Estimating the Standard Error of Projected Dollar Gains in Utility Analysis

Ralph A. Alexander and Murray R. Barrick
University of Akron

Although attention given the utility analysis of personnel interventions has substantially increased recently, researchers have not addressed the problem of the standard error of utility estimates. The method for estimating such standard errors is presented and demonstrated in this article.

In the past few years there has been a resurgence of interest in utility or cost–benefit analyses of personnel interventions. Much of the impetus stems from Schmidt, Hunter, McKenzie, and Muldrow (1979). Their empirical method for estimating the dollar value of the standard deviation of performance ($SD_p$) brought about a renewed focus on utility research by providing one means for overcoming a serious obstacle to the application of utility estimation. Since 1979, alternative methods for estimating $SD_p$ have been developed (Cascio, 1982; Cascio & Ramos, 1986; Eaton, Wing, & Mitchell, 1985) and a number of authors have called for wider use of utility analysis in personnel research (Cascio, 1980; Landy, Farr, & Jacobs, 1982; Schmidt, Hunter, & Pearlman, 1982). Although most utility analyses have evaluated selection procedures, recent work has extended its use to other human resource interventions such as recruitment (Boudreau & Rynes, 1985), training (Landy et al., 1982), promotions (Cascio & Ramos, 1986), and employee flow through the work force (Boudreau & Berger, 1985).

Throughout the literature on utility estimation there is an important omission. Neither in the early developmental work on utility analysis nor in the recent refinement, extension, and application literature do we find a consideration of the standard error of estimates of utility gains from personnel interventions. Although it is obviously useful for an organization to have an estimate of the expected value of returns from investments in human resources, the variance associated with such estimates is also crucial to organizational decision making, if for no other reason than to determine whether the confidence interval includes zero.

Most if not all of the variables entering a typical utility analysis have some variability (or uncertainty) associated with them (Alexander & Cronshaw, 1984; Alexander, Cronshaw, & Barrick, 1986; Cronshaw & Alexander, 1983, 1985), and most recent empirical studies of utility analysis report standard errors for at least some of these variables. In none of this literature is the extension made to the standard error of the overall utility estimate.

This article details the method for calculating this standard error and demonstrates its use. Before proceeding with the calculations we review briefly both the classic utility model and recent capital budgeting modifications of that model.

Classic Utility Model

The classic utility model (Broden, 1949; Cronbach & Gleser, 1965) is expressed as some variant of the following equation:

$$ U = (T)(N_s)(SD_p)(r_{xy})(\lambda/\phi) - (C_s)(N_s)/(\phi), \quad (1) $$

where $U =$ the dollar value payoff resulting from the human resource intervention (e.g., a selection program); $T =$ the time period duration of the intervention (i.e., the average tenure of selectees); $N_s =$ the number of individuals impacted (the number of selectees in the case of selection utility); $SD_p =$ the standard deviation of performance in the present employee group in dollars; $r_{xy} =$ the validity of the (selection) intervention as a predictor of performance; $\phi =$ the selection ratio; $\lambda =$ the height of the normal curve associated with $\phi$; and $C_s =$ the per applicant cost of implementing the selection program.

Cronshaw and Alexander (1983, 1985) and Boudreau (1983) modified this basic model to take the time value of money into account. Simply stated, this means that a dollar earned by an organization in the present year is worth more than a dollar earned at some future time, if for no other reason than that the present dollar can be invested at prevailing interest rates. Cronshaw and Alexander (1983, 1985) showed that this is easily accounted for by applying a capital budgeting model to Equation 1 to give the following equation:

$$ U = \sum_{r=1}^{T} \frac{(N_s)(SD_p)(r_{xy})(\lambda/\phi)}{(1 + i)^T} - (C_s)(N_s)/(\phi) $$

$$ = \frac{(1 + i)^T - 1}{i(1 + i)^T} (N_s)(SD_p)(r_{xy})(\lambda/\phi) - (C_s)(N_s)/(\phi), \quad (2) $$

where $i =$ the cost of capital (often referred to as the discount rate) and all other terms are as previously defined. In this form, $U$ would be referred to as the net present value (NPV) of the intervention. Cronshaw and Alexander (1985) showed a similar derivation for other useful capital budgeting indexes such as internal rate of return, payback period, return on investment, and so on. A comparison of Equations 1 and 2 shows that they are
 identical except for the term associated with the time duration of the intervention: \( T \) in Equation 1 becomes \([(1 + i)^T - 1]/\( i(1 + i)^T \), which can be thought of as the discount-adjusted duration of the intervention. It is worth noting that to financial decision makers this capital budgeting modification is essentially a standardizing transformation that permits the comparison of projects that have different time spans (Beenhakker, 1976).

With these two forms of the utility estimate, we now turn to the problem of estimating the variance of \( U \). The discussion will be facilitated by considering both Equations 1 and 2 in their generic form, with utility being the difference between returns and costs:

\[
U = R - C. \tag{3}
\]

Variance of a Function of Random Variables

To compute the variance in a composite function of variables, we begin by recalling that given some variable \( X \) and constants \( a \) and \( b \) such that \( Y = a + bX \), the variance in \( Y \) is

\[
S_Y^2 = b^2S_X^2. \tag{4}
\]

Using this relation to find the variance in estimated \( U \) (from Equations 1 or 2), we must separate the variables \( (T, N_r, SD_r, \text{etc. in the returns component, and } C_a, N_a, \text{ and } 1/\phi \text{ in the costs component}) \) into those that are random variables or variables measured with error (the \( X \)'s) and those that are considered constants. The product of the expected values of the constants will become \( b \) in Equation 4, and the variance of the product of the random variables will become \( S_Y^2 \). Both \( b \) and \( S_Y^2 \) are found separately for the returns \( (R) \) and costs \( (C) \).

We are left, then, with finding an expression for the variance of the product of two or more random variables. The formula for the variance of the product of random variables when the population means and variances of the variables is known has been in the statistical literature since the 1930s and was (apparently) independently derived for use in meta-analysis in Hunter, Schmidt, and Jackson (1982, pp. 77, 85). Goodman (1960) dealt with the estimation of this variance in the two-variable case when the means and variances of the individual variables are sample estimates. Goodman (1962) generalized this to any number of random variables. His Equation 4, in which \( X = \Pi X_i \), is the following:

\[
S_{\Pi}^2 = \prod (M_i^2 + S_i^2(n_i - 1)/n_i) - \prod (M_i^2 - S_i^2/n_i), \tag{5}
\]

where \( \Pi \) is the product over the \( i \) random variables; \( S_i^2 \) is the variance of the \( i \)th variable; \( M_i^2 \) is the squared mean of the \( i \)th variable; and \( n_i \) is the sample size on which the mean and variance of the \( i \)th variable are based.

To summarize, then, we first find (separately) the variances in the cost component \( (S_{\Pi}^2) \) and the returns component \( (S_{\Pi}^2) \) of Equations 1 or 2 by the methods previously described. Because costs and returns can reasonably be expected to be independent of each other, the variance in the overall utility is simply the following equation:

\[
S_U^2 = S_R^2 + S_C^2. \tag{6}
\]

Application of the Method

Schmidt et al. (1979) analyzed the dollar-value utility that might be expected from using the Programmer Aptitude Test (PAT) to select computer programmers in the federal government. Their analysis was based on the classic utility model (Equation 1). Cronshaw and Alexander (1983) applied the capital budgeting modification of that model (Equation 2) to the Schmidt et al. (1979) data to demonstrate the NPV model calculations. Because the information needed to calculate the standard error of the utility estimates is available or can reasonably be estimated for the Schmidt et al. PAT utility analysis, that data will be used to demonstrate the method.

Schmidt et al. (1979) argued that utility analyses are often best conducted by analyzing the data for a range of values for the selection ratio. We will say more about this later; but for purposes of comparison, the utility analysis was conducted at four selection ratios (.10, .20, .50, and .80). For each of these, \( \lambda/\phi \) is treated as a constant. The value of \( \lambda \) (the height of the normal curve) can be found for a particular selection ratio, \( \phi \), from tables of the standard normal distribution. The \( \lambda/\phi \) constants for these selection ratios are 1.755, 1.40, 0.7978, and 0.350, respectively.

Schmidt et al. (1979) also reported values for \( SD_r (\hat{M} = 10,413; SD = 1,354; n = 105) \), \( N_r (\hat{M} = 610; SD = 43; n = 2) \), an estimated true validity of .76, an average tenure \( (T) \) of 9.69 years, and a cost per applicant \( (C_a) \) of $10. The \( r_{xy} \) estimate is from a validity generalization study, and Schmidt, Gast-Rosenberg, and Hunter (1980) reported an adjusted standard error of .25 \((n = 42)\) for that value. The determination of this \( n \) size will be discussed in detail later. Schmidt et al. provided no information on the standard deviation or \( n \) size for their estimate of tenure. An informal survey of six large organizations, each of which employs more than 100 computer programmers, found a value of \( SD_T = 3 \) to be a reasonable approximation on the basis of an \( n \) of 325.

This provides all of the information needed to calculate the standard error of utility estimates, \( SE_U \), under the classic model (Equation 1). For Equation 2, recall that \( T \) is replaced by a term reflecting the discount-adjusted time duration of the intervention. Most recent capital budgeting utility estimates have used a discount rate, \( \delta \), of .10. The \([(1 + i)^T - 1]/\( i(1 + i)^T \) \) term for \( i = .10 \), and \( T = 9.69 \), is 6.029. The exact variance of that term is far too cumbersome for the present context. A sufficiently close approximation to the standard error can be obtained by multiplying the standard deviation in \( T \) by the ratio of the discount-adjusted time to the unadjusted time—that is \([6.029/9.69] \times 3 = .622 \times 3 = 1.867 \). Simulations using values of \( T \) and \( i \) typical of utility analyses shows this approximation to underestimate the actual standard error by as much as 10%.

Table 1 gives a step-by-step example of the calculation of \( SE_U \). Table 2 gives the values of \( U \) for selection ratios of .10, .20, .50, and .80 for both the classic and NPV utility models. The standard errors were found by first applying Equation 5 to find the variance in the product of four random variables \( (T, N_r, SD_r, r_{xy}) \), then Equation 4 was used with \( b = \lambda/\phi \) to find the variance in the returns \( (S_R^2) \). Application of Equation 4 with the variance in \( N_r \) and \( b = C_a/\phi \) yielded the variance in cost \( (S_C^2) \). Finally, \( S_U^2 \) was calculated as the sum of \( S_R^2 \) and \( S_C^2 \) (Equation 6).
Table 1
Sample Calculation of SE(U) for the Classic Utility Model With a Selection Ratio (φ) of .10

<table>
<thead>
<tr>
<th>Equation</th>
<th>Step 1. Compute ( S^2_T ) for the returns component (Equation 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S^2_T = \Pi [\bar{M}^2 \pm S^2_T \sqrt{(n_i - 1)/n_i}] - \Pi (\bar{M}^2 \pm S^2_T \sqrt{n_i/n_i}) )</td>
<td></td>
</tr>
<tr>
<td>where X = ( T; \bar{M}_T = (9.69)^2; S^2_T = (3.0)^2; n_i = 325 )</td>
<td></td>
</tr>
<tr>
<td>X = ( N_i; \bar{M}_N = (610)^2; S^2_N = (45)^2; n_j = 2 )</td>
<td></td>
</tr>
<tr>
<td>X = ( S_D; \bar{M}_S = (10413)^2; S^2_S = (1354)^2; n_s = 105 )</td>
<td></td>
</tr>
<tr>
<td>X = ( r_m; \bar{M}_r = (7.6)^2; S^2_r = (25)^2; n_k = 42 )</td>
<td></td>
</tr>
</tbody>
</table>

Step 2. Compute \( S^2_R \) for the returns component (Equation 4)

\( S^2_R = (\lambda/\phi)^2 S^2_T = (1.775)^2 S^2_T \)

Step 3. Compute \( S^2_C \) for the costs component (Equation 4)

\( S^2_C = (C_T/\phi)^2 S^2_R = (10/10)^2 (45)^2 \)

Step 4. Compute \( SE(U) \) (Equation 6)

\( SE(U) = (S^2_T + S^2_R) \)  

Discussion

The use of utility analysis to assess the potential dollar gains to be realized by an organization's personnel interventions is becoming an increasingly acceptable and valuable tool to the personnel psychologist. Just as in other areas of measurement, the variate values on which such analyses are based—and thus the overall utility estimates—are subject to uncertainty (measurement error). It is by reporting the standard error of a measure that we convey information regarding the relative magnitude of that uncertainty.

We emphasize that Equations 1 and 2, and, consequently, the standard error calculations presented here, apply to the invariance of the selection ratio. Some situations, however, may call for the utility analysis to be performed for some empirically estimated selection ratios. In which case it will be necessary to determine whether that value is best treated as a random variable or a fixed constant. If, on the other hand, the selection ratio may vary, an estimate of the variance of \( \phi \) should be made (Alexander, Hanges, Kollar, & Alliger, 1986), in which case it can be used to evaluate the relative contribution of each of the random variables to overall dollar gains, and simulations are used to evaluate the likelihood of specific values of expected gains. Cronshaw, Alexander, Wiesner, and Barrick (in press) demonstrated the application of both of these procedures to utility analysis. Statistical approaches similar to the approach in the present study are much less common in the capital budgeting literature (Wagle, 1967).

Table 2 shows that for this particular data the standard errors of the utility estimates are relatively large. Financial decision makers may be more interested in the confidence interval of the projected gain from an investment than in the absolute value of the standard error. Table 2 also shows the 90% confidence interval of each estimate. In the present analysis we see that for the cases studied, the lower bound of the 90% confidence interval is well above zero, and thus the downside risk is minimal.

The analysis conducted here to demonstrate the method for calculating \( SE(U) \) followed the lead of Schmidt et al. (1979) and treated \( \lambda/\phi \) as a constant by analyzing utility at several selection ratios. Some situations, however, may call for the utility analysis to be performed for some empirically estimated selection ratio (Alexander, Hanges, Kollar, & Alliger, 1986), in which case it will be necessary to determine whether that value is best treated as a random variable or a fixed constant. If, for example, an organization uses a fixed cutoff on the predictor (Kroeck, Barrett, & Alexander, 1983), \( \phi \) could be treated as a constant (Alexander, Barrett, & Doverspike, 1983). If, on the other hand, the selection ratio may vary, an estimate of the variance of \( \lambda/\phi \) and of \( 1/\phi \) will also be needed. Alexander, Hanges, and Kollar (1986) tabulated these values for empirically determined estimates of the selection ratio.

Table 2
Expected Value, Standard Error, and 90% Confidence Interval of Expected Dollar Gains From the Use of the Programmer Aptitude Test for One Year to Select Computer Programmers in the Federal Government

<table>
<thead>
<tr>
<th>Utility model</th>
<th>Net present value (Equation 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection ratio</td>
<td>( U^* )</td>
</tr>
<tr>
<td>.10</td>
<td>82.0</td>
</tr>
<tr>
<td>.20</td>
<td>65.5</td>
</tr>
<tr>
<td>.50</td>
<td>37.3</td>
</tr>
<tr>
<td>.80</td>
<td>16.4</td>
</tr>
</tbody>
</table>

* Figures are in millions of dollars. ** Us for the classic model are comparable to those reported by Schmidt, Hunter, McKenzie, and Muldrow (1979). *** Us for the net present value model are comparable to those reported by Cronshaw and Alexander (1983).
The method is easily adapted to other situations. For example, in some settings the cost of implementing the intervention may be a constant, regardless of the number of individuals impacted (e.g., some training situations). In that case, \( C \) is a constant and referring back to the discussion of Equation 4, the variance in utility (\( SD_y^2 \)) will equal the variance in the returns (\( SD_t^2 \)). In a further refinement of the capital budgeting utility model (Equation 2), Boudreau (1983) included an additional multiplicative term in the returns (\( R \)) to account for organization costs that vary as a function of individual performance. He designated that term \( 1 + V \). Alexander, Cronshaw, Barrett, and DeSimone (1985) discussed the special case in which such variable costs may be estimated by a function of the pay-performance correlation. In these instances, an additional multiplicative term in \( R \) would be added to Equation 5.

This method for computation of the standard error requires two assumptions, and some discussion of them is in order. The first—required by Equation 5 for the variance of the product of random variables—is that the random variables, in this case \( T, N_y, SD_y, \) and \( r_{xy} \), are uncorrelated with each other (Bohrnstedt & Goldberger, 1969). There is no reason to expect that \( SD_y \) would be correlated with any of the other variables. If the assumption is violated, the most likely possibility would be that \( r_{xy} \) correlates positively with \( T \) and negatively with \( N_y \). That is, if the true validity were higher than the expected value used in the utility analysis, average tenure (\( T \)) might be greater and the required number of new hires (\( N_y \)) smaller than their sample-based expected values. It should be clear that such intercorrelations (in a particular single-cohort utility study) will be the result of correlated errors of estimation or correlated measurement errors. They will not arise if the usual assumptions of random and uncorrelated errors is met.

Although the statistical literature provides an adjustment to Equation 5 to deal with correlated variables (Bohrnstedt & Goldberger, 1969), such an adjustment is of limited usefulness because the information from which to estimate these intercorrelations will seldom be available. Note that the assumption of zero correlation among the random variables taken in this present application is the same assumption that is implicit in all utility analyses to appear in the industrial/organizational psychology (I/O) literature. Only if the variables are uncorrelated will the expected value of their product be equal to the product of their expected values. If the random variables are positively correlated, both the expected value (that is, Equations 1 and 2) and the standard error of the utility will be underestimated by the present methods (Goodman, 1960, 1962).

We emphasize that the assumption that the variables (\( T, N_y, \) etc.) are uncorrelated has to do with these intercorrelations in any single utility study. Intercorrelations among these variables across studies (jobs, organizations, etc.) are irrelevant to the assumption.

The second assumption is that the costs and returns are uncorrelated (Equation 6). The assumption is a reasonable one (Beenhakker, 1976), but once again the typical utility analysis setting will not be amenable to the estimation of the correlation. If costs and returns are positively correlated, Equation 6 will overestimate the standard error.

We comment on two other issues that arise from the use of this method. First, although the method is accurate for finding the standard error of \( U \), use of that value for constructing accurate confidence intervals depends on the distribution of \( U \) being approximately normal. In general, the distribution of \( U \) will be skewed (positively skewed when all expected values of the variables forming the product are positive; DeZur & Donahue, 1965). This skew will be at a maximum when the variables forming the product have equal variances and will become negligible as the ratio of the largest to second largest variance becomes quite large. As this ratio becomes very large, the distribution of the variable with the largest variance will come to dominate the shape of the distribution of \( U \). (Olcott, 1973, demonstrated this in detail for the product of two variables.) In the usual utility analysis, the values encountered are likely to satisfy this requirement sufficiently well to alleviate the potential problem of a skewed \( U \) distribution. In the PAT data (Table 2), for example, \( SD_y \) has the largest variance, 1,354, the next largest, 45. The ratio of these two values is extremely large. Because for most utility analyses of personnel interventions, \( SD_y \) will likely have the largest variance, the shape of the distribution of \( SD_y \) will dominate the shape of the distribution of \( U \). Schmidt et al. (1979) concluded that the distribution of \( SD_y \) for the PAT was not significantly nonnormal. In applications in which there is evidence for a substantial skew in \( SD_x \), the upper and lower confidence bounds for \( U \) should be calculated separately using the separate standard errors for \( SD_y \), from the upper and lower halves of the distribution.

The second issue is the matter of \( n \) sizes. In most cases the \( n \) sizes for Equation 5 will be straightforward. The demonstration in this article, however, illustrates where a problem may arise. Recall that the value for \( r_{xy} \) and its standard error were from a validity generalization study (Schmidt et al., 1980). The adjusted standard error of \( r_{xy} \) in that study was a function both of the number of studies (\( N = 42 \)) and the total number of subjects (\( N = 1,299 \)) in the meta-analysis. Statistical theory provides no guidance as to the “proper” \( n \) size for Equation 5 in such instances. The conservative approach is to use the smaller \( n \), which was done here.

Alexander, Cronshaw, and Barrick (1986) and Cronshaw and Alexander (1985) emphasized that in order for personnel interventions, such as selection systems, training programs, and so forth, to be treated on an equal footing with other investment decisions by organizations, it is necessary to abandon a cost mentality and treat the costs of obtaining, improving, and retaining valued human resources as investments that are expected to generate returns to the organization over an extended period of time. This shift of focus requires both a change in certain of the features of the utility model and in expressing the results of such analyses in a language that is well understood by financial decision makers. The classic Cronbach–Gleser utility model ignores a number of factors that are routinely dealt with in organizational financial analysis. Such considerations lead Cronshaw and Alexander (1983, 1985) and Boudreau (1983) to recast the classic utility model into the capital budgeting framework of Equation 2. Reference to Table 2 shows that this model has two other advantages over the classic model. The first is that it produces more conservative (and in the view of the financial decision makers, more comparable) estimates of utility. The NPV model also has a substantially smaller standard error.

The method is now available for calculating the standard error of estimated utility gains from personnel interventions. It is recommended that future studies report not only the expected...
value of such gains but also the standard error of these estimates.

References


