Journal of Applied Psychology
1992. Vol. 77, No. 3, 227-233

0021-9010/92/$3.00

Direct Estimates of SD_y and the Implications for Utility Analysis

Brian E. Becker and Mark A. Huselid
School of Management
State University of New York at Buffalo

Utility analysis suggests that human resources policies can have an economically significant impact on business organizations. Confidence in such conclusions, however, requires an accurate estimate of SD_y. This article provides a validity check on prevailing subjective methods of SD_y estimation by directly estimating SD_y from unique field data. Using both simulated and field data, we first illustrate the range of potential bias associated with predictor unreliability y in regression analysis and show how to calculate corrected values. We then discuss the methodological problems of directly estimating SD_y with organizational data and provide a range of estimates for SD_y. Our direct estimation of SD_y yielded values ranging from 74% to 100% of mean salary, which are considerably greater than conventional subjective judgments.

Recent work in the area of utility analysis suggests that human resources policies and interventions can have a significant economic influence on business organizations. Yet confidence in such conclusions must in large part turn on what has often been considered the weakest link in conventional utility analysis (Boudreau, 1991), the subjective estimation of SD_y. Although utility models will no doubt continue to dominate the field, we believe that more attention should be given to validating this approach with field research designed to directly estimate the magnitude of SD_y. The purpose of this article is to illustrate such a study and to discuss several of the methodological problems inherent in such an effort. In particular, we consider the problems and solutions associated with unreliability in the predictors in a regression model. On the basis of data uniquely suited for such a task, we show that changes in employee performance have a statistically and economically significant impact on firm profits. Compared with conventional benchmarks of 40% and 70% of mean salary, our calculations of SD_y fall in the range of 74%- 100% of mean salary.

Following recent work by Raju, Burke, and Normand (1990), we drew on accounting and economic concepts of firm performance and developed a simple model of organization performance as a function of individual employee performance. The fact that our data were drawn from a sample of retail outlets enabled us to directly estimate in dollars the effects of employee productivity. Although the model, and its limitations, are discussed in more detail later in the article, this brief description highlights the nature of our estimate, which is the change in firm profits associated with a one-unit change in an individual's performance rating. Because this result provided the basis for a straightforward calculation of SD_y, it allowed for a direct test of the 40% and 70% rules associated with Schmidt and Hunter's model (Boudreau, 1991; Vance & Colella, 1990).

We are grateful to Nancy Day for providing the field data used in this study.

Correspondence concerning this article should be addressed to Brian E. Becker, 268 Jacobs Management Center, State University of New York, Buffalo, New York 14260.

A second theme explored in this article is how the estimates of a regression equation are influenced by measurement error in an independent variable. The issue cannot be ignored in research of this kind because performance ratings are the most likely predictor and are often unreliable (Bernardin & Beatty, 1984; Heneman, 1986). The issue is discussed in some detail because the problem of predictor unreliability in regression analysis, though related to the familiar attenuating effect on correlations, is considerably more involved in multiple regression. To explicate the issue, we briefly review the statistical literature, provide a simple simulation of the problem under alternative scenarios, and, finally, illustrate the empirical effects in the analysis of our field data.

The Problem of Predictor Unreliability

The focus of our analysis is a regression model in which organizational performance (i.e., profits) is the dependent variable and individual employee performance is the independent variable of interest. The regression coefficient on employee performance, A, reflects the average change in organizational profit for each one-unit change in our measure of employee performance. Because A is the basis for our eventual calculation of SD_y, an inaccurate estimate of A will result in an erroneous estimate of SD_y. Unfortunately, the fact that we must use measured rather than true employee performance means that our estimate of A may be biased in the presence of predictor unreliability. In this section, we review both the nature and magnitude of potential biases for both the simple and multiple regression models.

In a simple regression equation (i.e., a dependent variable and a single independent variable), an unbiased estimate of the regression coefficient requires that the covariance between the error term and the independent variable be zero (Aigner, 1971, p. 31; Green, 1990, p. 157: Maddala, 1988, p. 34).[1] This assump-

[1] By unbiased, we mean that the expected value of our sample estimate of A is equal to the true value of A in the population.

tion is the basis for the normal equation that defines the regression coefficient (A_t) as

$$A_t = \frac{\mathrm{Cov}(R_t, Y_m)}{\mathrm{Var}(R_t)}, \quad (1)$$

or the covariance of measured Y and true R divided by the variance of true R. Y is firm performance and R is employee performance, and the subscripts t and m define true and measured values, respectively. Because Cov (R_t, Y_m) = Cov (R_t, Y_t), A_t is unaffected by measurement error in the dependent variable. This implies an estimation equation of the form

$$Y_m = c_o + A_t R_t + e, \quad (2)$$

where c_o is a constant reflecting the mean effect of all omitted variables. However, when R_t is replaced by measured employee performance (R_m), Equation 1 no longer holds because A_t does not equal A_m.

This a familiar problem in econometrics and has been part of the psychometric literature for at least 20 years (Goldberger, 1971). The proof follows from conventional statistical texts, such as econometrics, that emphasize regression analysis. For example, Maddala (1988, p. 381) presented the true model as

$$Y_t = c_o + A_t R_t + u, \quad (3)$$

where u is the random error term representing the effects of other causes of Y_t. The error term in this model has the same characteristics as the error term generated from a randomized experimental design. A regression of Y_t on R_t estimates the effect of a one-unit change in true performance ratings on the dollar value of true employee performance.

Following classical test theory, measured values can be expressed as a true score and random error component, such that

$$Y_m = Y_t + Y_e \quad (4)$$

and

$$R_m = R_t + R_e. \quad (5)$$

Substituting Equations 4 and 5 into Equation 3 yields the following:

$$Y_m = c_o + A_m R_m + w, \quad (6)$$

where $w = u + Y_e - AR_e$. However, although Cov $(e, R_t) = 0$ in Equation 2 and Cov $(u, R_t) = 0$ in Equation 3, Cov (w, R_m) is not equal to zero in Equation 6. In fact,

$$\mathrm{Cov}(w, R_m) = -A * \mathrm{Var}(R_e). \quad (7)$$

Maddala summarized the problem as follows:

> Thus one of the basic assumptions of least squares is violated. If only $[Y]$ is measured with error and $[R]$ is measured without error [Equation 2], there is no problem because [Cov$(e, R_t) = 0$] in this case. Thus given the specification in [Equation 6], it is errors in $[R_m]$ that cause a problem. (Maddala, 1988, p. 381)

To avoid confusion, we should point out that the algebraic derivation of Equation 6 would result in A_t as the coefficient on R_m. However, we use the term A_m because it reflects the estimate of A_t that would result if Equation 6 were estimated with ordinary least squares (OLS) regression. It is the properties of the error term in Equation 6 and their implication for estimation that are the basis for the different notation. The important point is that OLS estimation of Equation 6 will not yield the same estimate of A as either Equation 2 or 3. If A_m in Equation 6 does not equal A_t in Equation 1, what is their relationship? Estimating Equation 6 will yield the following expected value[2] for A_m:

$$\mathrm{E}(A_m) = \frac{\mathrm{Cov}(R_m, Y_m)}{\mathrm{Var}(R_m)}. \quad (8)$$

Because Cov (R_m, Y_m) = Cov (R_t, Y_t), given Equation 5,

$$\mathrm{E}(A_m) = \frac{\mathrm{Cov}(R_t, Y_t)}{\mathrm{Var}(R_t) + \mathrm{Var}(R_e)}, \quad (9)$$

and dividing by Var (R_t)/ Var (R_t),

$$\mathrm{E}(A_m) = \frac{A_t}{1 + \mathrm{Var}(R_e)/\mathrm{Var}(R_t)}. \quad (10)$$

As a result, A_m will underestimate A_t, and the degree of attenuation will depend on the ratio of error variance to true variance in R_m, or Var (R_e)/ Var (R_t). Maddala (1988, p. 382) concluded "that the least squares estimator of $[A_t]$ is biased toward zero and if [Equation 6] has a constant term, the leastuares estimator of [the constant] is biased away from zero."

Rearranging terms in Equation 10 yields the following:

$$\mathrm{E}(A_m) = \frac{A_t \mathrm{Var}(R_t)}{\mathrm{Var}(R_t) + \mathrm{Var}(R_e)} = A_t\left(\frac{\mathrm{Var}(R_t)}{\mathrm{Var}(R_t) + \mathrm{Var}(R_e)}\right), \quad (10.1)$$

$$= A_t\left(1 - \frac{\mathrm{Var}(R_e)}{\mathrm{Var}(R_t) + \mathrm{Var}(R_e)}\right), \quad (10.2)$$

$$= A_t - A_t\lambda, \quad (10.3)$$

where (Maddala, 1988, p. 382) concluded

$$\lambda = \frac{\mathrm{Var}(R_e)}{\mathrm{Var}(R_e) + \mathrm{Var}(R_t)}. \quad (11)$$

Lambda is simply the error variance of R_m divided by the total variance of R_m, or the reliability of R_m subtracted from 1. Using Equation 6 to estimate A_t will result in an estimate that is biased by the quantity $-A_t\lambda$. Therefore, if one attempts to estimate SD_y on the basis of Equation 6, the resulting value will understate the true value of SD_y.

Measurement Error in Multiple Regression

Up to this point, we have focused on the case of measurement error in a simple regression model with one independent variable. However, because utility estimates are developed outside of the laboratory one can easily imagine a regression model

[2] We use the term *expected value* here to reflect the expected sample value of A_m. Our purpose is to illustrate that A_m will systematically differ from the true value of A_t. However, the more precise term for this expected value is the probability limit of A_m as the sample size increases.

like Equation 6 with additional independent variables. These might be included to reduce error variance in the model to obtain a more precise estimate of *A* (i.e., a smaller standard error) or to reduce the chances of observing a biased or confounded estimate of *A*. Our purpose in this section is simply to identify the nature of the problem. The reader interested in more complete treatments should consult Maddala (1988), Judge, Griffiths, Hill, Lutkepohl, and Lee (1985), Green (1990), or Fuller (1987), among others.

The easier case is one in which the independent variables other than R_m are not measured with error. In this case, the bias in A_m is a function both of *A* and the collinearity between R_m and the other independent variables. Following Maddala (1988, pp. 383-384), if we normalize on all independent variables and define Cov (R_m, X_1)= where X_1 is the other independent variable, then

$$A_m = A_t\left(1 - \frac{\lambda}{1-\rho^2}\right), \tag{12}$$

and the bias for B_1 equals- (bias for A_m). At the extreme, when the two independent variables are uncorrelated, $\rho = 0$, and the bias in A_m reduces to the simple regression case. In the presence of some covariance between the independent variables, the downward bias in A_m is magnified. When λ is less than $1 - {}^2$, A_m will fall between A_t and zero. Otherwise, the bias will exceed the $A_t - 0$ bound, and A_m will have the opposite sign of A_t (Judge et al., 1985, p. 708; Maddala, 1988, pp. 383-385). For example, if $\lambda = .4$, $\rho = .8$, and $A_t = .4$, then $A_m = -.44$. As the correlation between the independent variables increases, it takes less and less measurement error in R_m to bias A_m away from zero with the opposite sign.

Finally, consider the extension to the multiple regression case in which both R_m and X_{m1} are measured with error. Now λ_r, and λ_x refer to the λ for R_m and X_{m1}, respectively. Then (Maddala, 1988, p. 388), where B_{m1} is the coefficient on the measured value of a second independent variable, X_{m1},

$$E(A_m) = A_t - \frac{A_t\lambda_r - \rho B_{t1}\lambda_x}{1-\rho^2}, \tag{13}$$

and

$$E(B_{m1}) = B_{t1} = \frac{B_{t1}\lambda_x - \rho A_t\lambda_r}{1-\rho^2}. \tag{14}$$

Now we can see that the bias in A_m is a function of the magnitude of A_t and B_{t1}, the relative magnitude of error variance in R_m and X_{m1} (λ_r and λ_x), and the correlation between R_m and X_{m1}.

Magnitude of the Problem

Although this discussion suggests that measurement error in R_m is a problem that must be considered in any attempt to accurately estimate *A*, and a problem that becomes more complicated in multiple regression, there is a well-developed and accessible analytical literature available in econometrics. Moreover, these solutions are all in terms of observable measures and estimated coefficients. Therefore, the degree of bias can be evaluated with some confidence. It is largely a question of calculating correlations among variables and reliabilities. In contrast to economists, who devote very little attention to measurement issues and normally lack reliability estimates of their measures, psychologists are at a distinct advantage.

In this section, we simulate the magnitude of potential bias in A_m under three scenarios. Case 1 is the two-variable case. Case 2 is the three-variable case, with one independent variable measured without error. Case 3 is the three-variable case with both independent variables measured with error.

Case 1

In Table 1, λ is calculated on the basis of Equation 11. The range is based upon the range of reliabilities one might normally observe in the literature. For example, r_{R_m,R_m} is an estimate of reliability for R_m and can be interpreted as "the percentage of true score variance in the fallible measure" (Nunnally, 1967, p. 181). Given that λ is the percentage of error variance in R_m, $\lambda = 1 - r_{R_m,R_m}$. The reader can see that A_m will understate A_t in absolute terms by 10%, even when very reliable measures of performance appraisal are used. At the other end of the range, A_t will be understated by 40% when r_{R_m,R_m} equals .6.

Case 2

Table 2 presents estimates of the quantity in the parenthesis of Equation 12. This in effect is the ratio of A_m to A_t. Now the bias in A_m is a function both of the measurement error in R_m and the correlation between R_m and X_1. Again we simulate results for what we believe are reasonable ranges for both the reliability of R_m and sample intercorrelations. Recall that the actual bias will depend on the relative values of λ and $1 - {}^2$. In general, the bias will tend toward zero in this literature because reliabilities will be relatively high. Recall as well that the bias will move away from zero only as the intercorrelations increases relative to the reliability of R_m. For example, if Equation 12 is rewritten such that

$$A_m = A_t - A_t\left(\frac{\lambda}{1-\rho^2}\right)$$

when $r_{R_m,R_m} = .5$ and = .8, then $\lambda = .5$ and $A_m = A_t - 1.389\,A_t = -.389\,A_t$.

Case 3

Table 3 shows how the ratio of A_t to A_m varies when the model includes two independent variables that are both correlated and measured with error. The estimates were derived by solving Equations 13 and 14. We solved for A_t in Equation 13 by first

Table 1
Case 1: Bias in A_m as a Percentage of A_t as Reliability in R_m Varies

Reliability of R_m	Bias in A_m
.6	$-.4\,A_t$
.7	$-.3\,A_t$
.8	$-.2\,A_t$
.9	$-.1\,A_t$

Table 2
Case 2: Ratio of A_m to A_t in Equation 12 as and the Reliability of R_m Varies

Reliability of R_m	Correlation of R_m and X_{m1}, .1	.3	.5	.7
.6	.595	.560	.466	.215
.7	.697	.670	.600	.411
.8	.798	.780	.733	.608
.9	.899	.890	.866	.800

solving Equation 14 for B_{t1} and substituting that result into Equation 13. The result was

$$A_t = \frac{[(1-\rho^2-\lambda_x)(\rho\lambda_x)(1-\rho^2)B_{m1}] - [(1-\rho^2)A_m].}{[(\rho\lambda_x)(1-\rho^2-\lambda_x)(\rho\lambda_r)] - [1-\rho^2-\lambda_r]}. \quad (15)$$

We arbitrarily set A_m and B_{m1} equal to 1.00 so that the solutions in Table 3 can be interpreted directly as the ratio of A_t to A_m. For example, the value of 1.10 in row 1, column 1 of Table 3 is interpreted as A_t*'s* exceeding the estimated value of A_m by 10%, when $\lambda_r = .1$, $\lambda_x = .1$, and = .1. Again, we selected λ and values comparable to those observed in the literature. Although the pattern is complex, within this range A_m will generally fall between A_t and zero. The chances that A_m will overstate A_t increase in samples with high intercorrelations and high λ_r.

Estimation of SD_y With Field Data

The preceding discussion has addressed the implications of estimating *A* directly using data on Y_m and R_m. One obstacle to direct estimation of SD_y is that the dependent variable, the dollar value of an individual's job performance, is generally unobservable. However, we propose substituting the firm's economic performance as the dependent variable to resolve this problem. As we discuss later in the article, the estimates from such a model are not unambiguous, but this approach can generate meaningful estimates of the dollar value of changes in individual job performance without the subjective estimation procedures required in other approaches to estimating SD_y.

Our data are part of a larger study that surveyed 335 first-line supervisors in the117 locations of a nationwide home-products retailing firm (Day, 1987). The dependent variable, Y_m, was return on sales, defined as the ratio of net income to gross sales for each store. The predictor, R_m, was the supervisor's quarterly performance appraisal (averaged over the year and combined into a single score for each subject). The subjects were department managers in one of five departments in each store. Unfortunately, data were often incomplete and did not include all four quarters of the appraisal. As a result, there were only 88 complete performance appraisals for the subjects in question.

We also used two additional predictors, years of education and organizational commitment. Commitment was measured with the nine-item version of Mowday, Steers, and Porter's (1979) Organizational Commitment Questionnaire (OCQ). Our choice of education and organizational commitment as additional predictors was largely an attempt to replicate the reliability characteristics of the independent variables in our prior simulations. The purpose was to illustrate how our estimates of SD_y vary with the reliability characteristics of R_m and the other independent variables.

Case 1: One Predictor

Return on sales and average quarterly performance appraisal *(R_m)* were the dependent and independent variables, respectively. The results of this regression are reported in column 1 of Table 4. The coefficient for R_m, .0922, is equivalent to A_m in Equation 6. Our estimate of r_{R_m,R_m} is the Cronbach's alpha (.74) for R_m, which is calculated directly from our data and yields a $\lambda_r = .26$. On the basis of Equation 10.3,

$$A_t = \frac{A_m}{(1-\lambda_r)} = \frac{.0922}{.74} = .125.$$

The bias in A_m equals $-A_t\lambda_r$ or $-.033$, so that A_m is approximately 75% of the true estimate. Given that the standard deviation of R_m is .27, the estimated SD_y is (.27)(.0922) = .025, and the true SD_y is (.27)(. 125) = .034. Because the dependent variable was return on sales, these figures are in percentages. However, multiplying these values by the mean net sales for the sample ($646,166) yielded a range in dollars from $16,154 to $21,970. These estimates are SD_y in terms of net income, not sales. This is an important distinction because these estimates include any costs associated with generating this higher performance. With an average salary of $21,888 in the sample, the ratio of SD_y to salary ranged from 74% to 100% of mean annual salary. Even if we were to assume that higher performers also earned higher salaries, the standard deviation of salary in the sample is only $1,722, so the magnitudes of the ratios would be relatively unaffected.

Case 2: One Predictor Measured With Error and One Predictor Measured Without Error

In the second example, we added an additional independent variable to the equation in Case 1. To illustrate the effect of another variable measured without error, we used years of education. This is not to deny that an employee may have lied about his or her education or that the information could have been recorded incorrectly, but generally education is a realistic example of a variable that one would not normally associate with measurement error. The results are reported in column 2 of Table 4. The coefficient for R_m, .0646, is equivalent to A_m. The

Table 3
Case 3: *Ratio of A_t to A_m as $_r$, and $_x$ in Equation 13 Vary*

Correlation (ρ) of R_m and X_{m1}	$\lambda_r = .1$: $\lambda_x = .1$	$\lambda_r = .1$: $\lambda_x = .3$	$\lambda_r = .3$: $\lambda_x = .1$	$\lambda_r = .3$: $\lambda_x = .3$
.1	1.10	1.09	1.42	1.41
.3	1.10	1.06	1.46	1.42
.5	1.12	1.08	1.63	1.59
.7	1.21	1.20	2.43	2.43

Table 4
Regression Results (Unstandardized Coefficients) for Field Data

Independent variable	Return on sales: Case 1	Case 2	Case 3	Equation 16
Constant				
Unstandardized coefficient	−0.2660	−0.0224	−0.2959	0.9916
SE	0.2030	0.2791	0.2511	2.5160
Performance appraisal rating (R_m)				
Unstandardized coefficient	0.0922	0.0646	0.0910	−0.2917
SE	0.0596	0.0632	0.0603	0.7454
Years of education				
Unstandardized coefficient		−0.0215		
SE		0.0169		
Organizational commitment				
Unstandardized coefficient			0.0065	
SE			0.0320	
Annual salary				
Unstandardized coefficient				−0.000053
SE				0.000174
Salary $* R_m$				
Unstandardized coefficient				0.000016
SE				0.000035
N	88	88	88	88
R^2	.027	.045	.027	.121

correlation between years of education and R_m is −.344. Following Equation 12,

$$.0646 = A_t\left(1 - \frac{.256}{1 - .118}\right),$$

so that A_t equals .091. True SD_y is then $16,150, or 79% of average salary. The estimate of A_t is not the same in Cases 1 and 2 because in Case 2 it reflects the effect of R_m, with years of education controlled. In addition, the bias in the second variable, years of education, is equal to (the bias in A_m). Therefore, the bias in B_1 equals −(−.344)(−.0264), or −.0091. In other words, B_1 is negatively biased in column 2 of Table 1, and the true estimate should be more positive by nearly 40%.

Case 3: Both Predictors Measured With Error

In the third example, we added a second predictor (organizational commitment) to Equation 6. Unlike in Case 2, however, this predictor was also measured with error. Cronbach's alpha for organizational commitment was .76, so λ_x in Equation 13 is .24. The correlation between organizational commitment and R_m was .102. The results of this regression are reported in column 3 of Table 4, where A_m = .091 and B_{m1} = 0.0065; λ_r and λ_x were .26 and .24, respectively. Solving Equation 15 yielded a value for A_t of .1226. Again, A_m substantially understated the true value of A_t. In this case, the bias is attributable to measurement error in both predictors, as well as the correlation between the two variables. However, given the relatively low value for in this sample, measurement error is the overwhelming cause of the bias. Again, an A_t value of .1226 implies an SD_y value of $21,318, or 96% of average salary. With rounding error, this is identical to the result in Case 1.

The results in Table 4 are interesting from two perspectives. On the one hand, they reflect the potential bias one might expect if Equation 6 is applied in practice. On the other hand, even these biased estimates suggest that an approach incorporating firm-level earnings data will yield SD_y estimates of considerable magnitude. These results are important because they bear directly on the continuing debate over the magnitude of SD_y. They suggest that, for this particular firm, in this particular industry, SD_y is economically significant and greater than previously suggested by much of the utility literature. However, the interpretation of these results must be tempered by the qualifications discussed in the next two sections.

These data also allow us to shed some light on the notion that SD_y is a constant percentage of salary. If in fact SD_y is a fixed percentage of salary, it follows that the absolute value of SD_y must increase with salary level to maintain the same percentage. This hypothesis is directly testable with these data. For example, the following regression model,

$$\text{Return on sales} = a_o + A_{m1}R_m + A_1\,\text{Salary} + A_2(R_m \times \text{Salary}) + w, \quad (16)$$

is Equation 6 with two additional independent variables, salary and an interaction between salary and R_m. The coefficient on the interaction term (A_2) will be positive when A_m increases with salary. In other words, A_m should be equal to A_m, + A_2 salary. The results of this regression are reported in column 4 of Table 4. Though the estimate is not statistically significant by conventional standards, the magnitude of the coefficient on the interaction term (A_2) means that the ratio of SD_y to salary actually increases with salary level. Recall that SD_y is equal to $A_m(SD_{R_m})$. . Therefore, given that A_m is −.29 + .000016(Salary), SD_{R_m} = .27, and annual sales average $646,166, the dollar value of SD_y increases by $2,791 for every $1,000 increase in salary. As a result, SD_y is just 25% of salary at $20,000 but 67% of salary at $24,000. The magnitude of these differences suggests the potential practical significance of these results, despite the statistical insignificance of the interaction term.

Related Estimation Issues

Our analysis of the field data raises additional estimation issues that researchers should consider. First, we would caution the reader that for purposes of hypothesis testing one must also consider the effect of measurement error on the variance of the regression coefficient. In general, the effects of measurement error on the t statistic for A_m are as complex as the effects on the magnitude of the coefficient. In the bivariate regression (column 1, Table 4), the t statistic for A_m will be attenuated in the same way as A_m. For example, if r_{R_m,R_m} is the reliability of R_m, and we define the true t statistic for A as t_A, and the observed t statistic for A_m as t_{A_m}, then, following Fuller (1987),

$$t_{A_m} = r_{R_m,R_m} t_A. \tag{17}$$

According to Fuller,

> Any linear hypothesis about [A] can be transformed into a hypothesis about [A_m] by using the reliability ratio. Therefore, in the bivariate situation, knowledge of [r_{R_m,R_m}] permits one to construct an unbiased estimator of the parameter [A] and to apply the usual normal theory for hypothesis testing and confidence interval construction. Unfortunately, these simple results do not extend to the vector-x case. (Fuller, 1987, p. 7)

A reanalysis of the results in Table 4 (column 1) illustrate the effects of measurement error in R_m on our hypothesis test. The results in Table 4 (column 1) reflect an underestimation of both the absolute magnitude of A and the statistical significance of the coefficient. In this example, $t_{A_m} = 1.55$ and $r_{R_m,R_m} = .74$, so the true t statistic (t_A) is 2.09 (1.55/.74) rather than 1.55. The coefficient is in fact statistically significant at the .05 level (two-tailed test).

The second issue reflects the problems of using firm- or unit-level earnings measures and individual-level performance ratings. This will require that multiple organizational units be available because the dependent variable is constant within units. However, it is conceivable that organizational policy is such that a unit, a store in this case, hires not just one good employee, but rather that every employee is above the sample average. Specifically, in this sample, there were five department managers in each store, and our data generally include one manager per store. There were 58 different stores and 88 different managers. If Store As policy is to hire better people across the board, then an individual manager's performance appraisal could be a proxy for all five managers in that store. This is also more likely because the appraisal scheme is based on absolute measures of performance, and all managers could receive high ratings. Therefore, we cannot distinguish whether the increment in sales associated with one manager's higher rating is attributable to his or her individual efforts or to the cumulative efforts of all five managers in a high performance store. As the between-store variance in R_m increases in proportion to the within-store variance in R_m, this problem will increase. At the extreme, true SD_y would be only one-fifth the size calculated earlier.

A third problem is the potential for what Econometricians call simultaneity bias. Namely, supervisors' performance appraisals could be influenced by their knowledge of store earnings during the period. The significance of this contamination depends on the influence of department managers on earnings relative to uncontrollable market factors. If the manager is largely responsible, then even if supervisor ratings are influenced by earnings, the direction of the underlying relationship goes from managerial performance to earnings. However, to the extent that earnings are instead associated with developments in the local product market and are out of the department manager's control, then we may be observing the effect of earnings on appraisal, rather than SD_y. The bias in this case would overstate the value of SD_y. Although well-developed statistical procedures to eliminate this bias are available, such a model is beyond the scope of this article. Nevertheless, we would caution researchers pursuing this line of inquiry to be aware of this problem. This particular problem may be partly mitigated in our data in two respects. First, there was a temporal sequence to the performance ratings and the earnings announcement. Official earnings were available only after the appraisal, though it is hard to imagine that store personnel did not have a reasonable idea of the magnitudes involved. Second, the performance appraisal process does provide behavioral anchors for the dimensions being rated.

Economics of SD_y Estimates

The use of accounting and economic theories as the basis for our estimation of A means that A is not necessarily a stable value over time. Earlier researchers have discussed the need to incorporate the duration of intervention effects and the time value of money into utility calculations (e.g., Boudreau, 1983), and this issue continues to be a source of debate in the literature (Cronshaw & Alexander, 1991; Hunter, Schmidt, & Coggin, 1988). However, our point is that A, and the resulting estimate of SD_y, is now a function of both the impact of employee performance on organizational output and the value of that output in the product market. Even if the effect of employee performance on organizational output is relatively stable over time, product market changes that are beyond the control of the employees will affect the economic value of their contribution to the organization. For example, in our sample of retail outlets, one could easily imagine sales and profits falling if competitors moved into these markets. Employees' performance could remain unchanged, yet the value of that performance (A) would fall. On this issue, conventional utility analysis and economic theory are in conflict. Although a full discussion of this point is beyond the scope of this article, we raise it as a caveat to caution researchers who may find it convenient to assume that SD_y is immutable.

Discussion

For a line of research that seems to have such direct implications for management decision making, the method of utility analysis has not been widely adopted in practice. We agree with Cascio and Morris (1990) that too often this literature has focused on issues that narrow its impact rather than expand its legitimacy and accessibility to managers. The overwhelming reliance on subjective estimates of SD_y is one such practice. We believe that a systematic program of research that provides direct estimates of SD_y across a variety of contexts not only will provide a reality check on the results of more subjective approaches but also will lend greater legitimacy to the method in general.

We also recognize that our approach does not have broad applicability among practitioners because of the specific organizational characteristics required for implementation. In fact, in cases where it could be used in a particular organization, firms might prefer to estimate the utility of human resources interventions directly without the intermediate estimation of SD_y. Rather, our intent is to motivate a program of research that will provide an objective, empirically based benchmark against which the prevailing subjective estimates of SD_y can be compared. This study is an initial step in that direction.

We have illustrated how, with an appropriate organizational structure, it is possible to estimate SD_y in terms of observable and available measures. Several methodological problems with this line of research were also explored in some detail. Particular attention was given to predictor reliability, a problem familiar to psychologists, but in the perhaps unfamiliar context of regression analysis. This article has demonstrated, with both simulation and field data, that predictor unreliability will result in biased estimates of SD_y and that, within the data characteristics of most utility studies, predictor unreliability will typically underestimate SD_y. Nevertheless, on the basis of our field study, we show that SD_y estimates are at the high end of the range suggested by prior subjective estimates. Indeed, our estimates SD_y in this sample ranged from 74% to 100% of mean salary.

Although such results imply dramatic effects for improved employee performance, a number of issues are suggested for future research. First, if utility researchers move into the field and model the actual operation of a firm, they will have to be more sensitive to the fact that they are engaged in an interdisciplinary line of inquiry. Economists, for example, have a well-developed literature on the economic value of information. This applies directly to efforts by organizational psychologists and personnel researchers to improve selection techniques and appraisal methods. The simple estimation model developed in this article will no doubt require considerable elaboration. Second, because this approach draws on actual firm earnings, with the associated impact of unique labor and product markets, a wide range of field studies will be required to determine the generalizability of SD_y levels. For example, we have no reason to believe that our results will translate directly to other firms or industries. Third, this study has touched only briefly on the precision of our estimates and some of the other statistical issues involved in this type of field work. We believe, however, that future work along these lines must incorporate existing econometric theory on these problems along with the knowledge that psychologists are in a unique position to utilize these procedures, given their reliability estimates of the measures in question.

References

Aigner, D. S. (1971). *Basic econometrics.* Englewood Cliffs, NJ: Prentice-Hall.

Bernardin, H. J., & Beatty, R.W. (1984). *Performance appraisal: Assessing human behavior at work.* Boston: Kent Publishing.

Boudreau, J. W. (1983). Economic considerations in estimating the utility of human resource productivity improvement programs. *Personnel Psychology, 36,* 551–576.

Boudreau, J.W. (1991). Utility analysis for decisions in human resource management. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., pp. 621-745). Palo Alto, CA: Consulting Psychologists Press.

Cascio, W. F., & Morris, J. R. (1990). A critical reanalysis of Hunter, Schmidt, and Coggin's (1 988) "Problems in using capital budgeting and financial accounting techniques in assessing the utility of personnel programs." *Journal of Applied Psychology, 75,* 410–417.

Cronshaw, S. F., & Alexander, R. A. (1991). Why capital budgeting techniques are suited for assessing the utility of personnel programs: A reply to Hunter, Schmidt, and Coggin (1988). *Journal of Applied Psychology, 76,* 454-457.

Day, N. E. (1987). *Clarifying the relationship between organizational commitment and performance.* Unpublished doctoral dissertation, University of Kansas.

Fuller, W. A. (1987). *Measurement error models.* New York: Wiley.

Goldberger, A. S. (1971). Econometrics and psychometrics: A survey of communalities. *Psychometrika, 36,* 83–107.

Green, W. S. (1990). *Econometric analysis.* New York: MacMillan.

Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology, 39,* 811–826.

Hunter, J. E., Schmidt, F. L., & Coggin, T. D. (1988). Problems and pitfalls in using capital budgeting and financial accounting techniques in assessing the utility of personnel programs. *Journal of Applied Psychology 73,* 522-528.

Judge, G., Griffiths, W, Hill, R. C., Lutkepohl, H., & Lee, T. C. (1985). *The theory and practice of econometrics.* New York: Wiley.

Maddala, G. S. (1988). *Introduction to econometrics.* New York: Macmillan.

Mowday, R. T., Steers, R. M., & Porter, L. M. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior, 14,* 224-247.

Nunnally, J. (1967). *Psychometric theory.* New York: McGraw-Hill.

Raju, N. S., Burke, M. J., & Normand, J. (1990). A new approach for utility analysis. *Journal of Applied Psychology, 75,* 3–12.

Vance, R. J., & Colella, A. (1990). The utility of utility analysis. *Human Performance, 3,* 123-139.

Received March 18, 1991
Revision received October 30, 1991
Accepted October 31, 1991 ■