Appendix B: Statistical Procedures Testing for Statistical Differences

The statistical difference between mean estimates is tested using a t-statistic. The null and alternative hypotheses to be tested are:

$$H_o: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

where μ_1 is the population mean of group 1 and μ_2 is the population mean of group 2. Evidence allowing rejection of the null hypothesis indicates a significant difference between population means of farms in the two groups. The t-statistic used for hypothesis testing is (see Kmenta, 1986, p. 137 and 145):

$$t \sim \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{VAR\,(\overline{X}_1) + VAR\,(\overline{X}_2)}}$$

where X_1 and X_2 are sample means, and $VAR(X_1)$ and $VAR(X_2)$ are variance estimates of the sample means. If the estimated t-statistic exceeds the critical-t value for the chosen level of significance then the null hypothesis can be rejected and the group means are deemed significantly different. At a 5-percent level of significance, this means that from infinite samples of both populations, only 5 percent of the time would the estimates lead to an incorrect rejection of the null hypothesis.

Decomposing Profit Variation

The statistical association between farm profits and several farm structural and performance characteristics is tested using a regression equation. The empirical model for farm profits is:

$$y_i = \alpha + X_i\beta' + \varepsilon_i$$

where y_i is the farm profit of the i^{th} individual, and x_i is a vector of farm structural and performance characteristics assumed to influence farm profits. The error term, ε_i, is assumed to have the usual desirable properties. Parameters of the model, α and β', are estimated using weighted least squares.

The measure of profit variation is the variance of profits, y_i. The variance of farm profits can be expressed as the sum of the variation explained by the model and the variation in the error term:

$$\sigma_y^2 = \beta'\Sigma\beta + \sigma_\varepsilon^2$$

where β' is a vector of parameter estimates, Σ is the variance-covariance matrix of explanatory variables, and σ^2_ε is the residual variation. To measure the extent to which each explanatory variable influences the variation of profits, the sample variation can be decomposed into its various components (Kmenta, 1986, p.410). Consider a partition of $x_i = [x_{1i}\ x_{2i}]$, with the corresponding partition $\beta = [\beta_1\ \beta_2]$. The variance of profits can be written as:

$$\sigma_y^2 = \beta_1'\Sigma_{11}\,\beta_1 + \beta_2'\Sigma_{22}\beta_2 + 2\beta'\Sigma_{12}\beta_2 + \sigma_\varepsilon^2$$

where Σ_{11} and Σ_{22} are matrices of variances for x_{1i} and x_{2i}, and Σ_{12} is the matrix of covariances between x_{1i} and x_{2i}. The first term on the right-hand side represents the amount of variation in profits that can be attributed solely to x_{1i}; the second term is the variation in profits explained solely by x_{2i} (variance effects). The third term arises from the covariance of x_{1i} and x_{2i} and cannot be separated into parts due only to x_{1i} or only to x_{2i}, but is attributed to the influence of the two groups of variables together (covariance effects).

Alternative Specifications of the Regression Equations

Three alternative specifications of regression equations are used to examine the various relationships presented in this report—linear, reciprocal, and quadratic.

The most commonly used and easiest to interpret is the linear form:

$$Y = \alpha + \beta X$$

Estimated parameters of this equation, α and β, indicate the intercept and slope, respectively, of the estimated equation. The estimate of β describes the unit change in Y with a unit change in X.

The reciprocal form is expressed as:

$$Y = \alpha + \beta \frac{1}{X}$$

The intercept estimate of the reciprocal form, α, represents the value of Y that is approached as X grows infinitely large. The estimate of β describes the unit change in Y with a unit change in 1/X. If β is negative, α represents a maximum value that is approached from below but never reached. Conversely, a positive value of β implies that α is a minimum that is approached from above but never reached.

The quadratic form includes the linear term plus a squared term:

$$Y = \alpha + \beta X + \delta X^2$$

The estimated value of α represents the intercept. The estimate of β describes the unit change in Y with a unit change in X and δ describes the unit change in Y with a unit change in X^2. If both β and δ are positive (negative), Y increases (decreases) at an increasing rate with X. If β is positive and δ is negative, Y increases at a decreasing rate and eventually reaches a maximum. Likewise, if β is negative and δ is positive, Y decreases at a decreasing rate and eventually reaches a minimum. The level at which a maximum or minimum occurs can be identified by setting the first derivative of the estimated equation to zero and solving for the value of X.