_l$ factors as they were designed and intended to be. We conclude that there is insufficient evidence to infer that the $(WMCF)_l$ values were systematically misestimated, which would justify estimating a nonunity coefficient on the $(WMCF)_l$ variables.

The next chapter discusses the application of Equation 5.4 to estimating the cost of new aircraft.