INSTITUTE FOR DEFENSE ANALYSES

The Role of Inflation and Price Escalation Adjustments in Properly Estimating Program Costs: F-35 Case Study

Bruce R. Harmon
Stanley A. Horowitz, Project Leader

March 2016
Approved for public release;
distribution is unlimited.
IDA Document D-5489
Log: H 15-000415
Copy

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-14-D-0001, Project Number BA-7-3054, "Cost Indices Assessment," for the Director, Cost Assessment and Program Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
Thank you to Sarah K. Burns and David M. Tate for performing technical review of this document.

INSTITUTE FOR DEFENSE ANALYSES

IDA Document D-5489

The Role of Inflation and Price Escalation Adjustments in Properly Estimating Program Costs: F-35 Case Study

Bruce R. Harmon
Stanley A. Horowitz, Project Leader

Executive Summary

Background

The application of price indexes presents a substantial challenge in estimating the costs of new defense systems. The problem is twofold. First, the analyst must use a price index when normalizing historical cost data to a common point in time (where the normalized costs are referred to as "base year" (BY) dollars in defense acquisitions or, more generally, "real" dollars) so that these data can be used to help estimate the costs of future systems. Second, as budget requirements for future acquisitions are in "then-year" (TY) dollars (or more generally, "nominal" dollars), BY dollar estimates must be inflated to TY dollars using a price index. Using an inappropriate price index can introduce errors in both of these steps. In this paper we apply two sets of price indexes to a cost estimating problem—the F-35 Joint Strike Fighter (JSF) procurement program. The purpose is to help cost analysts and others involved in the acquisition process understand the impacts of different price indexes and to provide guidance in their choice.

The point of departure for this work is the analysis of price indexes presented in the Institute for Defense Analyses (IDA) Document D-5112, "Inflation Adjustments for Defense Acquisition."[1] The overall goal of that research was to identify a price index that is better than current indexes at supporting the Department of Defense (DoD)'s need for a sound basis for cost estimation. In particular, we explored an alternative "hedonic" approach for calculating price indexes for tactical aircraft.

In general, price indexes isolate changes in price due to factors other than quality changes. These changes can be categorized as changes due to general inflation, changes in the overall price level in the economy (subsequently just called "inflation"), and real price growth—price changes for a particular class of products relative to inflation. The combination of inflation and real price growth constitute price escalation, overall change in the price of a specified, constant quality, good, or service.

Hedonic price indexes isolate changes in price due to escalation by accounting for changes in quality over time. The process of relating prices to quality metrics using regression models that yield hedonic price indexes is closely related to standard techniques for developing cost estimating relationships (CERs). The hedonic price index approach uses costs in then-year dollars (the dependent variable in the regression

1 Bruce Harmon, Daniel B. Levine, and Stanley A. Horowitz, "Inflation Adjustments for Defense Acquisition," IDA Document D-5112 (Alexandria, VA: Institute for Defense Analyses, October 2014).

analysis) and derives a price index from the coefficients on variables reflecting the year of purchase. In CER development, the standard practice is to deflate costs to constant dollars (the dependent variable in the analogous regression) using a previously determined price index. The relevant cost metric in both cases is unit recurring flyaway (URF) costs.

For the current project, we develop a "Baseline" CER model, taking the hedonic approach, using data describing historical tactical aircraft programs that were available at the time of the F-35's late-2001 Milestone (MS) B decision. As the Baseline CER only provides hedonic price index values for the historical period through 2001, we calculate an index describing future escalation based on the relationship between the hedonic index and a general inflation index. Comparisons are made between F-35 URF costs estimated using the Baseline model and estimates using a CER model that we create, employing a more conventional approach (i.e., more typical of CER estimation) in which historical URF prices are escalated using a general inflation index. In estimating this CER model—referred to as the "Green Book" model—and making TY dollar projections, we apply a measure of general inflation published in the *National Defense Budget Estimates* "Green Book" by the Office of the Under Secretary of Defense (Comptroller) (OUSD(C)).[2] The models' estimates are compared with actual F-35 URF costs as well as with projections made by the JSF Joint Program Office (JPO).

Comparison of Models

The Baseline and Green Book CER models use the same data sample of fighter aircraft procured between fiscal year (FY) 1973 and FY 2002, and the same five quality variables. Both models also take into account cost/quantity effects (both learning and production rate). The Baseline model is estimated using 30 time-dummy variables to capture year-to-year escalation; taken together, the coefficients on the time-dummies define the hedonic price index. The Green Book model uses inflation rates published in the FY 2002 Green Book, implicitly assuming that inflation was equal to price escalation for fighter aircraft. The table below compares the Baseline and Green Book models.

[2] The *National Defense Budget Estimates* is commonly referred to as "The Green Book." It is a reference source for data associated within current DoD budget estimates.

Comparison of Models: FY 1973–FY 2002 Data

Metric	Baseline	Green Book
Price index used	Hedonic	Green Book
Number of data points	117	117
Parameters estimated	41	11
Number of quality variables	5	5
Adjusted R^2	.97	.84
Learning curve slope	84.5%	88.1%
Annualized escalation rate: FY 1973–FY 2002	7.4%	4.5%
Projected annual escalation: FY 2002–FY 2013	3.8%[a]	2.1%[b]

[a] Based on the relationship between the Hedonic and Green Book inflation rates.

[b] Extrapolated from projections in the FY 2002 Green Book.

Using the index prescribed in the FY 2002 Green Book results in a significantly worse model fit (lower R^2). The learning curve slope is substantially shallower for the Green Book model (88 percent vs. 84 percent)—consistent with lower Green Book escalation. Systematically lower relative constant dollar costs in the earlier years mean that the estimated learning effect is blunted; this also results in lower estimated first unit (T1) costs.[3] The remaining model parameters are more likely to reflect the true underlying relationship between costs the other cost drivers when a price index specific to tactical aircraft is used to account for escalation. This is particularly true for quality variables that have a systematic relationship with time,

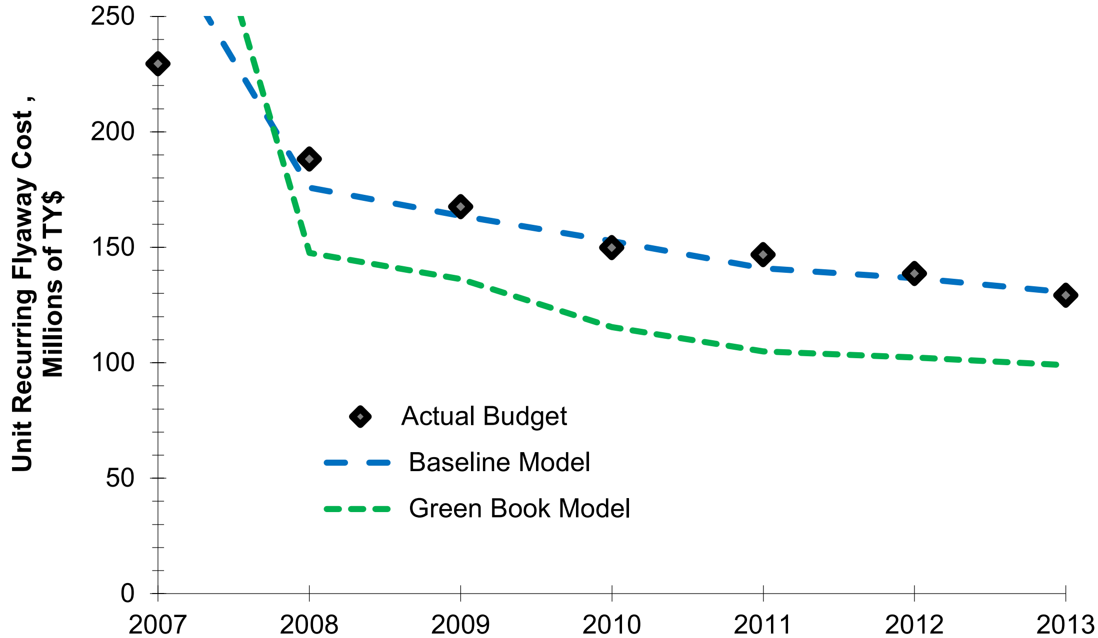
Model Estimates and F-35 Costs

We compare estimates from the two models with estimates and actual costs from the F-35 program.

In the first example, we apply the models to the aircraft design and program as it was defined at MS B. We compare the models' BY 2002 estimates with the MS B JPO estimates. Both models produce estimates above the JPO URF estimate ($44.6 million and $47.3 million vs. $40.6 million), with the Baseline model producing the highest estimates. They are still below later cost estimates; there are many elements of later F-35 cost growth that are not captured in this set of model estimates. These include changes to the F-35 program and aircraft design.

3 Prices fall over time due to learning. We expect that input prices will rise over time, pulling aircraft prices up. If we understate the rate of increase in input prices, the amount of learning will appear to be lower than it really is. Consider a simple example. Suppose that in the absence of input price changes, learning would reduce aircraft prices by one percent per year. Suppose further that we observe aircraft prices rising at two percent per year. This implies that input prices are rising three percent per year. If we erroneously assume that input prices rise two percent per year, we will calculate learning to be zero.

The next step is a comparison of F-35 actual TY budget costs with model estimates, taking into account changes in the F-35 aircraft design and program. In this example, escalation is modeled based on information that was available in FY 2002. The figure below shows comparisons between model estimates and F-35 actual costs.



Comparison of Model Estimates with Budget Actuals, All F-35 Variants

This example shows the close correlation between the Baseline model estimates and the actual costs. Although the model inputs reflect the latest F-35 aircraft and program parameters, in terms of the structure of the model and inflation projections, the models are defined by the information that was available at MS B. The lower estimates from the Green Book model are due to two factors: the underestimates of escalation from FY 2002 to FY 2013 and biases introduced into the model parameters because of underestimates of escalation in the historical period.

We demonstrate the effect of different escalation methodologies using top-level CER models. Cost analysts usually build up their estimates from a more detailed level. However, insights from top-level models regarding the proper application of price indexes, for both normalizing historical data and making projections, are also valid in more typical cost estimating applications. For example, escalation rates for raw material inputs, propulsion systems, electronic components, and labor inputs are likely to be different from general inflation. In our last example, we calculated overall escalation rates implied in the JPO estimates for the rest of the F-35 program; we found the implied escalation rates to be consistent with those projected using values from the historical hedonic price index. Projections based on the standard CER methodology using general inflation would underprice the program by 30 percent as shown in the figure.

Contents

1. Introduction

A. Background

The application of price indexes presents a substantial challenge in estimating the costs of new defense systems. The problem is twofold. First, the analyst must use a price index when normalizing historical cost data to a common point in time (where the normalized costs are referred to as “base year” (BY) dollars in defense acquisitions or, more generally, “real” dollars), so that these data can be used to help estimate the costs of future systems. Second, as budget requirements for future acquisitions are in “then-year” (TY) dollars (or more generally, “nominal” dollars), BY dollar estimates must be escalated to TY dollars using a price index. Using an inappropriate price index can introduce errors in both of these steps. In this paper, we apply two sets of price indexes to a cost estimating problem—the F-35 Joint Strike Fighter (JSF) procurement program. The purpose is to help cost analysts and others involved in the acquisition process understand the impacts of different price indexes and to provide guidance in their choice.

In general, price indexes isolate changes in price due to factors other than quality changes. These changes can be categorized into changes due to general inflation, changes in the overall price level in the economy (subsequently often just called “inflation”), and real price growth—price changes for a particular class of products relative to inflation. The combination of inflation and real price growth constitute price escalation—overall change in the price of a specified, constant quality, good or service.

The point of departure for this work is the analysis of escalation indexes presented in Harmon, Levine, and Horowitz (2014) (hereafter “D-5112”).[1] The overall goal of that research was to identify a price index that is better than current indexes at meeting the Department of Defense (DoD)’s need for a sound basis for cost estimation. In particular, we explored an alternative “hedonic” approach for calculating price indexes for tactical aircraft. In this analysis, we used updates to the hedonic model presented in D-5112 in the F-35 example.

B. The F-35 Cost Estimating Problem

The F-35 program has experienced significant program cost growth since its October 2001 Milestone (MS) B decision that initiated Engineering and Manufacturing

[1] Bruce R. Harmon, Daniel B. Levine, and Stanley A. Horowitz, “Inflation Adjustments for Defense Acquisition,” IDA Document D-5112 (Alexandria, VA: Institute for Defense Analyses, October 2014).

Development (EMD). A substantial portion of this cost growth has been in its unit recurring flyaway (URF) cost, with much of this attributed to the incorrect application of price indexes.[2] Given the tactical aircraft focus of the Institute for Defense Analyses' (IDA's) previous hedonic models, the F-35 makes for a suitable case study.

We used information available at MS B to develop models for exploring the effects of escalation adjustments on estimated F-35 URF costs. The resulting estimated costs can then be compared to several benchmarks, including cost estimates produced by the JSF program office (JPO) and observed URF costs for F-35s procured from 2007 through 2013. From this exercise, we draw lessons for future cost estimating practice.

---

[2] Scott A. Arnold et al., "WSARA 2009: Joint Strike Fighter Root Cause Analysis," IDA Paper P-4612 (Alexandria, VA: Institute for Defense Analyses, June 2010).

2. Hedonic Price Index Models for Tactical Aircraft

A. Introduction

In this chapter, we review past work on hedonic price index models and present updates developed specifically for the F-35 cost estimation problem. The estimation of the hedonic indexes for tactical aircraft builds upon tools that cost estimators have used for years. The basic setup is:

nominal system unit price = f(year, quality variables, other control variables)

The hedonic index application has commonalities with cost estimating relationships (CERs), which also model system costs as a function of quality variables, and cost/quantity relationships (primarily learning), which are control variables in the hedonic model. The hedonic index estimation differs from past cost estimating practice in that the price index is estimated simultaneously with other model parameters and the dependent variable is expressed in TY (nominal) dollars. In CER development, adjustments needed to normalize historical cost data to BY dollars used as the dependent variable are often performed using a general deflator based on an index of overall inflation, such as that published in the *National Defense Budget Estimates* "Green Book" by the Office of the Under Secretary of Defense (Comptroller) (OUSD(C)).[3] For commodities such as tactical aircraft, a given observed price may reflect both inflation and relative price changes, including those due to variation in the quantity purchased. Typically normalization to a common quantity (e.g., first unit or 100th unit)[4] is performed using BY dollars prior to CER estimation. Thus, another unique aspect of our modeling is the simultaneous estimation of CER and learning curve parameters, as well as production rate effects.

The hedonic analysis described in D-5112 used the direct time-dummy variable approach formulated by Triplett, an early developer of hedonic analysis.[5] The update to the earlier analyses also used this approach, along with the same set of explanatory

---

3 The *National Defense Budget Estimates* is commonly referred to as "The Green Book." It is a reference source for data associated within current DoD budget estimates.

4 Although unit prices are also sensitive to production rate, this typically has not been taken into account.

5 Jack E. Triplett, *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products* (Paris, France: Organisation for Economic Co-operation and Development, 2006).

variables (presented in Table 1). Five quality variables describe the aircraft, two quantity variables capture the cost effects of learning and production rate, and the time-dummy variables identify each fiscal year in which the aircraft were procured. The hedonic index is defined by the expression $b_t^{D_t}$,where D_t is a 1/0 dummy variable with a value of 1 for fiscal year t, and b_t is the estimated index for that year. BY dollars are calculated as $BY\ dollars_t = \frac{TY\ dollars_t}{b_t^{D_t}}$. In the application of the Green Book index, the index (where the base year value equals 1) replaces the $b_t^{D_t}$expression in calculating BY dollars.[6]

Table 1. Explanatory Variables

Quality variables
Empty weight in pounds
Maximum speed in knots
Advanced materials as percentage of structure weight
Dummy variable for 5th generation aircraft[a]
Dummy variable for Short Take-Off and Vertical Landing (STOVL) aircraft[b]
Quantity variables
Cumulative production
Lot size (number of aircraft produced in a year)
Time-dummy variables

[a] 5th-generation aircraft are characterized by stealth, internal weapons carriage, avionics with information fusion and support of net-centric operations. In the D-5112 sample, the F-22 and F-35 A/B/C were classified as 5th-generation aircraft; in the update, we added the F-117.

[b] The A/V-8B and F-35B, aircraft with STOVL capability needed for operations from small aircraft carriers and short unimproved airfields.

The database used in regression estimation contains pooled cross-section and time-series data, often called "panel data" in the econometrics literature, where each panel is an aircraft program. The cost metric of interest is the unit recurring flyaway cost (URF). In D-5112, the time series included forty fiscal years (FYs 1973–2013), with 2012 as the base year; the cross-sections (panels) consisted of the eleven aircraft programs' original designs plus derivatives of these designs from series or block changes. In model estimation, the quality changes associated with the series/block changes are captured in the changes in empty weight over time. Production rate effects were calculated by estimating the annual fixed cost for each program.[7] Learning spillovers due to commonality between the EA-18G and F/A-18E/F and between the F-35 variants were

[6] If the values for the Green Book escalation index were the same as the hedonic price index, all other model parameters would also be the same.

[7] Fixed costs for each program were estimated as a function of the estimated maximum variable costs.

included in the model.[8] We also accounted for loss of learning due to series/block changes.[9]

B. Updating Hedonic Price Index Models for Tactical Aircraft

For the current analyses, we made multiple changes to the previous work, including several versions of the model meant to capture different aspects of the F-35 cost estimating problem. Our primary focus is on the "Baseline" F-35 model; the intent was to use the vintage of information available for the MS B (October 2001) cost estimate. As the FY 2002 budget materials were released earlier in 2001, we used data through FY 2002. Eliminating the newer data means that we dropped the EA-18G from the data sample along with the three F-35 variants (F-35A, F-35B, and F-35C); also, the F-22A program is truncated. This left the F-22A as the sole 5th-generation aircraft with only two data points (2001 and 2002). In order to include another 5th-generation aircraft, we added the F-117A[10] to the updated sample. Figure 1 shows all of the potential data as well as the data excluded from the Baseline case.

---

8 Learning spillovers are captured by estimating parameters that assign some portion of the cumulative quantity across related aircraft.

9 This is accounted for by parameters that decrement cumulative quantity at each block change.

10 Stealth technology is the prime feature of 5th-generation aircraft and the F-117. The F-117 differs from newer examples of 5th-generation aircraft in having less sophisticated electronic systems.

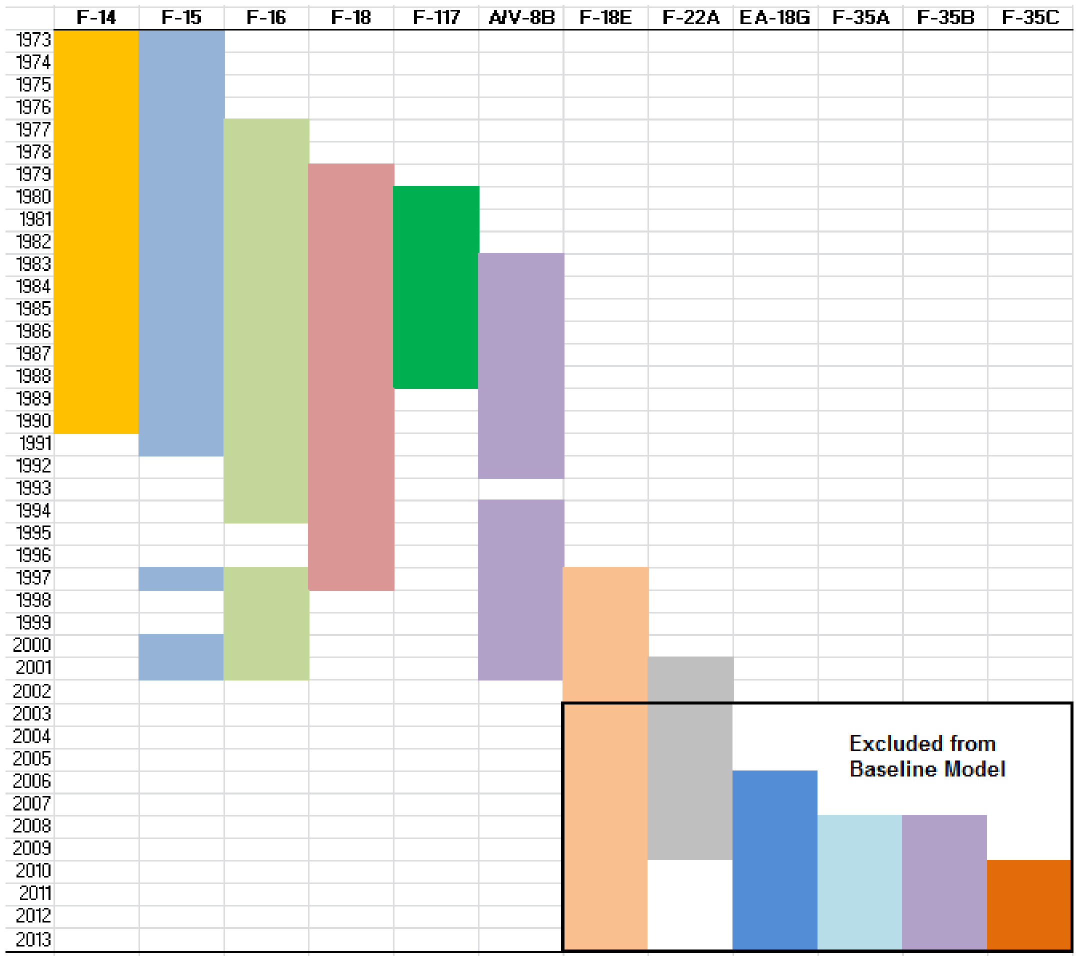


Figure 1. Data Coverage by Aircraft Program and Year

In addition to the original series aircraft, derivative follow-on aircraft were relevant for the F-14A (F-14A+ and F-14B), F-15A (F-15C, F-15C MSIP, and F-15E), F-16A (F-16C Blocks 25/30/50), F/A-18A (F/A-18C and F/A-18C Night Attack), and A/V-8B (A/V-8B Night Attack and A/V-8B Radar).[11] As these derivative aircraft were produced serially, they were included in the same panel as the original design. We use 2002 as the BY price index; this was also the BY for the F-35 MS B estimates and the associated URF goal.

In addition to the Baseline model, we estimated other model variations to address different aspects of the F-35 cost estimating problem. The Green Book model replaces the statistically estimated hedonic index with the procurement budget index published in the FY 2002 *National Defense Budget Estimates*. This would be more typical of the

---

[11] Military aircraft are described by Mission-Design-Series (MDS). For the F-14A, for example, the mission is fighter (F), the design is 14, and the original series is A. The aircraft in column headings of Table 1 are new designs, with the exception of the F/A-18E, which was a major change from the previous F/A-18s; the three F-35 variants are being built for different missions and produced in parallel.

approach used in CER estimation. All hedonic model variations follow the “Full CER Hedonic Model” approach from D-5112. We also estimated a “Full Information” model, using complete actual data through 2013. The purpose of that model is to provide a close comparison with the model included in D-5112.[12] A slight modification of this model excludes the F-35—the “Full Information less F-35” variation provides hedonic index values through 2013 without using any information from F-35 program cost experience. Unlike in the D-5112 and Full Information models, the Baseline model does not generate price index values from 2003 through 2013; instead, a methodology is presented in which model results are extrapolated to produce estimated index values through 2013.

C. Model Estimation and Results

This section presents regression results for the different model variations. Comparisons are shown between these models and the Full CER Hedonic Model described in D-5112. As the functional form of the models is the same, we do not repeat the detailed exposition presented in D-5112—instead, we highlight the differences in the regression results.

We estimate the model parameters using maximum likelihood estimation. The models are fit using the nonlinear optimization package within Microsoft Excel. The distribution of errors is assumed to be multiplicative/lognormal—this is analogous to estimating a log-log regression using linear regression.

Table 2 presents key regression metrics and parameter estimates for the five models.

[12] The model in D-5112 only used data through 2012 and did not include the F-117A.

Table 2. Comparison of Regression Results

Metric	FY 1973–FY 2002 Baseline	FY 1973–FY 2002 Green Book	FY 1973–FY 2013 D-5112	FY 1973–FY 2013 Full Information	FY 1973–FY 2013 Full Information Less F-35
Price index used	Hedonic	Green Book	Hedonic	Hedonic	Hedonic
Number of data points	117	117	150	159	143
Parameters estimated	41	11	55	54	53
Adjusted R^2	.97	.84	.97	.97	.97
Standard error	.09	.20	.09	.09	.09
Quality coefficients					
Empty Weight[a]	0.78	0.75	0.83	0.84	0.81
Maximum Speed[a]	0.29	0.08	0.30	0.28	0.26
Advanced Materials[b]	1.95	1.86	1.67	1.63	1.77
5th-Generation[b]	1.24	1.44	1.11	1.16	1.14
STOVL Capability[b]	1.00	1.00	1.10	1.05	1.00
1st unit cost (T1), FY02$M					
F-14A	240	119	271	261	261
F-15A	196	94	218	207	209
F-16A	97	50	109	104	104
F/A-18A	140	73	158	153	153
F-117A	187	128	189[c]	192	192
A/V-8B	81	49	94	88	87
F/A-18E	197	101	219	210	213
F-22A	370	212	368	367	365
F-35A	235[c]	144[c]	233	234	233[c]
F-35B	246[c]	154[c]	267	259	246[c]
F-35C	278[c]	169[c]	276	277	277[c]
Learning curve slope	84.5%	88.1%	83.9%	84.1%	84.1%
Escalation growth rate: 73–02	7.4%	4.5%	7.6%	7.5%	7.5%
Escalation growth rate: 02–13	N/A	2.1%[d]	3.6%	3.5%	3.2%

[a] The coefficients on these variables enter the model in the form x^b.

[b] The coefficients on these variables enter the model in the form b^x.

[c] Out-of-sample estimates.

[d] Extrapolated from projections in the FY 2002 Green Book.

The regression fits for the models in which a hedonic index is estimated are comparable. Restricting the index to that prescribed in the 2002 Green Book results in a significantly worse model fit. The learning curve slopes are similar for the hedonic models, but the slope is substantially shallower for the Green Book model (88 percent vs.

84 percent)—again, this is consistent with the embedded underestimation of escalation when normalizing the data to constant year dollars. Systematically lower constant dollar costs in the earlier years mean that the estimated learning effect is blunted. The steeper learning slope is also consistent with values of fighter/attack aircraft learning curve coefficients estimated using labor hour costs in previous studies.[13]

Coefficients on weight, speed, and materials composition are relatively stable across the models and are consistent with those reported in past CER studies.[14] Unit prices increase with weight, maximum speed, and more advanced materials. The one exception is the speed variable in the Green Book model—as the aircraft with the highest maximum speeds (the F-15 and F-14) appear early in the sample, the underestimates of aircraft inflation associated with the model tend to bias its parameter estimate downward. Estimates for the 5th-generation and STOVL aircraft effects change some when the F-117 is introduced into the sample. The 5th-generation factor increases from 1.11 to 1.16, while the STOVL factor decreases from 1.10 to 1.05. When the F-35 is excluded from the regression, the STOVL factor goes to 1.00—this reflects the influence of the F-35B (which is a 5th-generation STOVL aircraft), with the A/V-8B the only other STOVL aircraft in the sample.[15] The range of 5th-generation premiums for the hedonic models is generally consistent with values from an earlier IDA paper on the cost of stealth,[16] although the 1.24 factor for the Baseline model is somewhat higher than expected. The 1.44 factor estimated with the Green Book model is clearly too high—the bias is a mirror image of the maximum speed coefficient, where underestimated escalation and newer 5th-generation aircraft interact. Thus, if there is a relationship between time and the values of the quality variables, a systematic bias in the price escalation used will result in a related bias in the coefficients on the quality variables.

---

[13] S. A. Resetar, J. C. Rogers, and R. W. Hess, "Advanced Airframe Structural Materials: A Primer and Cost Estimating Methodology," R-4016-AF (Santa Monica, CA: RAND Corporation, 1991); Obaid Younossi, Michael Kennedy, and John C. Graser, "Military Airframe Costs: The Effects of Advanced Materials and Manufacturing Processes," MR-1370-AF (Santa Monica, CA: RAND Corporation, 2001); and Bruce R. Harmon, "Cost Estimating Techniques for Tactical Aircraft Manufacturing Labor," IDA Paper P-4490 (Alexandria, VA: Institute for Defense Analyses, May 2010), Unclassified/PI/LR.

[14] Resetar, Rogers, and Hess, "Advanced Airframe Structural Materials;" Younossi, Kennedy, and Graser, "Military Airframe Costs;" Harmon, "Cost Estimating Techniques;" and Bruce R. Harmon, J. Richard Nelson, and Scot A. Arnold, "Unit Cost Implications of New Materials: Preliminary Analyses of Airframe Experience (Revised)," IDA Document D-908-REV (Alexandria, VA: Institute for Defense Analyses, September 1991).

[15] This does not mean that STOVL capabilities are free in the model; holding all else equal, STOVL aircraft will tend to be heavier and have more advanced materials than a conventional aircraft. Also note that in model estimation, the coefficient on the STOVL dummy was restricted to ≥ 1.00.

[16] Nelson et al., (U) "Cost Estimating for Modern Combat Aircraft: Adjusting Existing Databases and Methods to Include Low-observable Cost Considerations," IDA Paper P-3528 (Alexandria, VA: Institute for Defense Analyses, June 2001), Secret/PI/LR.

Also note that the analogous cost drivers in the historical studies are usually estimated using labor hour data, eliminating the possibility of bias from price escalation.

Estimated first unit variable costs (T1s) for each initial Mission-Design-Series (MDS) (usually the "A" series) are calculated using the quality coefficients, the regression intercept, and the values of the quality variables for each MDS. Table 2 (on page 8) shows the T1s for all relevant MDS, including "out-of-sample" cases in which the MDS was not used in model estimation. These cases are the F-35 variants, with the exception of the F-117A, which was not used in estimating the D-5112 model. For the models using the hedonic indexes, the out-of-sample estimates were close to the values calculated using the models that included those MDS. The exception is the F-35B, where the more complex STOVL capabilities were not well captured in the models not using the F-35 data. Even in this case, the out-of-sample F-35B T1s are only around 5 percent lower than the estimates from the other hedonic models. The T1s from the Green Book models are all substantially lower than those from the hedonic models. This is consistent with the shallower learning curve for the Green Book model, where the real prices of the initial lots are systematically underestimated because of biased escalation. Figure 2 shows the escalation indexes for a selection of the regression models.[17] Also included for comparison is the latest (FY 2015) Green Book index.

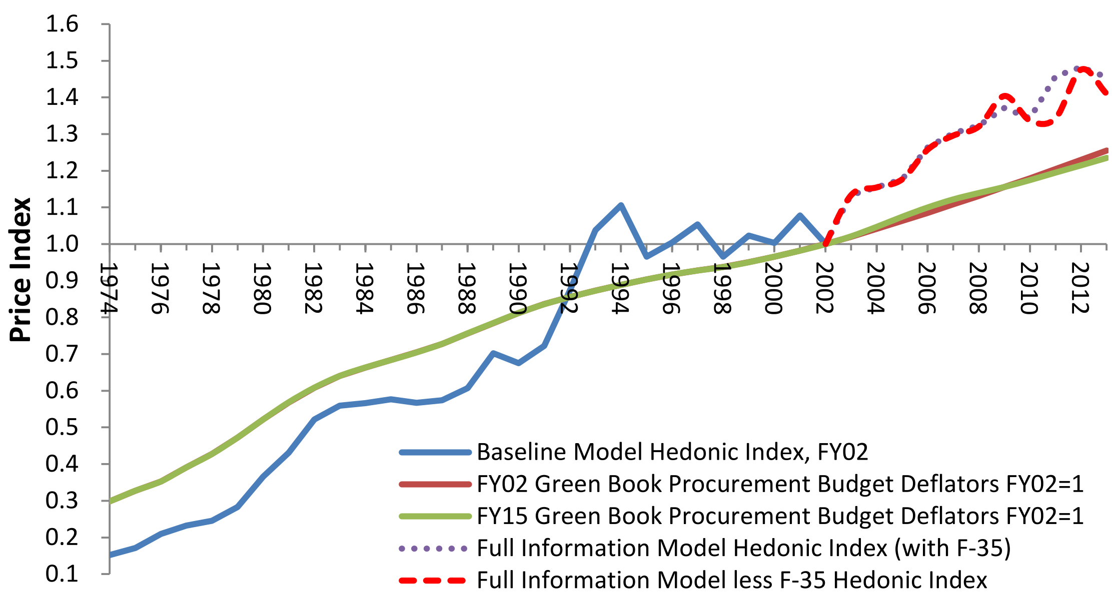


Figure 2. Comparison of Price Indexes

---

[17] The published FY 2002 Green Book deflators only include projections through FY 2007—beyond FY 2007, we use the 2.1 percent inflation rate evident in the FY 2004 to FY 2007 projections.

These indexes are portrayed in the price growth rates shown in Table 2. Of most interest for the F-35 estimating exercise are the Baseline and Green Book models. The other models are included for comparison purposes as well as to provide escalation estimates through 2013. There is no 2002–2013 escalation associated with the Baseline model; one of the goals of our analyses is to suggest a methodology for extrapolating forward growth rates from the Baseline model hedonic index. Also note how little the Green Book inflation changed from the FY 2002 forecasts (including extrapolations from FY 2007 to FY 2013) through the actuals reflected in the latest FY 2015 values. The fits of the Baseline and Green Book models to the data are shown in Figure 3 and Figure 4; also shown are the post-2002 data not included in model estimation.

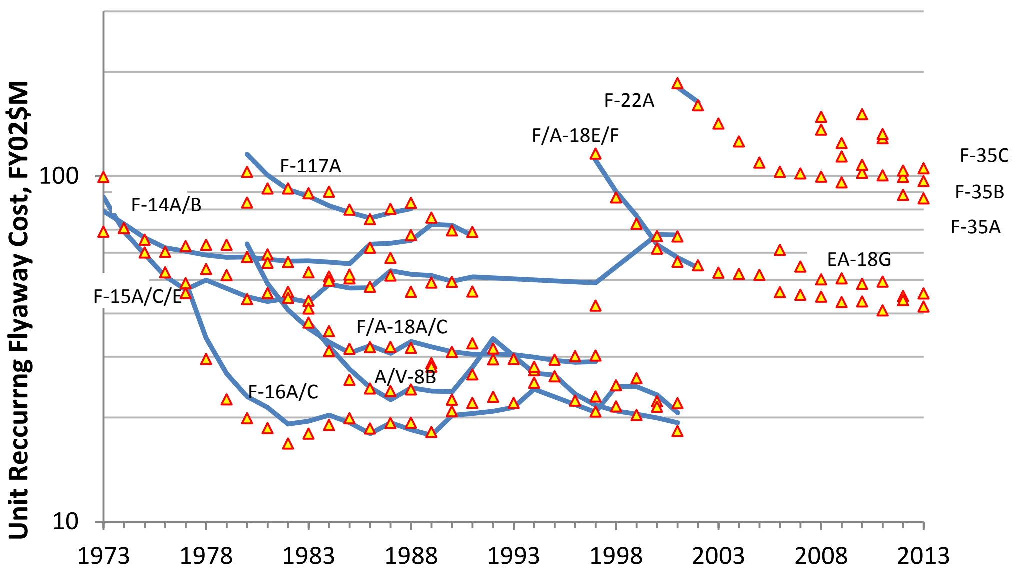


Figure 3. Fit of the Baseline Model to the Data

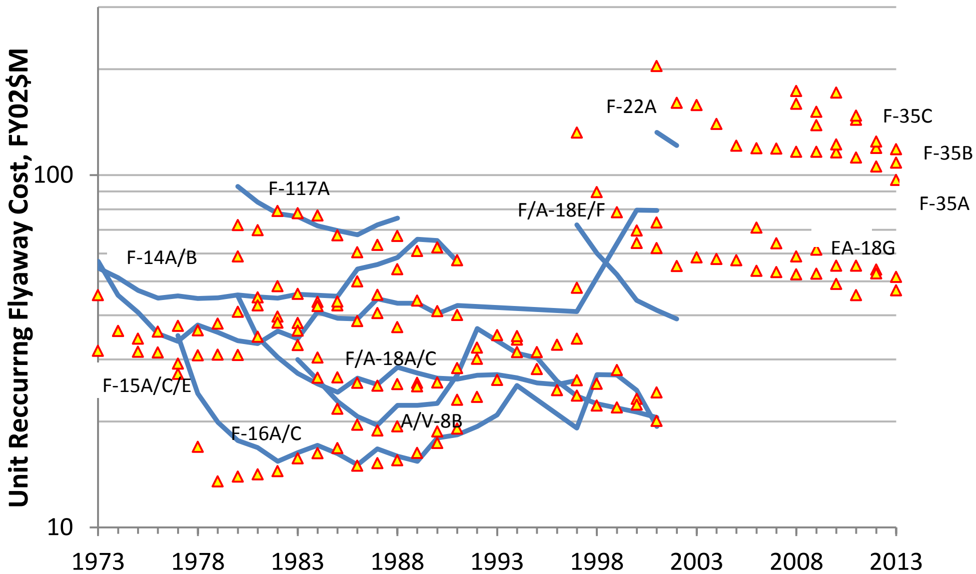


Figure 4. Fit of the Green Book Model to the Data

Figure 4 shows that normalizing the data using the Green Book index results in a constant-dollar cost data set and associated model that systematically underestimates costs in the earlier years and overestimates costs in the later years. In addition to introducing bias in the quality parameters, using the Green Book index also results in a shallower learning curve. This behavior is not evident in the Baseline model. It is clear in both the distortion of the parameter estimates and the systematic errors in estimating the actual data that a naïve application of price indexes can be problematic.

3. F-35 Cost Estimating Applications

A. Introduction

We compare F-35 URF estimates generated by the Baseline and Green Book models against three sets of benchmarks. They include:

- MS B program cost estimates and subsequent cost estimates associated with the 2009 “Nunn-McCurdy” unit cost breach,[18] in *BY 2002* dollars;
- Actual *TY* dollar budget values for the 2008–2013 fiscal year lots; and
- The latest program cost estimate as reported in the December 2013 selected acquisition report (SAR), reported in *TY* dollars.

To do this, we use the Baseline and Green Book models to produce BY 2002 cost estimates for each scenario. For comparisons with the TY actuals and estimates we use either an index calculated from the historical hedonic index (“projected hedonic index”) or the Green Book index. The BY 2002 estimate comparisons demonstrate the effect of different price indexes on the structure of the CER model, while the TY dollar estimates also show the effect of the different indexes in projecting BY estimates forward.

B. F-35 MS B and Nunn-McCurdy Breach Estimates

MS B estimates are the initial benchmarks used for budgeting and for calculating program cost growth. As both models take into account production rate and learning, they can produce an analog of the MS B estimate using the quantities and production schedule associated with the October 2001 program. The IDA model estimates in this application do not carry explicit assumptions regarding future (post-2002) escalation—they are in BY 2002 dollars as directly calculated by the model. Figure 5 shows comparisons between the MS B URF estimates (all F-35 variants combined) and those generated by the Baseline and Green Book models using MS B input values.

---

[18] A Nunn-McCurdy unit cost breach (10 U.S. Code § 2433a, “Critical cost growth in major defense acquisition programs”) occurs when cost growth in program or acquisition unit costs surpasses 15 percent.

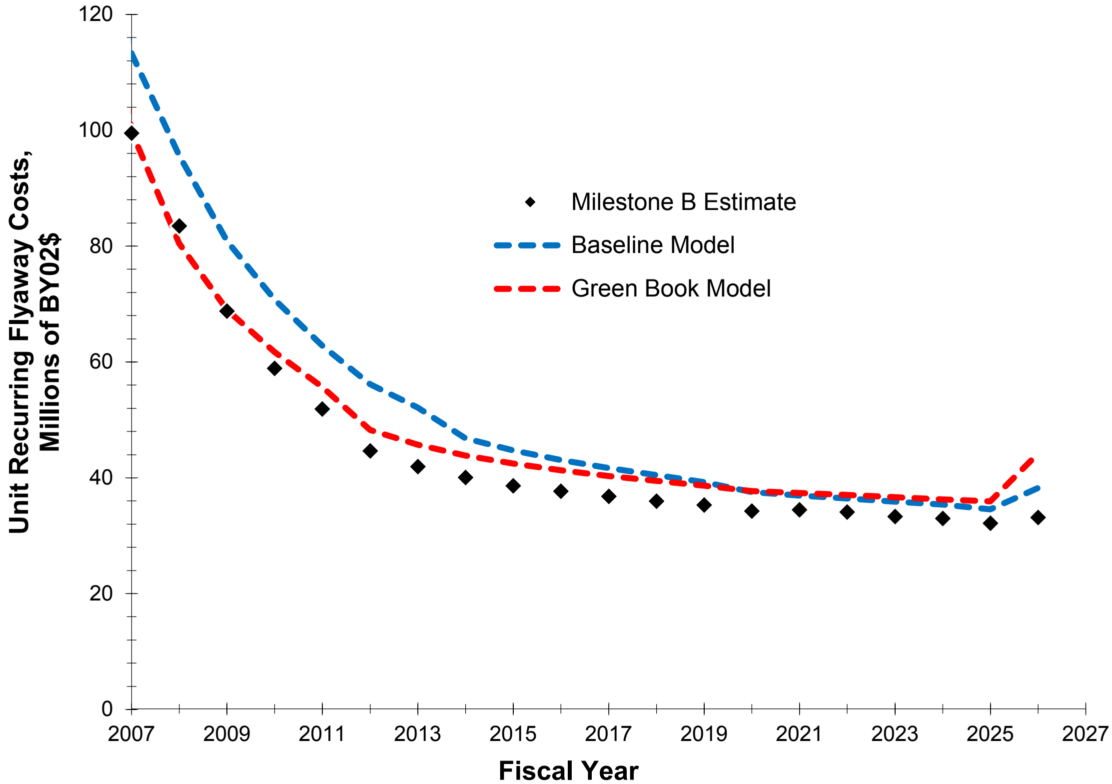


Figure 5. Comparisons of MIlestone B and Model Estimates for All F-35 Variants

The estimates from the two models converge as a result of the shallower learning slope of the Green Book model. Both models produce estimates above the program MS B URF estimate. However, they are substantially below the 2009 SAR estimates that triggered the Nunn-McCurdy breach. Many elements of F-35 cost growth are not captured in the above model estimates. Data from Arnold et al.[19] allow us to isolate and deconstruct the URF portion of the cost growth.[20]

Weight growth in all F-35 variants was a driver of cost growth between MS B paper designs and the current designs reflecting the aircraft as produced. Almost all weight growth attributable to redesign was evident by the 2009 Nunn-McCurdy breach and reflected in the production lots.[21] As empty weight is an input to the models, the weight growth must be taken into account when comparing model outputs to the MS B estimates and subsequent cost growth. Another change affecting cost model application is the

[19] Scott A. Arnold et al., "WSARA 2009: Joint Strike Fighter Root Cause Analysis," IDA Paper P-4612 (Alexandria, VA: Institute for Defense Analyses, June 2010).

[20] The 2009 F-35 Nunn McCurdy breach was driven by cost growth in EMD and nonrecurring procurement as well as by URF.

[21] We used the latest available weight status to characterize the F-35 variants as procured. These values were fixed across the procurement lots and do not include any weight growth margin.

decrease in commonality between variants (F-35A/F-35B/F-35C) since MS B. Current commonality is reflected in the "spillover" parameter affecting learning across variants estimated as part of the Full Information model. The cost effects of commonality have been estimated by the JSF program using a detailed assignment of the learning quantities depending on common component applications. As we cannot reproduce such a detailed analysis, we make use of the spillover parameter instead—for the MS B estimate we increase its value to reflect higher commonality assumed at that point.

Table 3 shows the MS B URF estimate, a buildup of cost growth drivers to the 2009 estimate as derived from Arnold et al., and comparisons with the model estimates. Model estimates presented include calculations with MS B inputs, and with inputs reflecting contemporary values for empty weight and commonality (learning spillovers).

Table 3. F-35 Program Growth Track from Milestone B to 2009 SAR and Model Estimate Comparisons

	F-35 Program URF Cost, in Millions of BY 2002$			
Metric	**Cost Growth Increment**	**Cumulative Cost Growth**	**Baseline Model**	**Green Book Model**
MS B Estimate		40.7		
Major Subcontractor Fee	1.5	42.2		
Change in Materials Manufacturing Efficiency	3.0	45.2	47.3[a]	44.6[a]
Design-Negated Affordability and Production Efficiency Plans	3.0	48.2		
Aircraft Weight Growth	3.0	51.2	52.1[b]	48.4[b]
Change in Buy Profiles (2009 SAR)	2.5	53.7		
Escalation Rates (2009 SAR Estimate)	7.0	60.7		

[a] MS B weight and commonality

[b] Contemporary weight and commonality

We orient the model outputs in the table to reflect how they relate to the cost growth elements from the MS B estimates. Elements that represent underestimates based on a departure from business as usual (i.e., the historical database) are included above the model estimates calculated with the MS B weight and commonality assumptions. The estimates reflecting updated weight and commonality are in line with cost growth through the Aircraft Weight Growth row. Not accounted for in this application of the IDA model estimates are cost increases due to buy profile changes (a reduction in quantities and a stretch-out of the procurement schedule) and a misapplication of escalation rates for future costs.[22] The last cost growth element is informative of our research question. Instead of using contractor-specific labor rate escalation, the JPO used OUSD(C) Green Book inflation when converting constant dollar estimates to TY dollar estimates.

From Arnold et al.:

> However, at the time of Milestone B, the Defense Contract Management Agency (DCMA) and Lockheed Martin had already agreed to a Forward Pricing Rate Agreement (FPRA) that increased rates more than the OUSD(C) escalation indices...therefore, the fully burdened labor rates turned out to be significantly higher than those used in the JPO Milestone B [estimate].[23]

The preferred methodology reflected in the 2009 JPO cost estimate is to escalate estimated constant year costs to TY dollars using escalation rates appropriate to the different cost elements. The OUSD(C) index is then used to de-escalate the TY dollars to BY dollars, which are, in turn, reported in the SARs and used as a basis for cost growth calculations. This correction of the original methodology is responsible for the $7 million unit cost growth due to escalation rates shown in Table 3. Analogous steps are not reflected in the BY 2002 model estimates in Table 3; thus, the constant year model estimates presented for comparison are conceptually similar to the JPO's MS B estimates, reflecting the same error.[24] In the next sections, we focus on model-generated TY estimates in the context of more up-to-date F-35 estimates.

C. F-35 Actual Budget Values

This section compares model-generated estimates with actual historical costs. The emphasis is on the results from the Baseline model. The budget experience is taken from Navy and Air Force President's Budget (PB) Justification Books, "Exhibit P-5, Cost Analysis" sheets. In collecting these data, we used the values in the latest PB in which they appeared; e.g., for the FY 2013 lot, we used data presented in the FY 2015 PB submission. For this exercise, we used the unadjusted TY URF values.

For the Baseline model, we developed the projected hedonic index to generate TY estimates through FY 2013. We also included results for the Green Book model, where the FY 2002 Green Book index (including extrapolations through FY 2013) is applied. We used the hedonic indexes generated by the Full Information and Full Information

---

[22] Both of these effects are addressed in the later benchmark comparisons.

[23] Arnold et al., "WSARA 2009: Joint Strike Fighter Root Cause Analysis," 12.

[24] Although it would be possible to capture the 2009 procurement profile and escalation application effects in the modelling exercise, we only address these issues in the context of more up-to-date cost data.

Less F-35 models for comparison purposes. For model inputs, we used contemporary values for the quality variables and the procurement profiles reflected in the budget data.

The projected hedonic index is based on the relationship between the FY 2002 Green Book and Baseline hedonic indexes; it has the advantage of using only information through 2002 while taking into account the systematically higher escalation rates associated with the hedonic indexes vs. the Green Book rates.

To calculate the projected hedonic index, we first define the relationship between the Green Book index and the hedonic index using data through 2002 as estimated by the Baseline model. Given the year-to-year volatility of the hedonic index, we do this by comparing 10-year compounded annual growth rates. These data are shown in Figure 6.

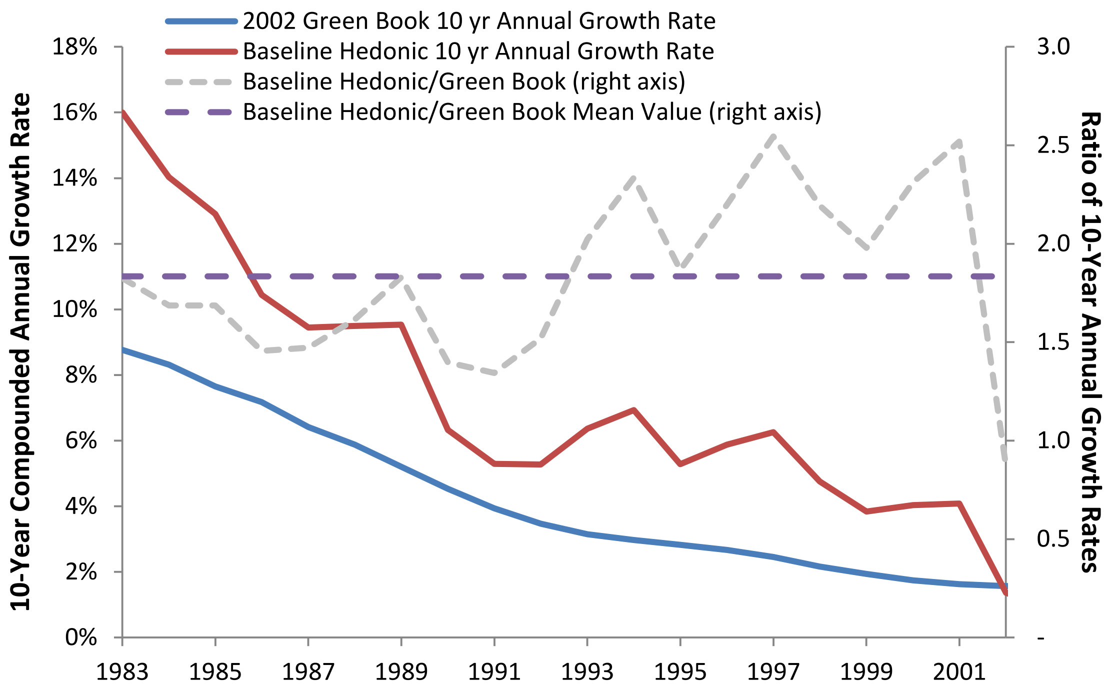


Figure 6. Comparison of Baseline Hedonic and Green Book Index Growth Rates, 1983–2002

Examination of the data shows that the hedonic and Green Book indexes relate to one another most consistently through a multiplicative factor vice an additive adjustment. We use the calculated average ratio (mean value) of 1.83 shown in the figure as a conversion factor on the 2003–2013 Green Book values to arrive at the projected hedonic index. This is shown along with the other indexes in Figure 7.

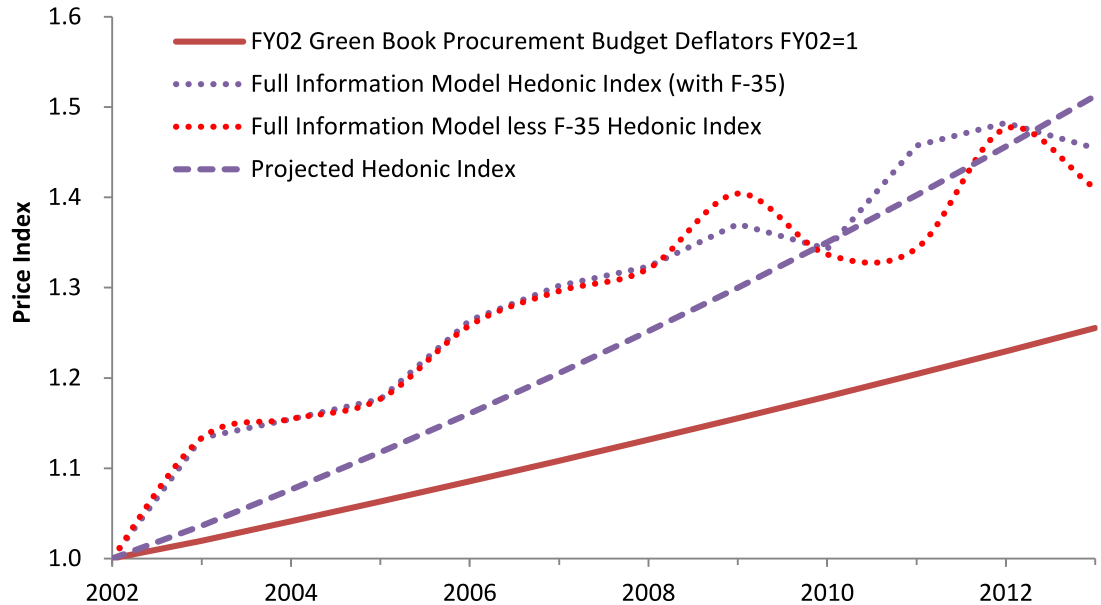


Figure 7. Comparison of Hedonic and Green Book Indexes, 2002–2013

Figure 8 compares the URF estimates associated with the two models and three escalation index assumptions with the budget actuals.

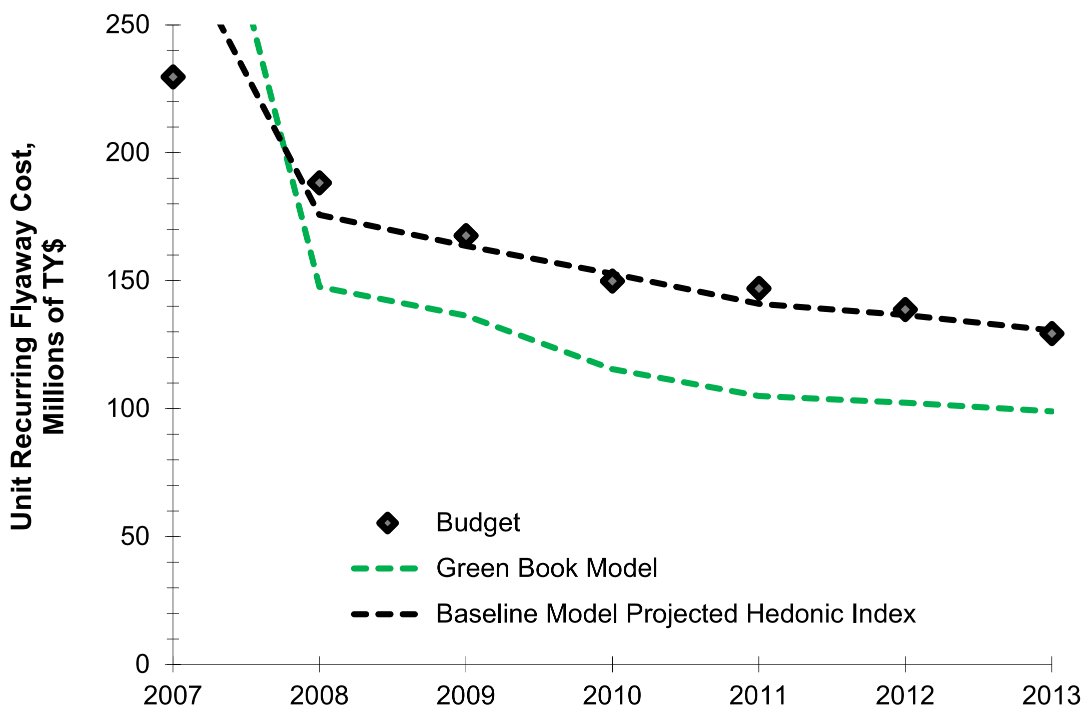


Figure 8. Comparison of Model Estimates with Budget Actuals, All F-35 Variants

Table 4 compares the estimated URF costs with the budget actuals calculated for the 2007–2013 budget years, broken out by variants.

Table 4. Comparison of Estimates of 2007–2013 URF Costs, Millions of TY$

Variants	Actual Budget	Baseline Model, Projected Hedonic Index	Green Book Model and Index
All Variants	149	147	115
F-35A	139	137	110
F-35B	160	152	121
F-35C	167	175	124

The results show that the Baseline model estimates when projected forward using the hedonic index come close to the actual budget values for 2007–2013; estimates depending on the Green Book index consistently underestimate the budget URF costs. However, the Baseline model tends to miss the costs for the individual variants, with the F-35B underestimated and the F-35C overestimated. This result is consistent with the differences in parameter estimates between the Baseline and Full Information models, which are, in turn, a result of the more complex STOVL implementation of the F-35B relative to the A/V-8B that is not completely captured in weight differences.

D. F-35 2013 SAR/PB 2015 Estimates

This section takes a somewhat different approach to the F-35 estimating problem. The question we want to answer is: what scaling of the FY 2015 Green Book index results in the closest fit to the latest JPO estimates? While the previous F-35 estimating exercises took the data available in 2002 as given, in this case we assume contemporary data for escalation projections. To address this question, we only use the Baseline model with the projected hedonic index as presented above. For 2014 onwards, we scale the FY 2015 Green Book index by a multiplier analogous to the factor used to calculate the projected hedonic index. The multiplier is determined by scaling the Green Book index such that the model-estimated totals for 2014–2037 are the same as those reported in the SAR. The resulting factor is 1.75—comparing directly with the 1.83 factor used to calculate the projected hedonic index. This analysis is shown graphically in Figure 9.

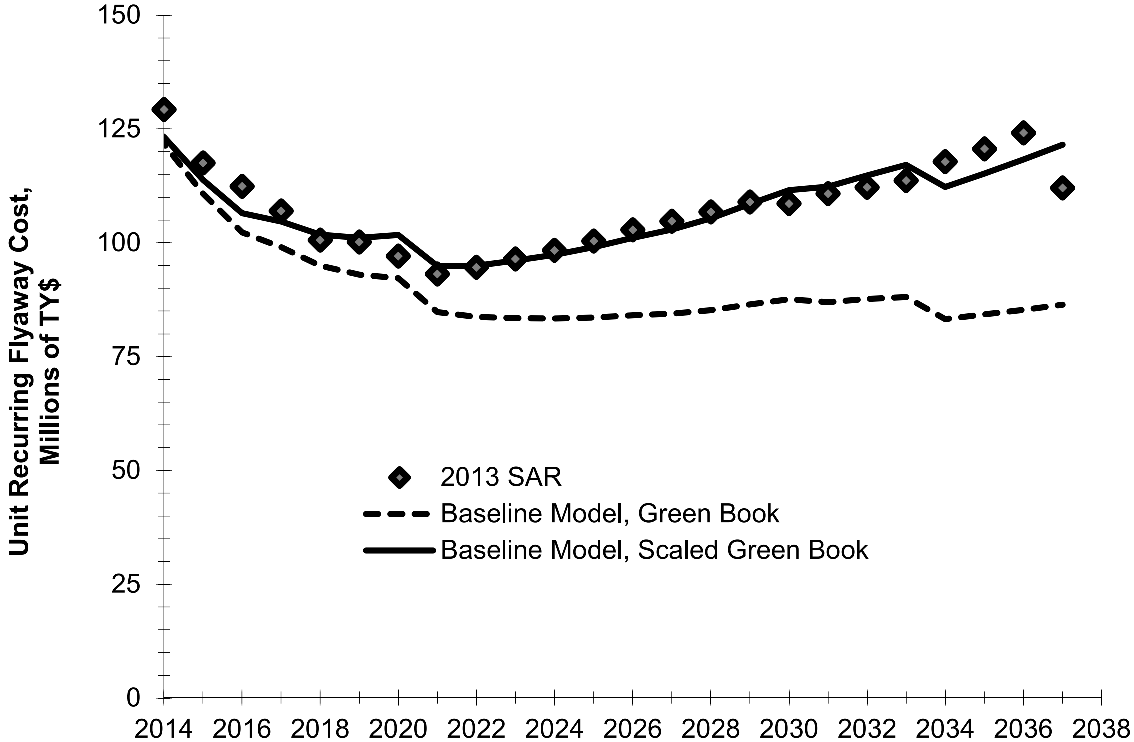


Figure 9. Comparison of Model Estimates with the 2013 SAR Estimates, All F-35 Variants

If the estimates are projected using the unadjusted Green Book index, the 2014–2037 URF estimate is $88 million vs. $106 million reported in the SAR. This shows the impact of the different indexes on projected costs, isolated from their influence on defining the CER model.

4. Summary and Conclusions

This paper describes different approaches to estimating expected price growth in defense system costs. The comparison of cost estimates based on escalation predictions derived from hedonic modeling with F-35 budget actuals through FY 2013 is particularly interesting. Although the model inputs reflect the latest F-35 aircraft characteristics and program parameters, in terms of the structure of the model and escalation projections, the models are defined by the information that was available at MS B. As the hedonic index is directly estimated only for the historic period, we apply a methodology to project forward escalation rates associated with the hedonic index. This example shows the close correlation between the Baseline hedonic model estimates and the budget actuals. The lower estimates from the Green Book model are due to two factors: the underestimates of escalation from FY 2002–FY 2013 and biases introduced into the model parameters because of underestimates of escalation in the historical period.

Looking out to FY 2037, we find that projecting escalation using our approach closely mimics the more detailed buildup of input-specific escalation rates used by the JPO. This is in contrast to projections using Green Book escalation, which result in an $18 million underestimate in unit costs.

We demonstrate the effect of different escalation methodologies using top-level CER models. Cost analysts usually build up their estimates from a more detailed level. However, issues regarding the proper application of price indexes, for both normalizing historical data and making projections, are equally valid in more typical cost estimating environments. For example, rates of price growth for raw material inputs, propulsion systems, electronic components, and labor inputs are likely to be different from that of general inflation. In our last example, we calculated overall escalation rates implied in the JPO estimates for the rest of the F-35 program; we found these escalation rates to be consistent with those projected using values from the historical hedonic price index.

Illustrations

Figures

Tables

References

Arnold, Scott A., Ji S. Byun, Harley A. Cloud, Alexander O. Gallo, Matthew W. Gonwa, Bruce R. Harmon, Prashant R. Patel, Colin D. Sullivan, John R. Hiller, and Patricia F. Bronson. "WSARA 2009: Joint Strike Fighter Root Cause Analysis." IDA Paper P-4612. Alexandria, VA: Institute for Defense Analyses, June 2010.

Harmon, Bruce R. (U) "Cost Estimating Techniques for Tactical Aircraft Manufacturing Labor." IDA Paper P-4490. Alexandria, VA: Institute for Defense Analyses, May 2010. (Unclassified/PI/LR).

Harmon, Bruce R., Daniel B. Levine, and Stanley A. Horowitz. "Inflation Adjustments for Defense Acquisition." IDA Document D-5112. Alexandria, VA: Institute for Defense Analyses, October 2014.

Harmon, Bruce R., J. Richard Nelson, and Scot A. Arnold. "Unit Cost Implications of New Materials: Preliminary Analyses of Airframe Experience (Revised)." IDA Document D-908-REV. Alexandria, VA: Institute for Defense Analyses, September 1991.

Nelson, Richard, Bruce R. Harmon, Robert J. Bontz, and Waynard C. Devers. (U) "Cost Estimating for Modern Combat Aircraft: Adjusting Existing Databases and Methods to Include Low-observable Cost Considerations." IDA Paper P-3528. Alexandria, VA: Institute for Defense Analyses, June 2001. (Secret/PI/LR).

Office of the Under Secretary of Defense (Comptroller). *National Defense Budget Estimates*, "The Green Book."

Resetar, S. A., J. C. Rogers, and R. W. Hess. "Advanced Airframe Structural Materials: A Primer and Cost Estimating Methodology." R-4016-AF. Santa Monica, CA: RAND Corporation, 1991.

Triplett, Jack E. *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. Paris, France: Organisation for Economic Co-operation and Development, 2006.

Younossi, Obaid, Michael Kennedy, and John C. Graser. "Military Airframe Costs: The Effects of Advanced Materials and Manufacturing Processes." MR-1370-AF. Santa Monica, CA: RAND Corporation, 2001.

Abbreviations

BY	Base Year
CER	Cost Estimating Relationship
DCMA	Defense Contract Management Agency
DoD	Department of Defense
EMD	Engineering and Manufacturing Development
FPRA	Forward Pricing Rate Agreement
FY	Fiscal Year
IDA	Institute for Defense Analyses
JPO	Joint Program Office
JSF	Joint Strike Fighter
MDS	Mission-Design-Series
MS	Milestone
OUSD(C)	Office of the Under Secretary of Defense (Comptroller)
PB	President's Budget
SAR	Selected Acquisition Report
STOVL	Short Takeoff and Vertical Landing
T1	First Unit
TY	Then Year
URF	Unit Recurring Flyaway
WSARA	Weapon Systems Acquisition Reform Act of 2009

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
xx-03-2016	Draft Final	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Role of Inflation and Price Escalation Adjustments in Properly Estimating Program Costs: F-35 Case Study	HQ0034-14-D-0001
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Harmon, Bruce R. Horowitz, Stanley A.	5e. TASK NUMBER BA-7-3054
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882	IDA Document D-5489 H 15-000415

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Office of the Secretary of Defense Cost Assessment and Program Evaluation Economics and Manpower Analysis Division Room BE798, The Pentagon Washington, DC 20301-1800	OSD/CAPE
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The application of price indexes presents a substantial challenge in estimating the costs of new defense systems. The problem is twofold. First, the analyst must use a price index when normalizing historical cost data to a common point in time (where the normalized costs are referred to as “base year” (BY) dollars in defense acquisitions or, more generally, “real” dollars) so that these data can be used to help estimate the costs of future systems. Second, as budget requirements for future acquisitions are in “then-year” (TY) dollars (or more generally, “nominal” dollars), BY dollar estimates must be inflated to TY dollars using a price index. Using an inappropriate price index can introduce errors in both of these steps. In this paper we apply two sets of price indexes to a cost estimating problem—the F-35 Joint Strike Fighter (JSF) procurement program. The purpose is to help cost analysts and others involved in the acquisition process understand the impacts of different price indexes and to provide guidance in their choice.
We approach this problem by way of hedonic price indexes derived from cost estimating relationships (CERs). Hedonic price indexes isolate changes in price due to factors other than changes in quality over time. We develop a “Baseline” CER model with this approach using data on historical tactical aircraft programs available at the F 35’s late-2001 Milestone B decision. Comparisons are made between the Baseline model estimates, estimates produced by the F-35 program office, and estimates using cost models employing more conventional approaches to inflation adjustment. We find that the Baseline hedonic model provides estimates close to actual F-35 costs. As the hedonic index is directly estimated only for the historic period, we develop a procedure to project inflation rates based on historical hedonic index values.

15. SUBJECT TERMS

Price Index, Inflation, Escalation, Cost Estimating Relationships, CER, Cost Growth, F-35, Joint Strike Fighter, Hedonic Price Index, Fighter Aircraft

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Summers, Garrett
Unclassified	Unclassified	Unclassified	Same as Report	34	19b. TELEPHONE NUMBER *(Include area code)* (703) 695-0722

Reset