Probability Distributions of Crop Prices, Yields, and Gross Revenue

Bernard V. Tew and Donald W. Reid

This study shows that the price-yield correlation is a major influence in determining the skewness of revenue. Therefore, normality for revenue may not be rejected even if the price and/or yield distributions are significantly skewed. Analysis of cotton revenue for Mississippi shows that this can be the case empirically when the correlation between price and yield is moderately negative and the relative variability of yield and price is not too high. Hence, for crops produced in their major production regions where negative correlations between prices and yields are the greatest, revenue distributions may have a greater tendency toward normal.

Day's pioneering research on the distribution of crop yields gives *a priori* reasoning and evidence for nonnormally distributed yields and Gallagher's recent investigation of national average corn yields further supports this contention. Both of these studies were motivated mainly by the implications for erroneously predicting crop yields. However, Day alludes to the importance of farm planning implications of nonnormal yield distributions. He states, ". . . decisions for maximizing profit and minimizing risk must be based not only on expected yields and variance but also upon skewness as well" (p. 735). Day's study is, at least in part, the impetus for rejecting the expected value-variance (E-V) criterion for risk analysis because normality of returns is a sufficient condition for making the E-V criterion consistent with expected utility maximization.

Crop yields often are the only stochastic consideration in forming return distributions for risk analyses (e.g. Yassour et al.; Klemme; Collender and Zilberman; and Harris and Mapp). Perhaps this perspective occurs because, as Grant points out, there is a tendency to view yields and prices as uncorrelated at the individual producer level in a competitive market. However, when aggregate output and price are correlated and yields at the producer level are correlated, it follows from the additivity of covariance that producer-level yield and the market price do not have a zero covariance (Grant, p. 630). Certainly the case for positive correlation among producers and a negative covariance between aggregate output and market price can be made for major producing regions. Therefore, while yield distribution in risk analysis is important, perhaps the price distribution and the relationship between price and yield are equally important.

Little empirical research has been directed toward analyzing the relationship between price and yield distributions and the resulting revenue distribution. An exception is the recent work by Buccola, which investigates the determinants of the shape of the revenue distribution and analyzes two empirical situations in Oregon. Haldane's equations for variance, skewness, and kurtosis are used to investigate the influence on the shape of the revenue distribution under conditions of normally distributed price and yield. However, a Monte Carlo technique is used for the skewed price-normal yield case due to the unavailability of the expression for the moments of returns for such conditions. Buccola's empirical investigation of irrigated alfalfa and dryland wheat leads to the rejection of normality for alfalfa revenue, but not for wheat revenue. An interesting aspect of the wheat result in that normality is not rejected for the revenue distribution, but is rejected for the price distribution. Although the wheat result appears somewhat inconsistent with the Monte Carlo results, Buccola attributes the inconsistency to a more negative correlation coefficient between price and yield (-0.287) than for alfalfa, and a lower coefficient of variation for price than for yield.

The research reported in this paper pursues two major objectives. First, an expression for the skewness of the product of two random variables is developed to provide more insight into how the price and yield component distributions influence the resulting revenue distribution when a multinormal situation and price-yield independence are not imposed. Second, data for price, yield, and the

---

Senior authorship is equally shared and the authors are Assistant Professor of Finance, University of Kentucky, and Associate Professor of Agricultural Economics, University of Georgia, respectively.

resulting revenue are analyzed for three crops to test the hypothesis that the sample was drawn from a normal population. The results of this analysis show how composite distributions may differ from the component distributions. This provides additional information on the prevalence of nonnormal revenue distributions for crops and situations when nonnormality of revenue is likely to occur.

Statistical Framework

Complete component expressions of gross revenue variance and skewness are needed to better understand and evaluate effects of price-yield interactions in forming crop revenue distributions. The equations presented in this section differ from those presented by Buccola in that the price and yield component distributions are not restricted to normality. Moreover, expressions for expected value and variance of a product of two random variables used in agricultural economics research previous to Buccola's work for the most part have assumed normality and/or independence (e.g. Burt and Finley; Boggess et al.). Recent exceptions include Tew and Boggess; and Alexander, Musser, and Mason which follow Goodman; and Bohrnstedt and Goldberger.

The expression for the variance of gross revenue, which is not restricted to an assumption of normality or independence for the price and yield distributions, can be established from Goodman or Bohrnstedt and Goldberger as follows:

$$(1)\quad \begin{aligned} V(PQ) = {} & [E(P)]^2V(Q) + [E(Q)]^2V(P) \\ & + E\{[P-E(P)]^2[Q-E(Q)]^2\} \\ & + 2E(P)E\{[P-E(P)]\,[Q-E(Q)]^2\} \\ & + 2E(Q)E\{[P-E(P)]^2[Q-E(Q)]\} \\ & + 2E(P)E(Q)C(PQ) - [C(PQ)]^2 \end{aligned}$$

where $V(\cdot)$, $E(\cdot)$, and $C(\cdot)$ are variance, expected value, and covariance operators, respectively.

The general expression for the skewness of gross revenue can be derived in a manner similar to that used by Bohrnstedt and Goldberger in deriving the variance expression. The resulting expression for the skewness of gross revenue is:

$$(2)\quad \begin{aligned} S(PQ) = {} & S(P)[E(Q)]^3 \\ & + S(Q)[E(P)]^3 + E[(\Delta P)^3\,(\Delta Q)^3] \\ & + 3 \sum_{\substack{i=0 \\ i\neq j}}^{2} \sum_{j=0}^{2} [E(P)]^i\,[E(Q)]^j \\ & \cdot E\,[(\Delta P)^{3-i}(\Delta Q)^{3-j}] \\ & + 6E(P)E(Q)E[(\Delta P)^2(\Delta Q)^2] \\ & - 3E[(\Delta P)^2(\Delta Q)^2]C(PQ) \\ & - 3[E(Q)]^2V(P)C(PQ) \\ & - 3[E(P)]^2V(Q)C(PQ) \\ & - 6E(Q)E[(\Delta P)^2(\Delta Q)]C(PQ) \\ & - 6E(P)E[(\Delta P)(\Delta Q)^2]C(PQ) \\ & - 6E(P)E(Q)[C(PQ)]^2 + 2[C(PQ)]^3 \end{aligned}$$

where $\Delta P = P-E(P)$ and $\Delta Q = Q-E(Q)$.

The effect on gross revenue skewness of the various simplifying assumptions often used in the agricultural economics literature can be seen by examining equation (2). For example, consider the assumption of independence between crop prices and yields. This assumption causes all terms in the equation that include covariance plus some other terms to be zero. This may cause the skewness of revenue to be grossly under or over stated. Another common assumption often used in risk analysis is nonstochastic prices which causes all of the terms in the equation except the second term, $S(Q)E(P)^3$, to be equal to zero. Given these common assumptions and the conclusion from Day's research, it is easy to see why the use of analytical methods of risk analysis that depend on an assumption of normality have been criticized. However, studying equation (2) reveals that, when positive skewness exists for P and Q, a negative correlation between P and Q may have some dampening effects on revenue skewness compared to the case with a positive correlation. Consider the term $3[E(P)]^2 E[(\Delta P)(\Delta Q)^3]$ that is implicit in the summation term in (2). If P and Q are negatively correlated, then $E[(\Delta P)(\Delta Q)]$ is negative. Furthermore, because (ΔQ) is large relative to (ΔP), then this is a numerically large negative number relative to other terms. Note that as the correlation becomes more negative and the numerically largest negative deviation occurs for ΔQ, the numerically largest ΔP is more likely to occur, causing the value of the term to become even more negative. Thus, equation (2) helps explain why the wheat revenue in Buccola's study did not significantly depart from normality.

Analytical Procedure

Skew and kurtosis parameters are calculated directly from the data for price and yield as well as the resulting revenue to gain insight into the characteristics of empirical price and yield distributions and the relationship between them in forming revenue distributions. Equation (2) is not necessary for such calculations, but is useful in explaining the resulting shape of the empirical revenue distribution and conditions under which nonnormality likely occurs. Characteristics of each component distribution (price and yield) are compared to the corresponding characteristics of the resulting revenue distribution so the interaction of the components can be assessed. The correlation coefficient between price and yield is calculated and reported

Table 1. Algebraic Definitions of the Statistical Tests Used to Evaluate the Normality of a Distribution

Author	Characteristic Tested	Test Statistic
Pearson	Skewness[a]	$\sqrt{b_1} = m_3/m_2^{3/2}$
Pearson	Kurtosis	$b_2 = m_4/m_2^2$
Shapiro and Wilk	General[b] normality	$W = [\sum_{i=1}^{n/2} a_{n-i+1} (x_{n-i+1} - x_i)]^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2$

[a]The value m_i is an estimate of the ith central moment.
[b]The value (n/2) is the largest integer representing one-half the sample size, and a_{n-i+1} is a coefficient tabulated in Shapiro and Wilk, 1965.

because it helps in understanding the formation of the shape of the revenue distribution.

Several tests are available for considering the normality characteristics of a distribution. These include the standardized third and fourth moment tests used by Day (Pearson, 1928; 1930) and, the Shapiro-Wilk test of complete samples (Shapiro and Wilk, 1965). Table 1 presents test statistics for skewness, kurtosis, and general normality. Shapiro and Wilk (1964) discuss and evaluate the relative merits of each test based upon results from Monte Carlo experiments. Briefly, the Pearson standardized third and fourth moment tests are sensitive to departures from normality, although they are generally nonrobust when small samples are considered. The Shapiro-Wilk is a nonparametric test of the null hypothesis of normality, considering symmetric versus asymmetric distributions, short-tailed versus long-tailed distributions, and sample size. The Shapiro-Wilk test is both scale and origin invariate and is more robust than a combined assessment of the standardized third and fourth moments (Shapiro and Wilk 1964; 1965). Each of these tests is used in this study in describing and analyzing distributional characteristics of each variable. However, only the Shapiro-Wilk test is used to determine statistically significant departures from normality because of its more general and robust nature. Table 1 presents the algebraic definitions of the Pearson test statistics for skewness and kurtosis and the Shapiro-Wilk test statistics for general normality.

Data

Yield data for this research are provided by Day's original research and his primary sources (Grissom; Grissom and Spurgeon). Crop yields for cotton, corn, and oats from field experiments conducted by the Delta Branch of the Mississippi State Experiment Station are used to form the yield distributions. The cotton and corn experiments extend from 1921 to 1957, while the oat experiments begin in 1928 and continue through 1957. All of the experiments are fertilizer tests. Yields are available for seven different fertilizer input levels. Data for a high, intermediate, and low amount of nitrogen are chosen for this study. Although the data are dated in terms of the production and technical approach, two compelling reasons exist for their use. First, the yield data are a uniquely long series and, based on Day's analysis, reflects little evidence of nonrandomness among yields (p. 719). Hence, no detrending or other adjustments of the data are necessary. Second, Day's article is often cited in the agricultural economics risk literature and serves as a foundation reference for many studies.

Mississippi prices for the three crops are used to form the various distributions of gross revenues (Agricultural Statistics). Crop prices are deflated by the index of prices paid by producers. Subsequently, the deflated prices are detrended by fitting a cubic trend to the data, making an additional prediction of a single observation, and adding the residuals to the predicted value to form price distributions used to compute gross revenue. Detrending of prices is needed to remove effects of technology and other secular influences in an effort to provide price distribution stationarity. Although the more common linear and quadratic detrending were considered, recall that the data span the time period before, during, and after the Great Depression. Hence, a cubic trend seems plausible for these agricultural commodity prices and provides the best fit of the temporal influences. All of the trend variables in the price trend regression equations for these commodities are significant at the $\alpha = 0.01$ level and all the equations have an adjusted R^2 exceeding 0.70.

Results

Table 2 presents distributional characteristics of prices and yields for each crop and each input level considered. Yield distributions of corn and cotton exhibit positive skewness with yields tending to

Table 2. Descriptive Statistics of Real, Detrended Crop Prices and Various Yields for Cotton, Corn, and Oats in Mississippi, 1921–57[a]

Random Variable	Mean	Variance	Coefficient of Variation	Skewness ($\sqrt{b_1}$)	Kurtosis (b_2)	General Normality (W)
Cotton Price ($/lb.)	0.283	0.001	0.10	0.55	2.51	0.95
Corn Price ($/bu.)	0.924	0.048	0.24	0.75*	4.16	0.96
Oat Price ($/bu.)	0.547	0.010	0.18	−0.07	1.88	0.95
Cotton Yields (lbs./acre)						
w/high N	1719.73	295160.0	0.32	0.90*	3.59	0.93*
w/medium N	1426.16	209372.0	0.32	1.51*	6.36*	0.88*
w/o N	1022.73	167348.0	0.40	2.10*	8.98*	0.82*
Corn Yields (bu./acre)						
w/high N	47.35	230.41	0.32	0.23	2.34	0.96
w/medium N	37.63	175.82	0.35	0.19	2.04*	0.93*
w/o N	24.83	114.14	0.43	0.40	2.21	0.91
Oat Yields (bu./acre)						
w/high N	56.98	217.94	0.26	−1.40*	5.20*	0.87*
w/medium N	44.86	160.93	0.28	−0.24	3.56	0.95
w/o N	17.45	73.03	0.49	0.05	3.02	0.97

[a]Cotton and corn experiments began in 1921 while oat experiments began in 1928. All experiments ended in 1957.
*Significant at the $\alpha = 0.10$ level.

Table 3. Correlation Coefficients of Crop Yields and Real Prices in Mississippi, 1921–57

	Cotton Price	Corn Price	Oat Price
Cotton Yields			
w/high N	−0.411		
w/medium N	−0.512		
w/o N	−0.360		
Corn Yields			
w/high N		−0.301	
w/medium N		−0.351	
w/o N		−0.323	
Oat Yields			
w/high N			0.125
w/medium N			−0.091
w/o N			−0.008

become less positively skewed as fertilizer levels increase. Alternatively, oat yields tend to be negatively skewed, with higher levels of nitrogen producing increasingly negatively skewed distributions. These characteristics are consistent with Day's findings.

All distributions are tested under the null hypothesis of normality using the Shapiro-Wilk statistic as discussed. All cotton yield distributions, medium and no nitrogen corn yield distributions, and high nitrogen oat yield distributions test significantly different from normal. The skewness coefficient of the real detrended price distributions of cotton and corn indicates positive skewness. However, the oat price distributions are slightly negatively skewed. None of the price distributions significantly departs from normality.

Table 3 presents the correlation coefficients between crop yields and real detrended prices. All of these correlation coefficients are negative with the exception of the high nitrogen oat case. As expected, cotton, which is the most regionally specific crop, has the strongest correlation coefficient. Finally, Table 4 presents the descriptive statistics of the various gross revenue distributions. Tests for all of the cotton and oat gross revenue distributions indicate that normality cannot be rejected. In contrast, two out of three corn gross revenue distributions test significantly different from a normal distribution. One apparent reason normality is rejected for corn revenue is because the price distribution skewness for corn is the highest of all crops considered and the yield skewness is relatively high as well. Note that the one situation in which corn revenue does not test significantly different from normal occurs when the correlation coefficient between price and yield is most negative and when the yield skewness is least positive.

Summary and Conclusions

Agricultural economists often cite Day's result that nonnormal distributions best characterize crop yields as partial justification for using risk efficiency methods other than the expected value-variance criterion. However, the general implication of Day's results on revenue distributions has been misleading because of the assumption of constant output prices. As Grant points out, this assumption may be invalid because price and yield may be correlated.

The skewness expression developed in this study shows that the price-yield correlation is a major influence in determining the skewness of revenue. Therefore, normality for revenue may not be rejected even if the price and/or yield distributions are significantly skewed. Analysis of cotton revenue for Mississippi shows that this can be the case empirically when the correlation between price and yield is moderately negative and the relative variability of yield and price is not too high. Hence, for crops produced in their major production regions where negative correlations between prices and yields are the greatest, forces generally are at work to make revenue distributions tend toward normal. However, depending on the coefficient of variation of the yield and price, normality may not be rejected for revenue even with a weak correlation between price and yield. This was the case for the oat revenue distribution. More research for the various regions of the country and crops is needed before the prevalence of nonnormal crop revenues can be determined. Nevertheless, the assumption of nonnormal distributions of revenue in risk analysis appears to be much less serious when prices are allowed to be stochastic than when a constant price assumption is imposed.

Table 4. Descriptive Statistics for Gross Revenue for Cotton, Corn, and Oats in Mississippi, 1921–57[a]

Gross Revenue Variable	Mean ($/acre)	Variance	Coefficient of Variation	Skewness ($\sqrt{b_1}$)	Kurtosis (b_2)	General Normality (W)
Cotton						
w/high N	460.95	10698.50	0.22	0.31	1.85	0.94
w/medium N	373.83	5085.21	0.19	0.39	2.41	0.95
w/o N	262.76	3522.58	0.23	0.70*	3.60	0.96
Corn						
w/high N	42.76	246.10	0.37	0.79*	3.13	0.93*
w/medium N	32.78	155.65	0.38	0.49	2.48	0.95
w/o N	20.88	93.76	0.46	0.76*	2.57	0.91*
Oats						
w/high N	31.44	104.35	0.32	−0.59	3.19	0.96
w/medium N	23.76	66.07	0.34	−0.39	2.89	0.97
w/o N	9.53	21.96	0.49	−0.20	2.36	0.97

[a]Cotton and corn experiments began in 1921 while oat experiments began in 1928. All experiments ended in 1957.
*Significant at the $\alpha = 0.10$ level.

References

Alexander, Vickie J., Wesley N. Musser, and George Mason. "Futures Markets and Firm Decisions Under Price, Production, and Financial Uncertainty." *Southern Journal of Agricultural Economics* 18(1986):39–50.

Boggess, W. G., G. D. Lynne, J. W. Jones, and D. P. Swaney. "Risk-Return Assessment of Irrigation Decisions in Humid Regions." *Southern Journal of Agricultural Economics* 15(1983):135–144.

Bohrnstedt, George W. and Arthur S. Goldberger. "On the Exact Covariance of Products of Random Variables." *Journal of the American Statistics Association* 64(1969): 1439–1442.

Buccola, Steven T. "Testing for Nonnormality in Farm Net Returns." *American Journal of Agricultural Economics* 68(1986):334–343.

Burt, O. R., and R. M. Finley. "Statistical Analysis of Identities in Random Variables." *American Journal of Agricultural Economics* 50(1968):734–744.

Collender, Robert N. and David Zilberman. "Land Allocation Under Uncertainty for Alternative Specifications of Return Distributions." *American Journal of Agricultural Economics* 67(1985):779–786.

Day, R. H. "Distributions of Field Crop Yields." *American Journal of Agricultural Economics* 47(1965):713–741.

Gallagher, Paul. "U.S. Corn Yield Capacity and Probability: Estimation and Forecasting With Nonsymmetric Disturbances." *North Central Journal of Agricultural Economics* 8(1986):109–122.

Goodman, Leo A. "On Exact Variance of Products." *Journal of the American Statistics Association* 55(1960):708–713.

Grant, Dwight. "Theory of the Firm with Joint Price and Output Risk and a Forward Market." *American Journal of Agricultural Economics* 67(1985):630–635.

Grissom, Perrin H. "Sources and Rates of Nitrogen." Mississippi State College Agricultural Experiment Station Information Sheet #575. February 1958.

Grissom, Perrin H. and W. I. Spurgeon. "Fertility Practices for Cotton and Corn in the Yazoo-Mississippi Delta." Mississippi State College Agricultural Experiment Station Bulletin 614, April 1961.

Haldane, J. B. S. "Moments of the Distributions of Powers and Products of Normal Variates." *Biometrika* 32(1942): 226–242.

Harris, Thomas R. and Harry P. Mapp. "A Stochastic Dominance Comparison of Water-Conserving Irrigation Strategies." *American Journal of Agricultural Economics* 68(1986):298–305.

Klemme, Richard M. "A Stochastic Dominance Comparison of Reduced Tillage Systems in Corn and Soybean Production Under Risk." *American Journal of Agricultural Economics* 6(1985):550–557.

Pearson, E. S. "The Distribution of Frequency Constants in Small Sample From Nonnormal Symmetrical and Skew Populations." *Biometrika* 20A(1928):356.

———. "A Further Development of Tests for Normality." *Biometrika* 22(1930–31):237–249.

Shapiro, S. S. and M. B. Wilk. "A Comparative Study of Various Tests for Normality." Invited paper, American Statistics Association Annual Meeting, Chicago, Ill., December 1964.

———. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika* 52(1965):591–611.

Tew, Bernard V. and William G. Boggess. "Risk-Return Assessment of Irrigation Decisions in Humid Regions: An Extension." *Southern Journal of Agricultural Economics* 16(2)(1984):159–160.

U.S. Department of Agriculture. *Agricultural Statistics*, various issues. United States Gov. Printing Office, Washington, D.C.

Yassour, Joseph, David Zilberman, and Gordon C. Rausser. "Optimum Choices Among Alternative Technologies With Stochastic Yield." *American Journal of Agricultural Economics* 63(1981):718–723.