COST REALISM: ASSURING MORE REALISTIC CONTRACTOR COST PROPOSALS

Donald L. Trapp, Science Applications, Inc.

ABSTRACT

Unrealism in Defense contractors' cost proposals, especially for RDT&E programs, often contributes to cost growth as well as to other problems. The Defense Department is therefore concerned with achieving greater cost realism. A methodology has been developed for achieving greater realism of contractor cost proposals. This methodology defines cost realism as an evaluation criterion stated in the solicitation which compares the offeror's proposed cost with a detailed Government estimate for each contractor and then scores the degree of realism. The methodology constitutes a source selection cost evaluation process involving (1) determination of cost evaluation factors; (2) preparation of instructions to be included in the solicitation concerning the cost evaluation factors; (3) preparation of Government estimates for each offeror; and (4) scoring each offeror for cost realism and Government estimated cost. The methodology is a synthesis and improvement of the best techniques and procedures currently being used in source selection cost evaluation (especially those of NAVELEX).

INTRODUCTION

Unrealism in Defense contractors' cost proposals, especially for RDT&E programs, often contributes to cost growth as well as to other problems. Cost growth may require drastic changes in system design and program scope. It also may lead to schedule slippages and program cancellations. The Defense Department is therefore concerned with achieving greater cost realism. Science Applications, Inc. (SAI) has undertaken to study the subject of contractor cost realism for the Navy Office for Acquisition Research.

How realistic are contractor cost proposals? This is not easy to determine. The complexity and risks inherent in Defense systems make contractor cost estimation, not to mention Government review and evaluation of the costs, difficult. Some optimism may be warranted (e.g., in engineering hours required, in number of test articles, and in number of tests), but when does optimism become unrealism? In some cases, the contractor may purposely propose an unrealistically low cost in order to "buy-in" and thereby improve its business base and cash flow.

SAI has collected data on proposed costs and Government estimates for eight programs from the Naval Electronics Systems Command (NAVELEX) which provide some insight into the degree of realism found in contractor cost proposals. In developing its estimates, NAVELEX evaluated the contractor cost proposals in detail, required the contractors to justify their data and estimating rationale, developed their own estimates based almost entirely on the contractor's data, and compared these estimates with the cost proposals. The data showed that one out of three original offers was more than 20% low, and one out of six best-and-final offers was more than 20% low. Offers more than 20% below the NAVELEX estimate certainly raise a question as to cost realism. Interestingly enough, less than one-quarter of the original offers but nearly half of the best-and-final offers were within 3% of the NAVELEX estimates, apparently as a result of the detailed Government cost evaluation and the fact that cost realism was an evaluation criterion.

Of course, the Government must also guard against unrealistically high cost estimates. This situation will most probably occur if there are few technically qualified contractors and if a contractor already has an especially good business base. High cost estimates do minimize the likelihood of an overrun, but the Government could probably have bought the system for less.

How do we achieve greater cost realism? This poses another difficult question, especially since there are often disincentives to realistic cost proposals in source selections for cost-reimburseable contracts. These disincentives occur because cost is generally a secondary consideration, especially for RDT&E contracts. Additionally, the Government often tends to award to the contractor with the lowest cost proposal, if technical and other factors are not too different; or at least, the contractors may perceive that the Government has such a bias.

The Defense Department needs to shape, refine, and implement acquisition policies and practices to encourage and achieve more cost realism. In particular the source selection procedures should be improved so as to:

- Encourage realistic bids,
- Discourage awards to low bidder on RDT&E contracts, unless realistically low, and
- Discourage success orientation bias.

It is primarily through the source selection cost evaluation process that the Government can direct efforts to give greater assurance that cost proposals are reasonable, defensible and realistic. In a general sense, then, cost realism may be defined as that part of the source selection process directed at giving

greater assurance of the reasonableness and defensibility of contractors' cost proposals.

SAI has developed a recommended methodology (i.e., techniques and prodedures) for achieving greater realism of contractor cost proposals which can be used by the Navy, as well as other Services. This methodology constitutes a source selection cost evaluation process involving four interrelated steps:

(1) Evaluation: Determination of cost evaluation factors;
(2) Solicitation: Preparation of instructions to be included in the solicitation concerning the cost evaluation factors;
(3) Government Estimate: Preparation of Government estimates for each offeror; and
(4) Scoring: Scoring each offeror for cost realism and Government estimated cost.

This recommended methodology is a synthesis and improvement of the best techniques and procedures currently being used in source selection cost evaluation (especially those of NAVELEX). This methodology suggests a more specific definition of cost realism as an evaluation criterion stated in the solicitation which compares the offeror's proposed cost with a detailed Government estimate for each contractor and then scores the degree of realism. This methodology is discussed in the following sections.

EVALUATION FACTORS

Generally, the two most important source selection factors or criteria are technical and cost. In RDT&E competitions, technical is far more important than cost. Additional evaluation criteria may include management, supportability, and schedule. Each evaluation criterion may have several sub-criteria.

In evaluating cost, the most important factors are the likely cost to the Government and the realism or reasonableness of the contractors' bids. The former we will term the Government estimated cost (or estimated cost) sub-criterion and the latter the cost realism sub-criterion. Both sub-criteria are important because, from the cost perspective, the award of a contract should be made not just on the basis of a low cost, but a realistically low cost. Other cost sub-criteria might include Design-to-Unit-Production-Cost (DTUPC) and/or Life Cycle Cost (LCC).

In a sense, all cost evaluations will address the realism or reasonableness of contractors' bids. For example, cost evaluators will normally check for completeness, accuracy and reasonableness of the bids. However, if the contractors are to be encouraged to make more realistic bids, it is very important to make cost realism an explicit evaluation sub-criterion. Cost realism can be assessed on the basis of the difference between the proposed cost (the bid or offer) and a Government estimate of the realistic cost for that contractor. The smaller this difference, the more realistic the cost. Therefore, a realistic Government estimate for each contractor is of critical importance, as discussed later.

Estimated cost can be assessed on the basis of the Government estimate of the realistic cost for each contractor. The lower the Government estimate, the better. This evaluation sub-criterion also requires that a realistic Government estimate be developed.

Although DTUPC and LCC may be important in RDT&E competitions, they should generally be given relatively little weight compared to cost realism and estimated cost. This will assure that cost realism and estimated cost have sufficient weight in the overall source selection evaluation.

Prior to issuing the solicitation, weights for each of the criteria and sub-criteria should be determined. For RDT&E programs, cost will usually be weighted much lower than technical. For example, cost might be 20 to 30% and technical 50 to 70% of the total score, with other criteria making up the difference. Deciding on the weighting of cost and the breakout between cost realism and estimated cost should be based on experience and an analysis of the effects of different weights and scoring rules on the overall scores for hypothetical proposals. Such an analysis should consider the trade-off between cost and technical scores (e.g., the degree of cost realism that could be traded for a certain degree of technical superiority).

Making cost realism an explicit evaluation criterion has great merit. It places contractors on notice that cost realism will be examined and could be a deciding factor in the award. The Government source selection team is also on notice to address cost realism specifically. The actual evaluation of cost realism can be done at many levels of detail depending on the procurement and the personnel available. Regardless of the level of detail, the key point is to make cost realism an evaluation factor which the contractors know can influence the award.

Furthermore, a detailed cost evaluation, such as that required for assessing cost realism can help in negotiation and is likely to result in best-and-final offers which are more realistic than the original offer.

SOLICITATION

The solicitation, in the form of an RFP or

RFQ, is the formal document specifying the Government's requirements associated with an impending procurement, as well as such matters as the terms and conditions and evaluation criteria. It is important to inform the prospective offerors in the solicitation that, among other factors, they will be evaluated on the basis of cost realism and that an unrealistic offer may be grounds for rejecting the proposal. Knowing that they may be penalized for unrealistic offers will certainly encourage the offerors to be more realistic. Whether they will in fact be more realistic depends on how the offerors weigh the penalty for being unrealistic against other factors, such as a belief that the award will be biased toward the low offer despite any unrealism.

Since the cost proposals must be evaluated by the Government in some detail in order to determine whether they are realistic and complete, it is important that the solicitation specify a work breakdown structure (WBS) tailored to the procurement and specify the type of data required to backup the cost proposal. By specifying a WBS, consistency among offerors is increased and the WBS can be tailored to break out those areas in which cost visibility is desired, such as areas where risk is considered greatest. Thus, a well structured WBS can simplify the job of cost evaluation.

By carefully specifying the type of cost proposal backup data, the job of cost evaluation can also be greatly simplified. Examples of some of the backup data that might be requested are: a list of material high dollar items; a list of all tools, test equipment and facilities; and monthly manloading charts by functional labor categories.

It should be noted that the Government has a legal right to detailed cost and pricing data. Furthermore, DOD Directive 4105.62 states that the "burden of proof of cost credibility rests with the offeror."

GOVERNMENT ESTIMATE

The key to assessing the realism and reasonableness of cost proposals is the development of a realistic Government cost estimate for each offeror, which can be compared against that offeror's proposed cost. While a cost proposal can be analyzed and its realism judged on a piecemeal basis (e.g., examining a few major cost elements and commenting on the degree of realism), such an approach is generally not sufficient. The Government should make an estimate of what it considers to be the most likely cost for each offeror. Since the Government estimates are compared to the offeror's proposed costs to determine the overall degree of cost realism and since the Government estimates for each offeror are compared to each other to evaluate estimated cost, the Government estimates themselves must be realistic and reasonable.

Basically, two types of Government estimates are used in source selection evaluations -- (1) Government "proposal-based" estimates based on evaluation of the offerors' cost proposals and (2) "independent" estimates based on parametric and analogy techniques and not on offerors' cost proposals. Sometimes a source selection will use both types of estimates, sometimes only one, and sometimes a blend (i.e., independent for some cost elements and proposal-based for others). For the purpose of evaluating cost realism and estimated cost, emphasis should be on the proposal-based estimates since these more clearly reflect differences between offerors, as well as the specific requirements of the program. However, it is desirable to make independent estimates for a program as well, since these can be used to cross check the proposal-based estimates.

Developing proposal-based estimates requires a detailed analysis of direct and indirect costs. Direct costs are direct labor, direct material, and other direct costs which are specifically and uniquely attributable to a particular product or service. Direct costs are the "nucleus" around which total contract price is built since the indirect costs (overhead, G&A, and fee) are generally a function of direct costs.

As mentioned above, it is desirable to make independent estimates to cross check with the proposal-based estimates. Such estimates made prior to receipt of the contractors' proposals can be compared with the contractor proposals to indicate areas where the detailed cost evaluations should be directed (i.e., areas that have the greatest potential payoff). This would reduce the amount of analysis and time required during the detailed cost evaluations. Independent cost estimates prepared for budget purposes or program review purposes may only require updating and tailoring to the specific requirements of the solicitation.

Technical risks must be considered in preparing the Government estimate. Engineers from the program office source selection organization and the Defense Contract Administration Service (DCAS) may assist in the risk assessment and cost evaluation. It also helps if some of the cost analysts preparing the Government estimate have an engineering background.

Since there are technical and program (e.g., schedule) risks and uncertainties associated with any program, as well as uncertainties in cost estimating data and techniques, there is often considerable uncertainty in the Government estimate. Therefore, it is important

to take into account the uncertainty associated with the Government estimate when judging the realism of the offeror's proposed cost. The greater the degree of uncertainty in the Government estimate, the more admissible are greater differences between the proposed cost and the Government estimate.

It is also important to achieve consistency in the Government estimates for different offerors. All offerors are proposing to the same requirements, but their approaches will differ and the requirements will not be specific in all areas.

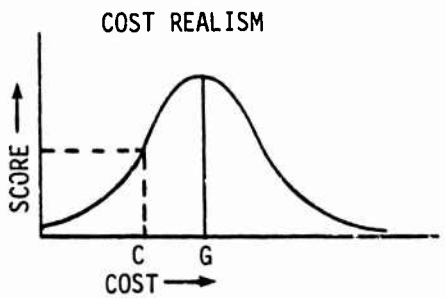
SCORING

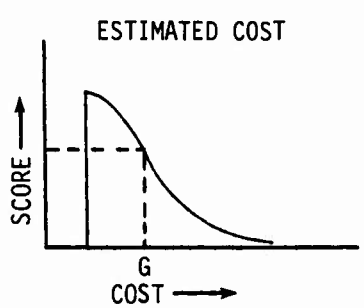
Considerable information is developed on the strengths and weaknesses of each of the offerors' proposals during the source selection evaluation of technical, cost and other factors. This information needs to be summarized and structured in such a way as to provide the Source Selection Authority (SSA) with a sound basis for making an award. Generally, the offerors are scored or ranked for each of the evaluation sub-criteria, which are then weighted to produce an overall score. The evaluators may assign numerical scores (e.g., zero to 10) or adjectival scores (e.g., exceptional, acceptable, etc). Sometimes, the adjectival scores are converted to numerical scores so that an overall score can be developed. Certain evaluation sub-criteria -- particularly those related to cost -- may not be scored at all, at least not numerically. However, in such cases a narrative evaluation would generally be provided.

Numerical scoring of both cost realism and estimated cost appears to have considerable merit. Nevertheless, it is presently discouraged by some Defense organizations. The advantage of scoring cost numerically, assuming an appropriate methodology is used, is that it provides the SSA with a clearer understanding of each offeror's relative ranking than an adjectival or narrative ranking does alone. However, the SSA needs sufficient flexibility and discretion with any scoring system to award to the offeror who provides the best value, all factors considered. That is, the SSA does not want to be driven to an irrational decision based on an overly elaborate and too inflexible scoring system. Careful selection of evaluation criteria weights and scoring rules for each criterion can minimize any problems.

The value of quantitative cost evaluations in source selection can be shown through an analogy with cost-effectiveness analysis. Cost-effectiveness techniques were developed as a means of quantifying, to the extent possible, various complex problems. Cost-effectiveness analysis has sometimes been criticized because it takes into account only quantifiable factors. This is an unfair criticism because it was never intended to take into account all factors, nor was it intended to drive decisions. Rather, cost-effectiveness analysis provides input to "decision makers" who must also consider the assumptions and caveats associated with the quantitative (cost-effectiveness) analysis, as well as the many non-quantifiable factors. Likewise, numerical scoring of cost can be a useful input to the SSA -- the source selection decision maker -- which is considered along with many other factors.

The basic notion for determining a numerical score is very straightforward: for cost realism, the greater the difference between the Government estimate (G) and the offer (C), the lower the score; for estimated cost, the greater the Government estimate, the lower the score. This is shown graphically below:





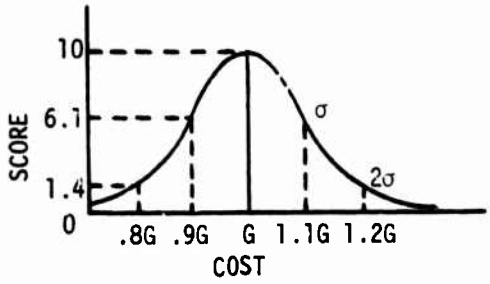
We refer to these curves as scoring rules and prefer smooth continuous curves based on a normal curve, as shown above. However, the curves could be some other shape (e.g., a triangle or trapezoid).

Cost Realism

In the case of cost realism, if the difference between G and C is zero, the score is a maximum - say 10. As the difference increases the score is reduced and approaches a minimum - say zero. For the same difference above or below G, we prefer to assign the same score (i.e., symmetry is assumed).

With a normal curve shaped scoring rule, the score depends on one factor, the "spread" of the normal curve as represented by the standard deviation, σ. There is a 68% likelihood that the difference between C and G will be between 0 and σ and a 95% likelihood that the difference will be between 0

and 2σ. For convenience we will define the spread as 2σ. For example, if 2σ = 20% of G, then an offer 10% higher or lower than G would receive a score of 6.1 and an offer 20% higher or lower would receive a score of 1.4. This is illustrated as follows:



The equation for the scoring rule is:

$$\text{Score} = 10e^{-\left(\frac{C-G}{G}\right)^2/2\sigma^2}$$

If the spread of the scoring rule is too small, there will not be sufficient differentiation between offers. And, if the spread is too large, even very unrealistic offers could still receive scores well above zero and large differences in realism will result in only small differences in score.

A spread (2σ) of 15 to 25% appears to be reasonable for defining a cost realism scoring rule. Offers more than 15 to 20% from a carefully prepared Government estimate, relying primarily on the offerors' own data, certainly raise a question as to cost realism. Of course, what is realistic will depend to some extent on the nature of the development program -- the risks involved, how much of the system has been previously designed, etc. It seems reasonable to assume that offers more than 15 to 25% from the Government estimate, depending upon the particular program, are unrealistic and should receive close to a minimum score. The low end of this suggested range, 15%, would apply only if the program was well defined, there were few risks or unknowns and there existed a good data base on which to base cost estimates.

A range of 15 to 25% for the scoring rule spread is supported by data for eight NAVELEX source selections. Cost realism (i.e., C - G) was determined for 19 offers made on eight programs. A normal curve fitted to a histogram of the data resulted in a 2σ of 26%. The average difference between C and G was 9%, and 84% of the time the difference between C and G was less than 25%.

Two things determine how cost realism will affect the overall proposal evaluation score: spread and weighting. First, the spread of the scoring rule should provide a reasonable differentiation of scores for contractors with differences in C - G, as mentioned above. Second, the weighting of cost realism relative to technical and other factors should be reasonable. Analysis and tradeoffs of the effects of different scoring rule spreads and different weights should be made as discussed later.

Uncertainty

The cost realism scoring rule is first of all a scaling rule -- the greater the difference between C and G, the lower the score -- but it should also take into account the uncertainty in the Government estimate. Ling and Wallenius [1] point out that any cost realism scoring rule can make two kinds of errors:

> Type A error -- assigning a higher score than warranted to an improbable proposed cost, and
>
> Type B error -- assigning a lower score that deserved to a highly plausible proposed cost.

Type A errors result if the spread in uncertainty is less than the spread in the scoring rule. This does not appear to be of concern as far as scoring is concerned. Although various offers may be unrealistic -- clearly more or less than they would have been had the offerors been realistic or more careful -- the scoring rule should differentiate among these unrealistic estimates, at least to a point.

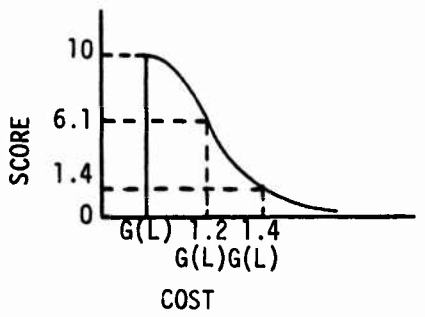
Type B errors result if the spread in uncertainty is greater than the spread in the scoring rule. This is undesirable because offers with a reasonable likelihood of matching the Government estimate would receive a score close to the minimum. If the spread (2σ) in the scoring rule is 15 to 25%, then the uncertainty spread (2σ) would have to be greater than 15 to 25% for a type B error to occur. Generally, we believe the uncertainty spread will be less than this since the Government estimate is based primarily on the offerors' own data and the stated requirements in the solicitation (e.g., no scope changes or redirections).

In conclusion, unless there is concern that the uncertainty spread is greater than the scoring rule spread, it does not appear necessary to determine uncertainty in the Government estimate for a particular program or to adjust the spread of the scoring rule accordingly. Benefits other than to cost realism scoring may, however, derive from determining the uncertainty. An uncertainty assessment would clearly lead to a better understanding of the cost estimates and would highlight certain areas warranting closer attention in the source selection evaluation.

Estimated Cost

It was mentioned above that the basic notion for scoring Government estimated cost is very straightforward: the greater the Government estimate, the lower the score. As a practical matter, we prefer to assign the maximum score say 10 - for the offer with the lowest Government estimated cost, G(L). As the difference between the estimated cost (G) for any offeror and the lowest estimated cost (G(L)) increases, the score is reduced.

With a normal curve shaped scoring rule, the score depends on one factor, the spread of the normal curve. For example, if 2σ = 40% of G(L), then an offer 20% higher than G(L) would receive a score of 6.1 and an offer 40% higher would receive a score of 1.4. This is illustrated below:



If the spread is too small, there will not be sufficient differentiation between offers. And, if the spread is too large, even very expensive offerors (as measured by the Government estimates) will still receive scores well above the minimum.

At what point should the Government estimate for an offeror be considered too high relative to the Government estimates for other offerors? Given that all offerors bid to the same set of requirements, their costs should not differ widely except possibly to account for such factors as variant design approaches, labor and overhead rates, and efficiencies. It seems reasonable that there would be a much greater likelihood that G would be reasonably close to G(L), say within 10 to 20%, than far from G(L), say 50 to 100%. To test this, we examined Government estimates for eight NAVELEX programs. A normal curve fitted to a histogram of the data resulted in a 2σ of 41%. The Government estimate or estimates which were not the lowest were 16% greater on average than the G(L) for each program. Half the time the Government estimate that was not the lowest was within 10% of G(L) and two-thirds of the time within 20%. The greatest difference from G(L) was a little more than 40%. Therefore, a spread of 30 to 50% seems reasonable.

Analysis and tradeoffs of the effects of different scoring rule spreads and different weights should be made, as discussed next

Combined Scores

The overall cost evaluation score depends on the scores and weights for cost realism and Government estimated cost. The overall source selection evaluation score depends on the scores and weights for each factor -- cost, technical, management, etc. The cost evaluation team with assistance from the technical team, as appropriate, should analyze the effects of different cost scoring rules and different weights. It is especially important to examine the tradeoffs between the cost and technical scores. For example, if one offeror has an estimated cost 20% greater than that for another offeror, how much better would his technical score have to be to offset his cost disadvantage. If such trades are not reasonable, then the spread of the scoring rule or the weights require adjustment.

Since the scoring rules and weights must be established before the solicitation is issued, these analyses and tradeoffs must also be done before the solicitation is issued.

Tabular and Graphical Displays

Despite the advantages of numerical scoring and the suggestions for avoiding or minimizing the chances of being driven to irrational decisions, some Defense organizations and source selection teams may prefer not to score numerically.

In these cases, it is suggested that key cost data be presented to the SSA in a tabular form such as shown below:

CONTRACTOR	C	G	$\frac{C-G}{G}$%	$\frac{G-G(L)}{G(L)}$%
A				
B				
C				
.				
.				
.				

The percentage columns provide a clear, convenient display of the degree of realism of each contractor and the estimated cost premiums compared to the lowest cost offeror (as estimated by the Government). While this tabular display is simple and may appear trivial, it can be of great value to the source selection team and the SSA. Graphical presentations of key cost data may also be useful. Even when scoring numerically, it would still be useful to provide the SSA with tabular and graphical presentations.

CONCLUSIONS

This report has developed and recommended a source selection cost evaluation methodology involving four interrelated steps:

(1) Evaluation: Determination of cost evaluation factors;

(2) Solicitation: Preparation of instructions to be included in the solicitation concerning the cost evaluation factors;

(3) Government Estimate: Preparation of Government estimates for each offeror; and

(4) Scoring: Scoring each offeror for cost realism and Government estimated cost.

This recommended methodology synthesizes and improves the best techniques and procedures currently being used in source selection cost evaluation.

It is believed that application of this methodology will encourage contractors to make more realistic offers since they will be on notice, though the solicitation, that cost realism will be examined and could be a deciding factor in the award. The contractors will know that the numerical scoring procedures will reduce the likelihood of awards to low offers on cost reimburseable contracts unless realistically low.

REFERENCE

[1] Ling and Wallenius, Cost Realism , Technical Report #400, Department of Mathematical Sciences, Clemson University, December 31, 1982