Chapter 14
A Systems Estimation Approach to Cost, Schedule, and Quantity Outcomes

R. Bruce Williamson

14.1 Introduction

The typical United States (US) weapons system has grown in cost, capability, and the time required for development and production. There is also a strong chance that fewer units of that weapon system will be acquired than were initially contracted. Recent US fixed wing aircraft programs, such as the F-22 Raptor or the F-35 Joint Strike Fighter, are two examples of programs that have greatly reduced delivery quantities, much increased per unit costs, and prolonged development and production schedules. These programs are not unusual in the history of US weapons acquisition. Cost growth, schedule growth, and quantity change are interrelated factors in some manner in most defence acquisition programs, but the empirical work to substantiate the magnitude and direction of the interaction of all three factors working together has been scarce.

There is an evidence that the interrelationship creating the problem begins early in the life of a program, but often propagates as the program moves toward completion, or in rare instances, cancelation. The U.S. General Accounting Office (1992) noted that

> In weapons acquisitions, optimistic cost estimates are rewarded because they help win program approval, win contract awards, and attract budgetary income. The consequences of cost growth are not directly felt by an individual program because they are 'accommodated' through stretch-outs and quantity changes and by spreading the pain across many programs (GAO 1992; 40).

Inevitably, weapon systems will be affected by Congressional decisions resulting from Federal budget priorities, reduced strategic requirements (e.g., the end of the Cold War), or Congressional actions to perpetuate a defence program that the Department of Defence (DoD) may actually want to terminate, as well as by DoD decisions to spread budget in such a way make room for additional

R. Bruce Williamson (✉)
National Defense Business Institute, University of Tennessee, Knoxville, Tennessee, US
e-mail: rbwilliamson@utk.edu

J. Alleman et al. (eds.), *Demand for Communications Services - Insights and Perspectives*, The Economics of Information, Communication, and Entertainment,
DOI: 10.1007/978-1-4614-7993-2_14,

desirable programs.[1] As the final decision authority, "the temptations for Congress to compromise by trimming a little out of each program instead of cancelling whole programs is enormous" (Tyson et al. 1989). When this happens, program managers in the DoD recalculate how many units they can buy and what current program expenditures they can defer by shifting current development or production into the future, when the budget might be more plentiful. Aspects of the Congressional, DoD, and contractor relationship also tend to encourage short-term decision-making, so that the full consequences of the way in which cost, schedule, and quantity choices interact to create out-of-control major defence acquisition programs (MDAPs) may not be recognized for months or years, often after the program staff, contractor teams, and Congressional supporters have retired or moved to different civilian or military jobs.

Defence acquisition programs are certainly not unique in facing difficulties in effective management of large-scale programs. Government acquisitions in public works have also had similar, economically inefficient outcomes. Foreign governments have similar issues with major defence programs (Chin 2004). Because of the large number of concurrent, large-scale programs in one Federal Department, however, excessive cost growth or schedule growth for a specific system like the F-35 aircraft impacts not only that weapon system, but also, the 90–100 other MDAPs running concurrently. The combined effect across many programs aggravates existing budget instabilities and steadily widens the gap between planned program expenditure *ex ante* and required expenditure *ex post.* Learning from the interaction between cost growth, schedule growth, and quantity change in the management of MDAPs would benefit the DoD and Congress and enable a better look at opportunities forgone from other weapons systems not funded, delayed or assigned lower priority than prudent national security concerns would warrant. It would have application to similar public investment programs, such as infrastructure projects that require continuous funding approval.

The remainder of the chapter is organized as follows. The next section reviews previous studies. The third section presents the methodology, followed by a section describing the data sample used. This is followed by an analysis and discussion of results, and a concluding section with recommendations.

14.2 Previous Studies

Peck and Scherer (1962) and Scherer (1964) appear to have been the first to apply rigorous economic theory to understand the trade-off in measures of schedule variance and cost variance in weapons program management. They estimated a

[1] Defense Science Board (1978). The consequence is that for many large programs, when future year defence acquisition expenditures are charted against current year expenditures, the pattern begins to resemble the "bow wave" ahead of a ship.

simple correlation of schedule variance and cost variance in weapons programs of 0.57, significant and positive. Peck and Scherer (1962) noted "the greater the schedule slippage in a program, the larger the cost overrun tends to be." Scherer (1964) understood quantity decisions (i.e., how much to buy) to be "implicit" in cost estimates but stopped short of a systems approach to explain possible endogeneity issues.[2] Scherer (1964) also proposed a concurrent trade-off between time, schedule, and weapons quality, but this insight was not pursued empirically.

Defence acquisition research since Peck and Scherer (1962) and Scherer (1964) has focused separately on cost growth, schedule growth or quantity change. The analyses usually rely on descriptive measures and seldom employ multivariate methods to control for other correlated factors.[3] The result in the literature has been a large number of differing point estimates for different time intervals, data sets, samples, types of MDAPs, types of contract mechanisms, and so on. At least one of these studies has determined there is no empirical support for the relationship:

> …we could find no relationship between cost growth and schedule slip. Both of these variables are measures of program success, and two hypotheses about their relationship are commonly asserted: (1) that there are tradeoffs between cost and schedule growth such that a program can incur one and not the other, and (2) that they in fact occur together in problematic programs. Even though these two hypotheses are opposed to each other, we can find no support for either in our data (Drezner and Smith 1990; 45).

This is a surprising statement, not least because of the analysis of Peck and Scherer (1962) and Scherer (1964), but also because Tyson et al. (1989) had previously validated in an ordinary least squares (OLS) econometric framework that there was a connection:

> Development schedule growth, production stretch, and development schedule length are the major drivers of total program cost growth. Production stretch in particular has increased cost growth by 7–10 % points per unit increase in stretch (for example, by doubling the production schedule length while keeping quantity constant) (Tyson et al. 1989; ix).

Among many studies reviewed by this researcher, it appears that only one study (Tyson et al. 1994) implemented a systems approach to account for the potential simultaneity between cost and schedule growth in their model. The authors

---

[2] For Peck and Scherer's analysis, "Implicit in the notion of cost is the aspect of quantity. The cost per unit of a weapon system determines the quantity of systems which can be obtained with any specified amount of resources" (Peck and Scherer 1962; Note 1, 251). This is in line with the prevailing industry and military practice of normalizing cost for any quantity changes in comparisons (e.g., quantity-adjusted cost). This has the tendency of casting quantity change as a predictor of cost change, but not the reverse, in which quantity change is understood as the dependent variable. See Scherer (1964).

[3] Drezner et al. (1993) is one fairly common example of research practices in reviewing the key factors affecting cost growth, including budget trends, system performance, management complexity, and schedule-related issues. Their reliance on univariate and bivariate views of selected acquisition report (SARs) data unfortunately allows important explanatory factors to be confounded.

considered cost and schedule changes affecting tactical missile acquisition programs of the previous two decades.

The Tyson et al. (1994) analysis is the first example of an econometric systems approach (two-stage least squares) for the relationship between program cost growth and program schedule growth while controlling for observed quantity change in the regression.[4] The first-stage regression predicted program schedule growth, while the second stage used the predicted schedule growth as an explanatory variable in the regression on predicted cost growth. Although the sample sizes were minimal, the estimates improved compared with a single-equation model of cost growth as a function of schedule and quantity change. The specific results are not easily compared because the data set involved only a small number of a one program type, but the analytical insight to pursue a systems approach is an empirical landmark in the study of the interaction of cost growth, schedule growth, and quantity change in MDAPs.

14.3 Methodology

A complete systems econometric approach might consider as dependent variables three measurable program phenomena: cost growth from budgeted amounts, schedule slip, and quantity change. At least theoretically, three additional program phenomena are known anecdotally but lack objective measures. One of these additional phenomena is the alteration during a program's life in capability, reliability, quality, or performance, which are implicit in the term "performance" that will be used here. Two other important phenomena could also be mentioned, but lack objective measurement: first, the level of risk assumed with the development of new technology on a program; and second, the experience and leadership manifested by the program management team on the project. This is a total of six phenomena and explanatory factors on which a consistent set of metrics would be needed for a complete systems approach on how critical trade-offs are made during the life of a defence acquisition program.

In practice, program risk and program management experience have no metrics that a researcher might access. Further, the portion of SARs which would acknowledge some aspects of performance trades, such as those from program evaluations, are not publicly available. This leaves the present analysis with only three publicly available metrics of weapon acquisition program success, that is, cost growth, schedule growth, and quantity change.

The econometric technique used in this analysis compares single-equation OLS methods on cost growth, schedule change and quantity change independently, with that of a systems estimation approach that considers that some portion of the error

[4] See Tyson et al. (1994) "Examining the relationship between cost and schedule growth in development, we concluded that it was not appropriate to consider the major independent variable [Development Schedule Growth] to be exogenous…Therefore we estimated the tactical missile development relationship as a simultaneous system of equations."

in estimation is common to all three equations. The specific methods are described in Kmenta (1986) and Zellner (1962). Zellner (1962) considered the possibility of a relationship between the dependent variables which is not already incorporated in the set of separate regression equations that explain the dependent variables of interest. In fact, there could be mutually relevant information being omitted from each equation in the set. This would be reflected in correlation between the regression disturbance terms, and as a consequence, the separate equation OLS coefficient estimates would no longer be the most efficient or unbiased. The systems technique used here follows Zellner's seemingly unrelated regression method (SUR) approach (Kmenta 1986).

The increase in estimation efficiency from the additional explanatory information in SUR should reduce the standard error of the coefficient estimates, if there is cross-equation correlation of the disturbance terms. If there is no correlation between the error terms, then the OLS single-equation estimator is as good as the SUR estimator (Greene 2000). SUR estimation efficiency improves the greater the correlation is between the equation disturbance terms.

The value in improved estimates of the effects of program factors on each other is threefold. First, it sheds light on the size and direction of quantitative interactions of program phenomena that result from trade-offs made to satisfy budget, schedule, and quantity requirements. Second, it improves qualitative insight that these kinds of trade-offs are a regular occurrence in the course of a MDAP. Third, it allows for the testing of policy change and acquisition reform variables in a systems multivariate setting to answer the long debated question whether any defence acquisition reform has produced a beneficial result in reducing cost growth, schedule growth, or quantity shrinkage.

14.4 Data

This study uses published SARs on MDAPs in a data set from RAND Corporation from March 1994.[5] RAND authors were consistent in methods used to adjust estimated baseline and future costs by quantity changes, in restating all program cost data in constant 1996 dollars, and in documenting many data issues in both (Jarvaise et al. 1996) and the SARs methodology monograph by Hough (1992). Missing values are a significant shortcoming of the RAND data set, and one not easily rectified.[6] Data from McCrillis (2003) and McNicol (2006) were used to

[5] RAND released their version of SARs data (Jarvaise et al. 1996); the electronic version is available on the RAND website (www.rand.org) as an electronic data appendix to the Jarvaise et al. (1996) monograph.

[6] For example, schedule data—perhaps a single missing planned date for MS B—impacts several variables calculated using that data as an end point for one interval, the start point for the next interval, and various ratios of program phase length and schedule slip ratios that could be calculated from that variable. MDAPs that began reporting by the 1969 SAR had only

supplement the RAND data to permit testing of basic program risk characteristics as well as controlling for the potential effects of DoD and Congressional acquisition reforms during the time period considered.

The SARs rely on event-specific baseline estimates that are made at the program's first, second, and third milestones.[7] These formal baseline estimates occur at the Concept Development stage, or program Milestone A, called a Planning Estimate; a second baseline estimate is made at Milestone B at the start of the engineering and manufacturing development (EMD) phase, called a Design Estimate; and at Milestone C a third baseline estimate, called a Production Estimate, is reported at the start of actual production. As a program proceeds, there is a Current Estimate supplied in the SAR either in the (December)year-end SAR summaries or the quarterly SAR updates.

The variables of interest in this analysis are derived from DoD program estimates posted at the baseline points mentioned above. The dependent variables in the analysis will measure *relative change* in key program factors, instead of absolute dates, final quantity, or unit cost; the independent variables applied must also have plausible explanatory relationships with the relative change in the dependent variables.

Quantity change (variable name QTYCHG) is defined as the number of units to be purchased as measured at program's latest available Current Estimate or last program baseline estimate, relative to the quantity that was expected to be purchased at the initial Design Estimate baseline. Here, quantity changes are expressed as a ratio, the numerator being the latest quantity estimate, and the denominator being the quantity baseline estimate for the program phase(s) of interest in the analysis. The ratio of latest Current Estimate quantity to the Design Estimate quantity at Milestone B has approximately 29.5 % of programs exceeding a ratio of 1.0; 15.3 % of programs are close to a ratio of 1.0; and about 55.2 % of MDAPs have ratio less than 1.0.

MDAPs with quantity change ratios above 1.0 include programs like the F-16 aircraft, which was set for 650 aircraft at Milestone B, but eventually saw final production of 2,201 aircraft with foreign sales. Programs with a quantity change ratio in the neighborhood of 1.0 produce what was originally planned. Programs with ratios less than 1.0 include those programs which had cuts in later quantities relative to the initial order. This is the largest category of MDAPs. Most MDAPs,

(Footnote 6 continued)
fragmentary or inconsistent data covering their early years before the SARs began. See Hough (1992) and Jarvaise et al. (1996). For the present analysis, an MDAP sample of 244 programs is reduced by 50 % or more in multivariate estimation procedures as a result of the missing endpoints on schedule data.

[7] The discussion is simplified to focus on the essential concepts required rather than burden the reader with DoD program terminology and acronyms. The reader is referred for additional technical detail of acquisition cycles, program structure and weapon program reporting metrics to the studies by Hough (1992) and Jarvaise et al. (1996).

for a variety of reasons, never produce the quantities they were expected to produce at the time of contracting.

Schedule change (variable SCHEDCHG) measures the relationship between the program's actual schedule to complete relative to the planned schedule from the first available baseline estimate. Schedule is measured in months, and the variable is constructed as a percentage difference. Schedule change is often referred to as "schedule slip" as well. When a program is "stretched out" by Congressional or DoD budget decisions, it means that schedule slip has been built into a revision in the program's planning (U.S. Congressional Budget Office 1987).

In the sample used here, only 8.3 % of programs had negative schedule growth, that is, they were completed in less time than expected. A much larger proportion of programs (31.5 %) had no schedule growth at all, and 60.2 % of MDAPs had positive schedule growth, taking months or years longer to complete than expected.

Cost growth (variable COSTCHG) is defined as the quantity-adjusted percentage ratio between a cost estimate at the last available reporting or baseline, versus the cost estimate at the initial SAR baseline for the program phase(s) of interest. Inflation is removed as a confounding factor in order to isolate program real cost growth, and cost growth is normalized for quantity change that may have occurred during the program. This procedure permits reasonable comparison of per unit costs across programs, despite there being numerous changes in planned quantities to be purchased and the budget required to purchase those changed quantities.[8,9] The variable used here is calculated from the latest available program cost reported in a Current Estimate, relative to the initial Design Estimate baseline cost.

The Cost Growth variable in this analysis can be negative, zero, or a positive percentage change in real cost between the last Current Estimate of program cost and the Milestone B baseline cost estimate, again, adjusted for quantity change. The sample used here has 42.1 % of programs with negative real cost growth; 5.1 % of programs have real cost growth of zero; and 52.8 % of defence programs have positive real cost growth, all after adjustments for quantity change.

The modeling variables used in this analysis are given with their summary descriptive information in Table 14.1. The characteristics for both the full data sample as well as the final estimation sample data set are shown. The primary difference in the two sample sizes results from the unfortunately high proportion (56 %) of missing values for the schedule slip variable (SCHEDCHG). Combined

[8] Because there have been many different purposes in calculating cost growth, it is possible to find cost analyses that focus on a specific time period, i.e., Development cost growth, or Production cost growth. This analysis calculates cost growth from the Design Estimate to the last available Current Estimate, or to project completion, whichever happens first, and is a more comprehensive measure of overall program cost growth.

[9] All SAR cost estimates throughout the life of a defence contract must be adjusted (or normalized) for effective cost (total and per unit) given the changing number of units to be purchased. How to best achieve this has been a topic of lengthy debate and refinements among defence analysts. The best exposition is in Hough (1992, 30–41), as well as in the earlier report by Dews et al. (1979, Appendix A).

Table 14.1 Descriptive summary for variables of interest in major defence acquisition program analysis

Variable name Variable type Variable use in analysis	QTYCHG Ratio Dependent	SCHEDCHG Percentage Dependent	COSTCHG Percentage Dependent	PEO Binary Independent	MUNITION Binary Independent	FXDWING Binary Independent	NAVY Binary Independent	B1950SYR Time indicator Independent
Full sample N	172	108	159	232	244	244	244	228
Count missing	72	136	85	12	0	0	0	16
Percent missing	0.30	0.56	0.35	0.05	0.00	0.00	0.00	0.07
Mean	1.03	18.13	29.47	0.30	0.06	0.13	0.41	29.02
Median	0.89	7.71	4.22	0.00	0.00	0.00	0.00	29.00
Standard deviation	1.01	28.73	115.11	0.46	0.24	0.33	0.49	8.56
Minimum	0.00	−11.92	−77.81	0.00	0.00	0.00	0.00	10.00
Maximum	5.69	149.71	1071.21	1.00	1.00	1.00	1.00	50.00
Skewness	1.97	2.52	6.76	0.87	3.67	2.25	0.39	0.02
Kurtosis	4.76	7.37	54.45	−1.26	11.59	3.10	−1.88	−0.69
Estimation sample								
Estimation Sample N	67	67	67	67	67	67	67	67
Mean	1.28	21.03	20.57	0.28	0.07	0.15	0.38	28.59
Median	0.99	14.59	5.62	0.00	0.00	0.00	0.00	28.50
Standard deviation	1.13	28.22	46.10	0.45	0.26	0.36	0.49	8.09
Minimum	0.00	−5.27	−52.03	0.00	0.00	0.00	0.00	13.00
Maximum	5.69	149.71	195.39	1.00	1.00	1.00	1.00	45.00
Skewness	1.94	2.38	1.63	1.01	3.34	2.04	0.49	0.14
Kurtosis	4.04	7.28	3.43	−1.02	9.45	2.22	−1.81	−0.91

Table 14.2 Correlation matrix for dependent variables

	COSTCHG	SCHEDCHG	QTYCHG
COSTCHG	1.000		
SCHEDCHG	0.279	1.000	
QTYCHG	0.321	−0.00495	1.000

with the effect of missing values in other variables, the net result is that there are only 67 MDAP programs available for econometric analysis over the 1965–1994 period.

The full sample and the estimation sample are not appreciably different in terms of the distributions of the key variables in the two samples. Distributions for the quantity, schedule or cost growth variables are moderately right skewed and thinner tailed than normal distributions,[10] although the estimation sample is less affected than the full sample by these characteristics.

Definitions for the regression exogenous variables are the following:

B1950SYR: Calendar year of program start minus 1950.
FXDWING: 1 = fixed wing aircraft procurement program, 0 = otherwise.
MUNITION: 1 = Munitions program, 0 = otherwise.
NAVY: 1 = Navy program; 0 = otherwise.
PEO: 1 = program started at the time of, or following, the creation of the DoD program executive offices (PEO) in 1987; 0 = otherwise.

Table 14.2 displays the simple correlation between the three dependent variables for the estimation sample. There is no apparent correlation between schedule change and quantity change, although the modest correlation between cost growth and both of the other dependent variables will be investigated in the analysis that follows. As noted earlier, Peck and Scherer (1962) calculated a simple correlation of 0.57 between cost growth and schedule growth with a different sample of major programs from the 1950s, while the correlation in the present sample is only 0.28. The degree of correlation is a suggestion of the potential gain from the use of systems methods in place of single-equation methods, as discussed previously.

The estimation and results using this (Table 14.1) data set are now examined.

14.5 Analysis and Results

Table 14.3 presents results of three single-equation OLS models in the center column, with the SUR system estimates provided in the rightmost column. The dependent variables are listed on the left side of the table. Table 14.4 on the following page provides goodness of fit measures that accompany the equation estimation results in Table 14.3.

[10] A measure of kurtosis = 3 and skewness = 0 would describe a normally distributed variable.

Table 14.3 Results of the OLS single-equation and SUR system equation estimates[a]

Dependent variable	Independent variables	Single-equation coefficients (standard error)	SUR system coefficients (standard error)
COSTCHG	CONSTANT	−24.98	−48.10
(Cost model)		(29.12)	(17.67)
	FXDWING	21.54	34.30
		(18.34)	(13.60)
	PEO	−25.31	−35.26
		(24.13)	(19.16)
	B1950SYR	1.75	2.59
		(1.26)	(0.82)
SCHEDCHG	CONSTANT	34.41	34.37
(Schedule model)		(6.66)	(6.23)
	FXDWING	−26.43	−27.55
		(5.02)	(4.86)
	PEO	−18.91	−18.96
		(5.99)	(5.78)
	NAVY	−10.10	−9.53
		(6.20)	(5.51)
QTYCHG	CONSTANT	1.59	1.61
(Quantity model)		(0.19)	(0.18)
	PEO	−0.75	−0.75
		(0.20)	(0.20)
	MUNITION	−1.13	−1.39
		(0.22)	(0.22)

[a] Robust-White heteroscedasticity correction used.

The regressions were estimated on several variables of interest to understand sign and significance, prior to the comparison with a systems approach using SUR estimation. The variables have minimal collinearity, are not serially correlated, and are adjusted with robust estimation to create heteroscedastic-consistent standard errors.[11]

Turning now to the SUR estimates in the right-hand columns of Table 14.3, the standard error of the SUR coefficient estimates is reduced (and thus t-statistics improved) from the standard error of the OLS coefficient estimates.

Overall measures of goodness of fit are presented in Table 14.4. The models have low R^2 overall suggests the models could be improved and sample issues resolved, although R^2 has limited value in appraising SUR or systems models.

[11] Only two variables (B1950SYR and the PEO dummy variable) show even mild collinearity, but the condition index is 10.86. All other variables produce condition indexes under a value of 2.97. The Belsley-Kuh-Welsch guidelines are that a condition index value greater than 15 indicates a possible problem and an index greater than 30, a serious problem with multicollinearity (Belsley et al. 1980). Serial correlation in a traditional sense is not present, although it is possible that there may be a cohort effect for groups of programs going through the same program stage in the same years, but with limited sample available, the sample partitioning would leave few degrees of freedom for hypothesis testing.

Table 14.4 Model comparative diagnostics

Dependent variable		Single-equation models	SUR system model
Sample N		67	67
Time period		1965–1995	1965–1995
Equation 1	*R*-squared	0.054	0.05
Cost model	S.E. of regression	45.98	44.93
	Mean dependent variable	21.16	21.16
	S.D. dependent variable	46.19	46.19
	Variance of residuals	2114	2019
Equation 2	*R*-squared	0.19	0.19
Schedule model	S.E. of regression	26.12	25.33
	Mean dependent variable	21.18	21.18
	S.D. dependent variable	28.41	28.41
	Variance of residuals	682.1	641.5
Equation 3	*R*-squared	0.15	0.15
Quantity Model	S.E. of regression	1.06	1.04
	Mean dependent variable	1.29	1.29
	S.D. dependent variable	1.13	1.13
	Variance of residuals	1.12	1.08

The coefficients from the SUR estimation are substantially larger for the SUR cost growth model, and less so for the schedule growth and quantity change models, relative to single-equation coefficient estimates. Reduced variance of the estimation residuals and the greatly improved explanatory power of the variables used for comparison suggest that there are efficiency gains in using the SUR approach versus three independent OLS equations.

The next step is to use the improved coefficient estimates of the SUR model to evaluate the sensitivity of the model to changes in key measures. In order to calculate the sensitivity of the model and its effect on the dependent variables, one needs to use the estimated coefficients and substitute in the values from the sample means on factors one wishes to hold constant, while experimenting with setting binary dummy variables at either their zero or one settings. The difference in the predicted values of the dependent variable is the predicted impact on the model when whatever condition the dummy variable represents is present versus when it is not present.

For example, to determine the impact of fixed wing aircraft on cost growth from the SUR cost growth equation, the following is calculated:

Step 1: Predicted average cost growth *with* fixed wing aircraft programs present $= -48.1 + (34.3 * 1) + (-35.36 * 0.28) + (2.59 * 28.59) =$ 50.5% average increase in cost growth *with* fixed wing aircraft programs

Step 2: Predicted average cost growth *without* fixed wing aircraft programs present $= -48.1 + (34.3 * 0) + (-35.36 * 0.28) + (2.59 * 28.59) =$ 16.2 % average increase in cost growth *without* fixed wing aircraft programs

Step 3: Net predicted impact on average cost growth from having fixed wing aircraft programs = (Step 1 result) − (Step 2 result) = 50.516.2 % = 34.3 % net increase in cost growth due to fixed wing aircraft programs

The same process is used to calculate the net impact of the other independent variables in Table 14.5 by substituting the appropriate values for the coefficient estimates and the sample mean value for the independent variables that are being held constant in the comparison.

Fixed wing aircraft programs (FXDWING), on average, produce a net impact of 34.30 % on cost growth. They are also associated on average with a 27.55 % reduction in schedule slip. The two outcomes may have much to do with program priorities accepting more cost growth in order to keep vital aircraft programs on schedule. Munitions programs (MUNITION) are significant only in the quantity change equation in the SUR model, where the net effect of a munitions program on quantity change is −1.39. One interpretation is that munitions programs tended to face significant quantity reductions when programs have exceeded budgets in this time period.

The Navy service variable (NAVY) has a predicted net effect on schedule slip of −9.53 %, which may be the result of program decisions where some cost growth was preferable to schedule slip during this period.

The variable (PEO) for the creation of PEO in 1987 plays a significant role in each equation.[12] It is included as a policy variable of interest worth any reduction in estimation efficiency.[13] In two of three equations in the SUR model, the PEO variable significantly reduces predicted cost and schedule growth. The predicted net effect from the creation of the PEO in 1987 was a 35.3 % net reduction in average cost growth from the situation before 1987; for the schedule model, the creation of the PEO correlates with a net 18.96 % reduction in schedule slip compared with the pre-1987 period. In the quantity equation of the SUR model, the creation of the PEO produces a net increase of 0.17 in the quantity change ratio compared with the situation before 1987. The simplest explanation overall is that the creation of the PEO significantly increased fidelity in program execution to the originally contracted cost, schedule, and quantity. In short, there is clear evidence that the acquisition reform that created the PEO in 1987 actually worked as intended.

---

[12] Other variables defined by McNicol (2006) and McCrillis (2003) also related to time periods in which DoD cost estimation requirements were tightened, or in which DoD budgets were plentiful, but coefficient estimates for these additional variables were statistically insignificant in estimation. In contrast, in the cost growth equation the highly significant PEO variable coefficient has a p value of 0.066, in contrast to p-values of 0.001 and 0.000 in the schedule change and quantity change equations, respectively.

[13] Including the variable PEO in each equation is a compromise because the right hand side variables across the set of equations in the SUR model are more similar, which reduces the gain from using SUR methods. Ideally, each equation in a SUR system should have dissimilar explanatory variables, but correlated dependent variables (Greene 2000).

Table 14.5 Predicted net change in SUR equation dependent variables

Dependent variable	Binary (dummy) independent variables	Predicted change in dependent variable with explanatory variable of interest set to 1.0	Predicted change in dependent variable with explanatory variable of interest set to 0.0	Predicted *net* change in dependent variable
Cost growth (percent, constant $)	Fixed wing aircraft programs	50.52 %	16.22 %	34.30 %
	Program executive offices	– 4.16 %	31.09 %	−35.26 %
Schedule growth (percent)	Fixed wing aircraft programs	– 2.11 %	25.44 %	−27.55 %
	Program executive offices	7.72 %	26.68 %	−18.96 %
	Navy programs	15.50 %	25.03 %	−9.53 %
Quantity growth (ratio)	Program executive offices	0.75	0.58	0.17
	Munitions	0.01	1.40	−1.39

Note All other variables on the right-hand side of the regression equations are held at their sample average value (for the estimation sample) reported in Table 14.1

The variable not included in Table 14.5 is the year of program start (B1950SYR). Because of the way the variable is defined, one can estimate the net effect of the variable on predicted cost growth by calculating the effect on cost growth in say, 1970, and then repeat the same process for 1980. The difference in the predicted cost growth estimated at two points in time is the net cost growth during that interval. Using this approach, there is a net real cost growth of 25.94 % between 1970 and 1980. This result is particularly surprising because it is real cost growth not accounted for elsewhere. A secular trend in program cost growth of averaging 2.6 % a year may reflect quality differences, performance, and technology, which is not otherwise controlled for by the cost or program measures in the SUR model. One interpretation is that the better technology in weapons programs appears to cost more over time, in real terms, all else held constant. The result would have to be compared with other technology sectors, but even with decreases in component costs for similar performance, the implication is that the integrated technology in weaponry increases, not decreases in real cost over time.

14.6 Conclusions

A systems approach using SUR finds efficiency improvements to be had over the standard OLS approach used in previous research. In addition to the unsurprising findings on fixed wing aircraft programs, munitions programs, and Navy programs, the SUR systems model provides good support for the finding that the PEO have actually achieved what they were intended to in the time period covered by the model, that is, they reduced cost growth, reduced schedule slip, and maintained acquisition quantities, compared with the time period before the PEO came into existence in 1987. The matter of acquisition reform effectiveness is frequently debated, and empirical findings showing a particular reform has worked is infrequent in defence economics. A second unexpected result was identifying a real cost growth trend of about 2.6 % per year in the data, with the possibility that this might reflect implicit quality change over time in weapons systems.

Although the systems approach to SARs data improved the parameter estimates on all tested variables, it also suggests ways to improve the analysis. A full systems model with endogenous variables included on the right-hand side was not tested, in part because specification error becomes even more important. Future work will investigate full information systems models for potential improvements in fit and predictive power, but a key aspect in that will be resolving issues with data. In addition, this paper did not investigate cross-equation restrictions, but the approach might prove useful in other systems model specifications that test service-specific or weapon system-specific samples.

Returning to the theoretical framework stated earlier, more work needs to be done on the key metrics of MDAP outcomes which are not captured, or captured poorly in the SARs data. There are other data sources internal to the DoD that need to be considered, although there may be some compatibility issues with SARs data due to definitional differences and intended use. An extended theoretical framework, and the data on MDAPs to test it, would improve one's understanding of the interaction of cost, schedule, quantity, performance, risk, and management experience in a way that could genuinely improve outcomes of defence acquisition programs and help policy makers identify defence acquisition reforms that actually work.

References

Belsley D, Kuh E, Welsch R (1980) Regression diagnostics: identifying influential data and sources of collinearity. John Wiley, New York, NY

Chin W (2004) British weapons acquisition policy and the futility of reform. Ashgate, Burlington, VT

Defense Science Board (1978) Report of the acquisition cycle task force. Office of the Under Secretary of Defense for Research and Engineering, Washington D.C

Dews E, Smith G, Barbour A, Harris E, Hesse M (1979) Acquisition policy effectiveness: department of expense experience in the 1970s. RAND Corporation, Santa Monica, CA, R-2516

Drezner JA, Smith G (1990) An analysis of weapon system acquisition schedules. RAND Corporation, Santa Monica, CA, R-3937

Drezner JA, Jarvaise JM, Hess R, Hough WPG, Norton D (1993) An analysis of weapons system cost growth. RAND Corporation, Santa Monica, CA, MR-291

Greene WH (2000) Econometric analysis, 4th edn. Prentice Hall, Englewood Cliffs

Hough PG (1992) Pitfalls in calculating cost growth from selected acquisition reports. RAND Corporation, Santa Monica, CA, N-3136

Jarvaise JM, Drezner JA, Norton D (1996) The defense system cost performance database. RAND Corporation, Santa Monica, CA, MR-625

Kmenta J (1986) Elements of econometrics, 2nd edn. Macmillan, New York

McCrillis J (2003) Cost growth of major defense programs. Annual Department of Defense Cost Analysis Symposium (ADoDCAS). Williamsburg, VA, 30 Jan

McNicol DL (2006) Cost growth in major weapon procurement programs. Institute for Defense Analysis, Alexandria, VA, IDA P-3832

Peck M, Scherer FM (1962) The weapons acquisition process: an economic analysis. Harvard, Boston

Scherer FM (1964) The weapons acquisition process: economic incentives. Harvard, Boston

Tyson K, Nelson JR, Om N, Palmer PR (1989) Acquiring major systems cost and schedule trends. Institute for Defense Analysis, Alexandria, VA, IDA P-2201

Tyson K, Harmon B, Utech D (1994) Understanding cost and schedule growth in acquisition programs. Institute for Defense Analysis, Alexandria, VA, IDA P-2967

U.S. Congressional Budget Office (1987) Effects of weapons procurements stretch-outs on costs and schedules. Washington, D.C

U.S. General Accounting Office (1992) Defense weapons systems acquisition. Washington, DC. HR 93-7

Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J A Stat Assoc 57(298):348–368