Appendix: Statistical Procedures

The Gini Coefficient

The Gini coefficient for production of each livestock commodity can be expressed as (Pyratt, Chen, and Fei):

$$G = 2COV(Y,F(Y)) \,/\, \overline{Y}$$

where G is the Gini coefficient; $\overline{Y}$ is the mean level of production; $F(Y)$ is the cumulative distribution of production; and $COV(.)$ is the covariance between production Y and the cumulative density function value for this production level. The Gini coefficient is a positive fraction between zero and one. With respect to livestock production, a Gini value of one means that all of production is concentrated in one individual, while a Gini value of zero indicates that all in the population produce an equal share of production.

Measuring Structural Change

Structural change is measured as a weighted index that describes the change in per-farm production from 1969 to 1992. The percentage change in per-farm production for each county is expressed as:

$$SC_i = \frac{PFUNITS_{i,92} - PFUNITS_{i,69}}{PFUNITS_{i,69}} * 100$$

where SC_i is the percentage change in per-farm production in county i from 1969 to 1992, and $PFUNITS_{i,69}$ and $PFUNITS_{i,92}$ are per-farm production in county i for 1969 and 1992, respectively. The index of structural change is developed by weighting each value of SC_i by the proportion of total production units provided by each county in 1992. Thus, the structural change index is:

$$SCI_i = SC_i * \frac{UNITS_{i,92}}{\sum_{i=1}^{n} UNITS_{i,92}}$$

where $UNITS_{i,92}$ is the number of production units from county i in 1992 and n is the number of counties with production of the commodity in 1992.

Testing for Significant Structural Change

The index of structural change between 1969 and 1992 for each county is statistically tested as to whether it is significantly greater than the mean of all U.S. counties. The null and alternative hypotheses to be tested are:

$$H_O\colon \overline{SCI} \geq SCI_i$$

$$H_A\colon \overline{SCI} < SCI_i$$

where $\overline{SCI}$ is the mean structural change index for the U.S. and SCI_i is the structural change index in each county. Evidence allowing rejection of the null hypothesis indicates that the index of structural change in the tested county is significantly greater than the U.S. average. A one-sided t-statistic is used to conduct the test (see Kmenta, pg. 144):

$$t \sim \frac{(SCI_i - \overline{SCI})\sqrt{n}}{s}$$

where n is the number of observations and s is the standard deviation of $\overline{SCI}$. If the calculated t-statistic exceeds the critical t-value for the chosen level of significance then the null hypothesis is rejected and structural change in the county is deemed significantly greater than the U.S. average. A 5-percent level of significance is used to conduct the test in this report, thus the critical t-value for the one-sided test is 1.65. A procedure similar to this was applied by Hoover, Mason, McKay, and Fraumeni to cancer mortality rates in U.S. counties.

Testing for a Statistical Difference of Group Means

The statistical difference between mean estimates for producers in each structural change area is tested using a t-statistic. The null and alternative hypotheses to be tested are:

$$H_O\colon \mu_1 = \mu_2$$

$$H_A\colon \mu_1 \neq \mu_2$$

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, pgs. 137 and 145):

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

where $\bar{X}_1$ and $\bar{X}_2$ are sample means, and $VAR(\bar{X}_1)$ and $VAR(\bar{X}_2)$ are variance estimates of the sample means.[13] 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.[14] At a 5-percent level of significance, this means that from infinite samples of both populations the estimates would lead to an incorrect rejection of the null hypothesis only 5 percent of the time.

Test of the Equivalency of Two Regressions

Statistical testing for a difference between coefficients of two regressions is used to compare the unit-cost equation estimated for each structural change area. Separate regressions are estimated for each area with the model:

$$Y = \alpha_0 + \sum_{k=1}^{m} \alpha_k X_k + \epsilon$$

where the α are parameters to be estimated, ϵ is the error term, and m is the number of explanatory variables. Data for farms in two areas are then combined. A dummy variable D is constructed with $D=1$ if the farm is located in the first area, $D=0$ otherwise. The regression model is then specified as:

$$Y = \alpha_0 + \sum_{k=1}^{m} \alpha_k X_k + \delta_0 D + \sum_{k=1}^{m} \delta_k X_k D + \epsilon$$

where the α and δ are parameters to be estimated, ϵ is the error term, and m is the number of explanatory variables. Coefficients estimated with the dummy variables, δ_0 through δ_m, measure the difference of the intercept (δ_0) and the slope estimate of each variable ($\delta_1 - \delta_m$) between the two areas. Therefore, t-statistics on the estimated coefficients indicate whether the estimated coefficients on each variable in the separate regressions for each area are significantly different.

Decomposing Cost Variation

To measure the extent to which each explanatory variable influences the variation of production costs, the sample variation is decomposed into its various components and expressed using the coefficient of separate determination (Burt and Finley). The components of unit-cost variation are:

$$\sigma_Y = \begin{matrix} \alpha_1^2\sigma_{11} + \alpha_1\alpha_2\sigma_{12} + \dots + \alpha_1\alpha_k\sigma_{1k} + \\ \alpha_2\alpha_1\sigma_{21} + \alpha_2^2\sigma_{22} + \dots + \alpha_2\alpha_k\sigma_{2k} + \\ \dots \\ \dots \\ \alpha_k\alpha_1\sigma_{k1} + \alpha_k\alpha_2\sigma_{k2} + \dots + \alpha_k^2\sigma_{kk} + \sigma_\varepsilon, \end{matrix}$$

where α_{ii} and α_{ij} $(i \neq j)$ are the variance of X_i and the covariance of X_i and X_j, respectively. Calculation of the coefficients of separate determination effectively allocates the explained variation of the regression model among the independent variables. Thus, the coefficients are computed as:

$$\begin{aligned} C_1 &= (\alpha_1^2\sigma_{11} + \alpha_1\alpha_2\sigma_{12} + \dots + \alpha_1\alpha_k\sigma_{1k})/\sigma_Y \\ C_2 &= (\alpha_2\alpha_1\sigma_{21} + \alpha_2^2\sigma_{22} + \dots + \alpha_2\alpha_k\sigma_{2k})/\sigma_Y \\ . &= \dots \quad \dots \quad \dots \quad \dots \\ . &= \dots \quad \dots \quad \dots \quad \dots \\ C_k &= (\alpha_k\alpha_1\sigma_{k1} + \alpha_k\alpha_2\sigma_{k2} + \dots + \alpha_k^2\sigma_{kk})/\sigma_Y. \end{aligned}$$

Each coefficient represents the portion of the variation in the dependent variable explained by each independent variable alone (variance effects), and the interaction among variables (covariance effects). The sum of the these coefficients equal the R^2 goodness of fit measure. One minus the sum equals the unexplained variation. Coefficients of separate determination were used to examine determinants of the profitability of cattle feeding by Langemeir, Schroeder, and Mintert; and of dairy farming by El-Osta.

[13]The FCRS uses a multiframe stratified sample. The formula used to compute variance estimates of sample means from FCRS data can be found in Dillard.

[14]For the sample sizes used in this study the critical t-value for a 5-percent level of significance is 1.96 and for a 10-percent level of significance is 1.65.