American Journal of Agricultural Economics

Editor: Michael K. Wohlgenant
Senior Associate Editors: Gerald A. Carlson
Walter N. Thurman

Department of Agricultural and Resource Economics
North Carolina State University•Box 8109•Raleigh, NC
27695-8109•(919)515-4673•(919)515-1794(FAX)

Book Review Editor: Jim Hite

Department of Agricultural and Applied Economics
Clemson University•Box 340355•Clemson, SC 29634-0355
(803)656-4700

Technical Editor: Sandra Clarke

1110 Buckeye Avenue•Ames, IA•50010-8063
(515)233-3234•~~(515)233-3101(FAX)~~

September 19, 1996

Dr. Octavio A. Ramirez
Production & Conservation Econ. Area
Centro Agronomico Tropical de
Investigacion y Ensenanza
7170 Turrialba, COSTA RICA

Dear Dr. Ramirez:

I am pleased to accept your manuscript No. 495, "Estimation and Use of a Multivariate Parametric Model for Simulating Heteroscedastic, Correlated, Non-Normal Random Variables: The Case of Corn-Belt Corn, Soya Beans and Wheat Yields," for publication in the *AJAE*, subject to responding to my final editorial instructions. At this point in time, I anticipate that your article will appear in either the February or May 1997 issue of the *Journal*.

Enclosed is an edited copy of the manuscript. Please incorporate the editorial suggestions as best you can. In addition, be careful to follow *AJAE* style as indicated in previous issues of the *Journal* and as described on the back cover of the *AJAE* under "Information on Submitting Manuscripts". As you can see, the suggested revisions are substantial. In addition to attempting to make the text flow better, I have also tried to achieve some much needed space reduction. Your manuscript, before editing, contained approximately 38 pages, which exceeds the AJAE's normal limit by eight pages. With the editing I have performed, and by eliminating the first appendix (see p. 11), I would anticipate that enough space reduction has been achieved. Please check all my editing marks carefully and please do not hesitate to ask for clarification where necessary. (My phone and fax numbers are indicated on the letterhead; my e-mail address is Michael_Wohlgenant@NCSU.Edu.)

Other administrative items to consider now include the following. Please prepare an abstract formulated according to the enclosed guidelines. You should sign and return the enclosed form for the transfer of copyright to the AAEA. Please send me professionally drawn originals of any figures, ready for photographic reproduction. Output from computer graphics packages should be made on a printer with at least 1200 dots-per-inch capacity. Please set up your table(s) in regular tabs. Do not use a table editor. We need a computer disk copy of your final manuscript (including the abstract). Please see the attached sheet for instructions. Some minor editing still may occur and you should be aware of this when reviewing the galley proofs. Please make any needed preliminary arrangements with your supporting institution to meet the page charge -- $65 per journal page on an estimated 11-12 pages for this article.

Again, I am pleased to accept your manuscript for publication, and look forward to receiving the final revision (two copies, please) in the near future.

Sincerely,

Michael K. Wohlgenant

Michael K. Wohlgenant
Editor, *AJAE*

ESTIMATION AND USE OF A MULTIVARIATE PARAMETRIC MODEL FOR SIMULATING HETEROSCEDASTIC, CORRELATED, NON-NORMAL RANDOM VARIABLES: THE CASE OF CORN-BELT CORN, SOYBEANS AND WHEAT YIELDS

by

Dr. Octavio A. Ramírez*

* Professor and Head, Production and Conservation Economics and Sociology Area, Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Turrialba, Costa Rica, Central America.

ESTIMATION AND USE OF A MULTIVARIATE PARAMETRIC MODEL FOR SIMULATING HETEROSCEDASTIC, CORRELATED, NON-NORMAL RANDOM VARIABLES: THE CASE OF CORN-BELT CORN, SOYBEANS AND WHEAT YIELDS

Octavio A. Ramírez

This study develops a multivariate non-normal density function that can accurately and separately account for skewness, kurtosis, heteroscedasticity, and the correlation among the random variables of interest. The statistical attributes of the underlying random variables and correlation processes are examined. The potential applications of this modeling tool are discussed and exemplified by analyzing and simulating Corn-Belt corn, soybeans and wheat yields. While corn and soya beans yields are found to be skewed and kurtotic, and exhibiting different variances through time; wheat yields appear normal but also heteroscedastic. A strong correlation is detected between corn and soybeans yields.

Key words: corn-belt yields, multivariate modeling and simulation, non-normality, skewness and heteroscedasticity.

ESTIMATION AND USE OF A MULTIVARIATE PARAMETRIC MODEL FOR SIMULATING HETEROSCEDASTIC, CORRELATED, NON-NORMAL RANDOM VARIABLES: THE CASE OF CORN-BELT CORN, SOYBEANS AND WHEAT YIELDS

Multivariate simulation is widely used in agricultural economics because many models and optimization problems do not have analytical solutions that can account by risk. Several flexible and comprehensive simulation methods have been proposed in the agricultural economics literature, including Clements, Mapp and Eidman and Richardson and Condra who developed a procedure to model and simulate multivariate, correlated random processes under the assumption of normality.

Around that time, Anderson stressed the importance of modeling correlation, non-normality (skewness and kurtosis) and changing variances around time/space trends/locations, because these are important characteristics of many stochastic processes in simulation models. The responses to such needs and challenges, however, have been slow and incomplete. Gallagher advances a univariate procedure to model and simulate random variables using the Gamma distribution. Because of the importance of skewness and changing variance of soybean yields over time, mainly caused by increased variation in weather conditions, he attempts to model these two characteristics of the corresponding probability distribution.

Gallagher recognizes, however, that the Gamma function assumes fixed relations between the mean, the variance and the level of skewness. In addition, these moments depend on the values taken by two parameters only. Thus, in order to model and simulate a changing variance, the corresponding mean and level of skewness also will have to be allowed to vary over time/space according to arbitrary formulae. One of the key points raised by Gallagher, nonetheless, is becoming a very important issue in simulation analysis: A large number of studies have been conducted in recent years exploring the impact of climate change and increased weather variation on agricultural production in the United States and world-wide (Adams et al; Crosson, Kaiser, and Drennen; Kaiser, H.M.; Parry et al; Rosenzweig et al).

For 1996, for example, The National Oceanic and Atmospheric Administration reports extreme or severe drought affecting crops in the Southwest (about one-third of the U.S. territory) and unusually or very moist conditions in nearly 10% of the United States. Increased weather variation has a definite impact on crop yield distributions and their simulation (Curry et al; Toure, Major, and Lindwall), possibly altering fundamental features such as their skewness and variance over time.

Taylor was the first to tackle the problem of multivariate non-normal simulation. He uses a cubic polynomial approximation of a cumulative distribution function instead of assuming a specific multivariate density for empirical analysis. This makes it impossible to assess the flexibility of this technique in terms of the potential ranges for and combinations of means, variances, coovariances, skewness, and kurtosis levels that are permitted. Another limitation of Taylor's estimation method is that it is not multivariate. Covariances are not estimated jointly with the other model parameters; therefore statistical efficiency sacrificed. Finally, Taylor's procedure can not be implemented with heteroscedasticity and specialized programs are required to estimate the model's coefficients.

Recently, the inverse hyperbolic sine transformation (IHST) proposed by Johnson has received increased attention for econometric estimation and for modeling and simulating non-normal random processes. Burbidge, Magee, and Robb empirically evaluate the usefulness of the IHST applied to the dependent variable to reduce the influence of extreme realizations on parameter estimates.

Reynolds and Shonkwiler apply the IHST in conjunction with the tobit model. Moss and Shonkwiler use the IHST to estimate yield distributions with a stochastic trend and non-normal errors. Ramírez, Moss, and Boggess explore the estimation and use of the standard form of the multivariate IHST to simulate non-normal correlated random variables. Also, Ramírez proposed an alternative specification, in which the IHST is applied to the error term instead of the dependent variable; he also derives and analyzes the statistical properties of the implicitly defined endogenous variable.

In this study, a modified inverse hyperbolic sine transformation is proposed, which extends the original parametrization to a multivariate non-normal density function that simultaneously accounts for skewness, kurtosis, heteroscedasticity, and correlation among the random variables of interest. The transformation can produce consistent maximum likelihood forecasts that are linear with respect to the matrix of explanatory variables. The proposed procedure also allows for the direct parametric testing of each of the previously mentioned distributional characteristics.

Parameter estimation and simulation are demonstrated for the IHS multivariate density function of non-normal (skewed and kurtotic), heteroscedastic, correlated random variables over time. No other multivariate density function performs through full information estimation for the modeling of all of those aspects independently and simultaneously. The theoretical attributes of this approach are shown through applying the model to Corn-Belt corn, soybeans and wheat yields. These yield distributions are believed to be non-normal because of adverse weather conditions, and because of bio-physical and technological limitations on the maximum yields attainable during any given year, causing the distributions to exhibit left-skewness.

In addition, the larger variance that usually accompanies the ever more higher average yields, and that could be compounded by the recently increased level of variability in weather conditions at agricultural areas in the United States (Adams et al.; Crosson, et al.; Kaiser, H.M. Parry et al.; Rosenzweig, et al.;), leads to believe that such yield's distributions could be heteroscedastic. Finally, the possibility of finding significant correlations between farm-level and aggregate commodity yields has also been stressed and documented by several authors, and has to be considered in any modeling and simulation effort.

The Hyperbolic Sine (HS) Random Variable

A HS random variable was defined by Ramírez as

(1) $$\sinh^{-1}(\Theta(y_t - c_t))/\Theta = v_t$$

where v_t is a normal random variable with mean μ and variance σ^2; and c_t is a "centering parameter" for the probability density function of y, which is allowed to vary across observations. In single equations, for instance, c_t can be set equal to $x_t\beta$ (Ramírez). In typical simulation models, the x_t's would become the time-trend explanatory variables.

The formulas for the expected value, variance, skewness and kurtosis coefficients of y_t can be characterized as follows (Ramírez): the expected value of y_t, given Θ, μ, σ^2, and $x_t\beta$, is equal to a constant K, plus the centering parameter $x_t\beta$; the derivative $dE[y_t]/dx_{tj}$ (t=1,.2,...,T) is equal to β_j, , so that the β_j's can be compared directly with the parameters associated with the independent variables in the standard linear econometric model; the $E[y_t]$ can take on any positive or negative value regardless of the signs or specific magnitudes of Θ, μ or σ; $x_t\beta$ provides unique and total control of the expected value of the Hyperbolic Sine random variable; and the variance of y_t is constant for every t, i.e., the model is homoscedastic.

The skewness and kurtosis coefficients are also fixed across the different realizations of the random variable y. The probability density function for y is symmetric if and only if μ=0; it will be skewed to the right if μ>0, and skewed to the left if μ<0. Therefore, a null hypothesis of symmetry, but not normality, can be specified and tested as H_o: μ=0. The model specification can accommodate random dependent variables that take on any expected values, positive or negative, in combination with any positive or negative degree of skewness.

The kurtosis coefficient for y_t can only take on positive values, so that the marginal density function of y_t is more peaked around its center than the density of a normal curve, and has fatter tails. Any positive value of the kurtosis coefficient in combination with any variance is possible, even if $\mu = 0$. As both μ and Θ approach zero, the random variable under consideration, y, becomes normally distributed.

The formerly discussed attributes as well as the flexibility of the HS model specification given by (1), in general, are illustrated in tables 1, 2 and 3. Notice that, since the expected value of the dependent variable y_t is independently controlled by the "centrality parameter" $x_t\beta$, the ranges of variance-skewness-kurtosis combinations, are achievable in conjunction with any desired expected value. Table 1 shows an ample range of possible variance-kurtosis combinations for an HS random variable when $\mu = 0$ (i.e., no skewness).

The first four rows of table 1 illustrate how a constant variance can be maintained in combination with a wide range of kurtosis coefficients by varying Θ and σ in the appropriate proportions. The last three rows show how a fixed degree of kurtosis can be preserved in association with a wide range of variances, also by changing Θ and σ in the correct proportions. Table 2 shows that the variance of y_t can be controlled with total independence of the skewness and kurtosis coefficients.

The first four rows of table 3 show how, starting from zero, an increase (decrease) in the value of μ causes an increase (decrease) in skewness, and always an increase in variance and kurtosis. Negative values of μ will cause negative (left) skewness, but will not yield negative degrees of kurtosis (thin tails). Positive values of μ will cause positive (right) skewness.

Also notice that the magnitudes of the variance and kurtosis coefficients are not affected by the signs of Θ or σ. Only the absolute value of those parameters is relevant in this case. The last four rows show how increases in the variance and kurtosis, caused by the changes in the absolute value of μ, can be offset by altering the magnitudes of Θ

and/or σ. In other words, the value of μ can be used to control the skewness coefficient of the dependent variable y_t. In summary, the flexible nature of the IHS specification makes it an attractive alternative for modeling and simulating random processes with characteristics that are not known a priori, but are expected to depart significantly from normality.

The Multivariate Hyperbolic Sine Random Process

Based on the original form of the transformation, a multivariate hyperbolic sine random process can be defined as

$$\sinh^{-1}(\Theta_i(y_{it}-x_{it}\beta_i))/\Theta_i = v_{it},\ (i=1,...,P) \qquad (2)$$

where v_{it} (i=1,...,P) is a set of normal random variables with means μ_i (i=1,...,P) and covariance matrix Σ. The standard form of the multivariate hyperbolic sine density function was derived and presented in Ramírez, Moss and Boggess:

$$f_{yt}=(2\pi)^{-P/2}|\Sigma|^{-1/2}\exp\{-.5(\mathbf{w}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{w}-\boldsymbol{\mu})\} \prod_{i=1}^{P} (1+(\Theta_i(y_{it}-x_{it}\beta_i)^2)^{-1/2} \qquad (3)$$

In equation (3), $\mathbf{w}$ is a P by 1 vector with i^{th} element $w_i = \sinh^{-1}(\Theta_i(y_{it}-x_{it}\beta_i))/\Theta_i = \ln([\Theta_i(y_{it}-x_{it}\beta_i)]+\{[\Theta_i(y_{it}-x_{it}\beta_i)]^2+1\}^{1/2})/\Theta_i$; $x_{it}\beta_i$ is the process that controls the expected value of y_{it}; and Σ is a P by P positive semi-definite matrix. This density can be made a function of several variables, each exhibiting the same characteristics of the HS random variable described in the previous section. Furthermore, such variables can be positively or negatively correlated with each other. The only restriction in degree of correlation allowed by this density function is the requirement that Σ be positive semi-definite so that the hypersurface between the null hyperplane and the function will be of measure one.

The Heteroscedastic Multivariate HS Model

It would appear that the logical and straightforward approach to introduce heteroscedasticity into the formerly defined model is to allow for the diagonal elements of the covariance matrix Σ (i.e., the σ_i's) to vary across observations (y_t's) according to any number of possible specific processes as is done when modeling under conditions of normality (Judge et al.). A quick examination of the results in Ramírez regarding the formulas for the expected value, variance, skewness and kurtosis coefficients of the y_t's, however, points out several major disadvantages of that approach:

1. Since the constant K_i in the expected value of each y_{it} is a function of σ_i, making this parameter vary across observations according to some specific process will introduce the same process into the expectation of y_{it} in an exponential manner[1].
2. The form in which the σ_i's enter the variances of the y_{it}'s, through a combination of several exponential equations, makes it very difficult, if not impossible, to incorporate and model many of the standard heteroscedastic processes currently used under normality[2]. Also, this complexity implies that isolating the effect on the variances of the y_{it}'s of any heteroscedastic processes incorporated into the model through the σ_i's would be difficult.
3. Because the σ_i's enter the formulas for the skewness and kurtosis through a combination of several exponential equations (see Ramirez), any attempt to model heteroscedas-ticity by varying the σ_i's across observations will result in a heteroskewed and heterokurtotic specification, forcing marked, and not necessarily

appropriate interrelationships, between the variances and all higher order moments of the y_{it}'s across the different observations.

The above disadvantages suggest reparametrizing the IHST to lessen or totally eliminate these limitations. With that goal in mind, the following modified transformation is proposed:

(4) $$\sinh^{-1}((r1/\sigma)(y_t - r2))/\Theta = v_t$$

where v_t is a normal random variable with mean μ and variance 1; and;

(5) $$r1 = [e^{.5\Theta^2}(e^{\Theta\mu} - e^{-\Theta\mu})]/2; \text{ and } r2 = -\sigma + \mathbf{x}_t\beta$$

For the former modified IHS transformation, it can be shown that (proof available from the author):

(6) $$E\,[y_t] = \mathbf{x}_t\beta$$

(7) $$V(y_t) = \sigma^2\,[(e^{2\Theta^2} - e^{\Theta^2})(e^{2\Theta\mu} + e^{-2\Theta\mu}) + 2(e^{\Theta^2} - 1)]/4r1^2$$

so that the parameter σ does not enter the formula for the expected value of y_t, and the variance of y_t is equal to σ^2 times a "constant" that depends only on the value of the other parameters of the model, σ and μ. It also can be demonstrated that, in this case, the third, fourth, and higher order central moments of the y_t's are not affected by the value undertaken by σ (proof available from the author). This resolves the above discussed disadvantages: σ now has as its main function controlling the variance of y in a simple linear fashion. Another change resulting from reparametrization concerns the relationship between μ and the skewness coefficient of y. This parameter can be still thought of as the chief controller of the absolute value of the third central moment of the dependent variable; as in the case of the original parametrization, the corresponding probability density function is symmetric if and only if μ=0.

As a result of introducing r1/σ into equation (4), however, determination of the direction of skewness now depends on the sign of σ and not on the sign of μ (proof available from the author)[3]. Its probability density function will be skewed to the right if $\sigma>0$, and skewed to the left if $\sigma<0$. Since the signs of Θ and σ are meaningless in this model specification, its best to "standardize" the results so that the corresponding estimates always have a positive sign. This has no real effect, other than assuring that a positive sign of σ_{12} will always correspond to a positive covariance between the two random variables under consideration.

With the reparametrization, a modified multivariate hyperbolic sine density function can be obtained from a multivariate normal density function applying the transformation technique (Mood, Graybill, and Boes); the corresponding concentrated log-likelihood function is:

$$(8) \qquad L_y = \sum_{t=1}^{T} \{ -.5\times\ln|\Sigma|+\ln(G)-.5\times(\mathrm{sumc}[\{\mathbf{dvm}*(\Sigma^{-1})\}.*\mathbf{dvm}]') \}$$

where Σ is a P by P positive semi-definite matrix with diagonal elements equal to 1, and non-diagonal elements σ_{jk}; G is a row vector with elements:

$$g_i = r1_i/(\sigma_i\Theta_i[1+\{(r1_i/\sigma_i)(y_{it}-r2_i)\}^2]^{1/2}) \ (i=1,...,P),$$

$$r1_i = [e^{.5\Theta_i^2} \ (e^{\Theta_i\mu_i} - e^{-\Theta_i\mu_i})]/2 \text{ and } r2_{it} = -\sigma_i + \mathbf{x}_{it}\beta_i,$$

if y_i is not normally distributed and $g_i = \sigma_i^{-1}$ if y_i is normally distributed.

Finally, **dvm** is a row vector with elements dvm_i (i=1,2,...,P), where $dvm_i = (\ln(r3_i+(1+(r3_i^2))^{1/2})/\Theta_i)-\mu_i$ and $r3_i = \{(r1_i/\sigma_i)(y_{it}-r2_i)\}$ if y_i is not normally distributed and $dvm_i = (y_{it}-x_{iit}\beta_i)/\sigma_i$ if y_i is normally distributed. The operator sumc(x) indicates to take the sum of the (P) columns of x; the operator * indicates a matrix multiplication; and .* indicates an element by element matrix multiplication.

It is important to emphasize that, beyond the differences previously discussed, the two model specifications are fundamentally equivalent as long as the σ_{it}'s (i=1,...,P) are constant for all t's. Under the former circumstance, the maximum value reached by the two corresponding likelihood functions will always be identical for any given data set. This is evidence that essentially the same basic model is being fitted to the underlying empirical probability distribution defined by the data.

The Case of Corn-Belt Yields: Model Specification and Estimation

The Corn-Belt corn, soya-beans and wheat yields (1950-1989) are analyzed in this study, with regard to the characteristics of their probability distribution functions by specific years, including expected value, variance, skewness, kurtosis and degrees of correlation. Utilizing the model presented above, the variations in expected yields and variances over time are investigated.

Linear time-trending processes are assumed for the expected values of the y_i's ($E[y_{it}]=\beta_{i0}+\beta_{i1}T$; t=49,...,89; T=1,...,50; i=1,2,3). The possibility of changes in variances of each of the y_i's over four different decades considered in the study also are explored ($\sigma_{it}=\sigma_{i1}$ if t=49,50...,59; $\sigma_{it}=\sigma_{i2}$ if t=60,61...,69; $\sigma_{it}=\sigma_{i3}$ if t=70,71...,79; and $\sigma_{it}=\sigma_{i4}$ if t=80,81...,89; for every i=1,2,3). The parameters of the model are estimated by maximizing L_y using the GAUSS 2.01 matrix algebra language, beginning with a first-stage limited information (i.e. variable by variable) approach by setting P=1. The results of this estimation process for the three variables under consideration are presented in tables 4 to 6.

In the case of corn yields, notice that allowing for kurtosis by permitting $\Theta_1 \neq 0$ significantly improves the model specification; the likelihood ratio test MLRT{A-B}=14.260$\approx \chi_1^2$, indicating rejection of the null hypothesis of normality at the 1% level significance. When skewness is also allowed, $\mu_1 \neq 0$, the model specification is significantly improved as well (MLRT{B-C}=9.288$\approx \chi_1^2$, indicating rejection of the null hypothesis of symmetry at the 1% level). Since σ_A is negative, the density function for y_1 is left-skewed.

Heteroscedasticity is tested for by using a standard nesting procedure that allows the possibility of up to four distinct variances, one in every decade. First, a significant difference is detected between the variance of y_1 from 1950 to 1969 and its variance from 1970 to 1989 (MLRT{C-D}=7.756$\approx \chi_1^2$; thus $H_o:\sigma_B=0$ can be rejected at the 1% level). Allowing for change in the variances within any of the two time-periods yields a substantially improved model specification as well (MLRT{D-E}=6.984 $\approx \chi_1^2$, so $H_o:\sigma_C=0$ can be rejected at the 1% level; and MLRT{D-F}=6.796$\approx \chi_1^2$, so $H_o:\sigma_C=0$ can be rejected). Analyzing the fully parametrized model, however, it can be concluded that the variance of corn yields in the period between 1960 and 1969 is very similar to the variance observed between 1980 and 1989; incorporation of that restriction into the model can be statistically justified (MLRT{G-H} =0.010$\approx \chi_1^2$, thus $H_o:\sigma_B=\sigma_D$ in G cannot be rejected even at the 20% level of significance). Therefore, the final specification (H) allows for skewness, kurtosis, and differential variances for 1950-1959, 1960-1969/1980-1989, and 1970-1979.

The results for the case of soya-beans yields are presented in table 5. Notice that, as before, simply allowing for kurtosis by permitting $\Theta_2 \neq 0$ significantly improves the model specification according to the likelihood ratio test (MLRT{A-B}=7.380$\approx\chi_1^2$, implying rejection of the null hypothesis of normality at the 1% level). When skewness is also allowed, $\mu_2 \neq 0$, the model specification is substantially improved again (MLRT{B-C}=12.838$\approx\chi_1^2$, implying rejection of the null hypothesis of symmetry at the 1% level). Since the sign of σ_A is negative, the density function for y_2 is also left-skewed, as in the case of corn yields.

When heteroscedasticity is permitted, a significant difference is detected between the variance of y_2 from 1950 to 1969, and from 1970 to 1989 (MLRT{C-D}=2.230 $\approx\chi_1^2$; thus H_o:σ_B=0 can be rejected at the 20% level of significance). Allowing for further changes in the variance within the two time periods, however, does not improve the model specification significantly.

The results for wheat yields are presented in table 6. In this case the parameter estimates for Θ_3 and μ_3 steadily approach zero[4] before the estimation algorithm collapses, indicating that the random variable y_3 exhibits a normal distribution. In regard to heteroscedasticity, a model (F) with a base-line variance, σ_A, throughout, and an additional parameter, σ_B, that is allowed to increase or decrease the variance between 1960 and 1969 only, appears to be the best. It is statistically superior to the homoscedastic formulation (A), with a MLRT equal to 3.604 indicating rejection of H_o:σ_B=0 at the 10% significance level.

In light of these results of the first-stage estimation, a full information maximum likelihood estimation effort was attempted by setting P=3 in equation (7). The results of this final step of the estimation process for the three crops under analysis are presented in table 7. Notice that all of the parameter values, and most of the corresponding asymptotic t-statistics, are reasonably close to those estimated when using the single-variable approach. It is interesting to note that the estimate of the parameter that governs the covariance between corn and soya-beans yields appears different from zero at the 1% level of statistical significance. In the cases of the parameters that govern the covariances between corn and wheat, and between soya-beans and wheat yields, however, the very low t-values do not allow us to conclude that covariances are significantly different from zero, even at the 20% level.

This means that the parameters of the multivariate model must be estimated again, subject to the restrictions $\sigma_{13}=\sigma_{23}=0$. The results of this estimation, presented in table 7, show that the estimate of σ_{12}, and its asymptotic t-value, remain fairly stable; the maximum value of the concentrated log-likelihood function in this case is only 0.20077 lower, yielding an insignificant likelihood ratio test statistic of MLRT=0.40015$\approx\chi^2_1$. Observe, however, that further restricting σ_{12}=0 reduces the concentrated log-likelihood function to -216.106 (which, as expected, is equal to the sum of the maximum values reached by the individual functions in the limited-information approach). This yields a MLRT= 16.855$\approx\chi^2_1$, which indicates a highly significant covariance relation between corn and soya-bean yields.

Yield Simulation

The above results can be used to simulate the estimated distributions of corn, soya-beans and wheat yields in the Corn Belt during any given year. This simulation assumes a time-trend process as well as heteroscedasticity, skewness and kurtosis, and takes into account correlation between the variables of interest, where appropriate.

The first step in this procedure is to generate a matrix of correlated normal random variables with a mean vector μ ($\mu=[\mu_1,\mu_2,\mu_3]$) and covariance matrix Σ as previously defined. For this purpose, an N by 3 matrix of standard normal random variables is multiplied by the Cholesky decomposition of Σ (Clements, Mapp, and Eidman), and the corresponding μ_i is added to the appropriate column vector of the resulting N by 3 matrix (M). Then, the yield distributions, y_{it}'s, are simulated separately for each crop and any given time period t by applying the inverse function of the transformation defined in equation (4):

$$(9) \qquad y_{it} = (\sigma_{it}/r_{1i})(e^{\Theta_i v_i} - e^{-\Theta_i v_i})/2 + r_{2it}$$

where r_{1i} and r_{2it} are as specified in (7), and v_i is the i^{th} column vector of M. Utilizing the previously described procedure, 2,500 yields were simulated for each crop, and for the years of 1955, 1975, and 1995. In the case of 1995, $x_{it}\beta_i$ was forecasted by setting t=95 (i.e. T=46), and the estimate of σ_i for the last decade (1980-1989) was used. The graphs of the estimated yield distributions for the non-normal cases, corn and soybeans, during those three different time periods, are presented in figures 1 through 6. Notice that in all cases, the simulation outcomes are compatible with the conclusions in the previous section about the signs and values of the parameter estimates and the related theoretical results. Although it is not easy to visually detect kurtosis in the graphs, the estimated corn yield distributions are clearly and significantly skewed to the left. Furthermore, the degrees of skewness and kurtosis (i.e., the general shapes of the distributions) are maintained through time, as anticipated.

Observe also that the variances are different in the three time periods under consideration. Simulated 1955 corn yields ranged from 38 to 63 bushels per acre, while 1975 yields ranged from 49 to 117 bushels per acre. This is approximately a 172% increase in variability. 1995 yields ranged from 108 to 153 bushels per acre, implying about an 80% increase in variability from 1955. Notice that the absolute value of σ for 1955 is $\sigma_A = 7.4409$, for 1975 it is $\sigma_A+\sigma_C = 20.3435$ (a 173.40% increase) and for 1995 it equals $\sigma_A+\sigma_B = 13.6085$ (an 82.90% increment from σ_A). Since σ is proportional to the standard deviation of the random dependent variable by design, it can be stated that the heteroscedasticity-governing mechanism built into the model performs as expected.

Very similar results are found for soybeans. The estimated yield distributions are clearly and significantly skewed to the left. Furthermore, the degrees of skewness and kurtosis (i.e., the general shapes of the distributions) are maintained through time, as well. Simulated 1955 soya beans yields range from 16 to 26 bushels per acre; while 1975 yields extend from 22 to 36 bushels and 1995 yields range from 31 to 45 bushels per acre for a 140% increase in variability with respect to 1955. Notice that, in this case, the value of σ for 1955 is $\sigma_A = 3.9393$, while for 1975 and 1995 it is of $\sigma_A+\sigma_B = 5.6418$ (a 143.22% increase). The heteroscedasticity-governing mechanism built into this model specification again performs as expected.

The finding of significant degrees of heteroscedasticity in all of the three variables reinforces Anderson's (1974) proposition about the importance of modeling and simulation methods that can account for changing variances over time and/or space. When homoscedasticity is assumed, the variance of corn yields during the last decade (which would be the appropriate one to use for simulation) is underestimated by 15%, the variance of soya-beans yields is underestimated by 20%, and the variance of wheat yields is underestimated by 10%.

This happens because the homoscedastic models (C in table 4, C in table 5, and A in table 6, respectively) estimate average variances for the 40-year period under analysis, while the heteroscedastic formulations (H in table 4, D in table 5, and F in table 6, respectively) capture an increased variance that usually accompanies higher average yields. This phenomena of increased variance is apparently compounded by the recently increased level of weather variability at agricultural areas in the United States.

The substantial left skewness that characterizes the probability distributions of corn and soya beens yields is another interesting result. A logical interpretation of this phenomena is the following: The right hand side of the distributions terminate in a relatively abrupt manner, suggesting that it is not possible to obtain yields that are much higher than the more frequently observed production levels. Most yields must be constrained between a "absolute maximum" determined by the current technology, and a "likely minimum" also located relatively close to the more frequently observed production levels. There always is, however, a low but feasible probability that, because of pest incidence or quite unfavorable weather conditions, a yield significantly below the likely minimum is observed in any given year.

The formerly described phenomena can be detected in the actual data set for the cases of corn and soya beans. Corn yields in 1970, for example, were 77.54 bushels per acre, while the production levels in the other years from 1967 to 1973 oscillated between 89.59 and 107.73 with a mean of 97.56. Also in 1974, corn yields were 76.86 bushels per acre; production levels from 1971 to 1977 fluctuated between 95.76 and 107.73 with a mean of 100.09. Very similar situations occurred during 1981-85 and 1986-89 (see the appendix).

The soybeans yields distribution is perceptibly less skewed than corn yields, but a similar phenomena still can be detected in the actual data set. In 1974, soya beans yields were 25.20 bushels per acre, while the production levels in the other years from 1971 to 1977 fluctuated between 30.36 and 35.86 with a mean of 32.27. In 1988, soya beans yields were 28.13 bushels per acre, while yields from 1985 to 1989 oscillated between 36.54 and 39.67 with a mean of 38.41. Similar situations occurred during 1951-55 and 1962-66.

It is important to notice the coincidence in several unusually low production years for both crops (1964, 1974, 1983, 1988), corn and soya beans. This strengthens the hypothesis of adverse weather conditions partially explaining the left skewness of their yields' probability distribution functions; it also supports the finding of a strong and significant covariance between the production levels of those two crops in any given year. On the other hand, the extremely low corn yields of 1970 were caused by a wide-spread attack of a fungal disease caused by <u>Helminthosporium</u> <u>maydis</u> (race T).

In the case of wheat, no unusually low yields are observed through time, and the related covariation with the corn and/or soya beans yields cannot be seen when examining the data. This is quite consistent with the theoretical results presented and discussed in the previous section.

Finally, it is important to ascertain whether the covariance modeling mechanism built into the model was able to accurately capture the interrelationship between the two random variables in the actual data set, so that it could be transfered in such a way to be correctly reflected in the simulation outcomes. To explore this issue, the standard formula to compute the covariance between any two random variables was used:

$$(10) \qquad cov(y_1,y_2) = \sum_{t=1}^{N} \{[y_{1t}-M(y_1)]x[y_{2t}-M(y_2)]/N; \; M(y_i) = \sum_{t=1}^{N} (y_{it})/N;$$

Applying equation (9), the covariance between the actual detrended corn and soya beans yields is 24.6664[5]. Using (9), the covariance between the 7,500 simulated yields equals 22.5995, indicating that the correlation-governing mechanism built into the model performed reasonably well.

Concluding Comments

The inverse hyperbolic sine transformation is successfully reparametrized to develop and formally define a multivariate non-normal density function that can accurately and separately account for skewness, kurtosis, heteroscedasticity, and correlation. The parameter estimation process, and the use of the newly designed multivariate density function for simulating non-normal (skewed and kurtotic), heteroscedastic, correlated random variables over time is effectively demonstrated with the case of Corn-Belt corn, soya beans and wheat yields. No other multivariate density function has been proposed and shown to perform in the modeling of all of those aspects simultaneously.

The theoretical attributes of this modeling tool are discussed and exemplified by analyzing and simulating Corn-Belt corn, soya beans and wheat yields. While corn and soya beans yields are found to be skewed and kurtotic, and exhibiting different variances through time; wheat yields appear normal but also heteroscedastic. A strong positive correlation is detected between corn and soya beans yields. Hypotheses are advanced to explain the resulting shapes of the probability distribution functions that are markedly asymmetric, and the significant correlation detected between corn and soya beans yields.

The simulation outcomes are observed to be highly consistent with the attributes of the actual data that can be assessed through visual inspection or with the help of simple numerical indexes such as the correlation coefficient. In addition to the clear advantages that this modeling tool exhibits when applied to simulation analysis; other distinctly important uses include the estimation of systems "seemingly unrelated regressions" (please see Judge et al.) in general; when the random dependent variables of interest are suspected to be non-normal, heteroscedastic, and mutually correlated. The former with the intention of improving statistical efficiency with respect to a model specification based on the assumption of normality.

The improved statistical efficiency and accuracy that could be achieved through the use of this modeling and simulation tool can have important positive consequences in key areas of application such as the analysis of agricultural policies, where the proper simulation of variables such as commodity yields and prices is a critical condition.

References

Adams, R.M., C. Rosenzweig, R.M. Peart, J.T. Ritchie, B.A. McCarl, J.D. Glyer, R.B. Curry, L.W. Jones, K.J. Boote, and L.H. Allen. "Global Climate Change and US Agriculture." *Nature*. 345(May 1990):6272, 219-224.

Anderson, J.R. "Simulation: Methodology and Application in Agricultural Economics." *Review of Marketing and Agricultural Economics*. 42(March 1974):3-55.

Burbidge, J.B., L. Magee, and A.L. Robb. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association*. 83(March 1988):123-127.

Clements, A.M., H.P Mapp Jr, and V.R. Eidman. "A Procedure for Correlating Events in Farm Firm Simulation Models." Oklahoma State University Agricultural Experiment Station Bulletin T-131, 1971.

Crosson, P., H.M. Kaiser, H.M. Kaiser (ed.), and T.E. Drennen. "Impacts of Climate Change on the Agriculture and Economy of the Missouri, Iowa, Nebraska, and Kansas (MINK) Region." Agricultural Dimensions of Global Climate Change. St. Lucie Press, Delray Beach, USA, 1993, pp. 117-135.

Curry, R.B., R.M. Peart, J.W. Jones, K.J. Boote, and L.H. Allen. "Simulation as a Tool for Analyzing Crop Response to Climate Change." Dep. of Agric. Engineering, Univ. of Florida, Gainesville, FL, USA, 1988.

Gallagher, P. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances." *American Journal of Agricultural Economics*. 71(November 1987):796-803.

Johnson, N.L. "Systems of Frequency Curves Generated by Methods of Translation." *Biometrika*. 36(1949):149-176.

Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lutkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. New York: John Wiley & Sons, Inc., 1985, pp. 419-464.

Kaiser, H.M. "Climate Change and Agriculture." *Northeastern Journal of Agricultural and Resource Economics*. 20(October 1991):151-163.

Mood, A.M., F.A. Graybill and D.C. Boes. *Introduction to the Theory of Statistics*. New York: McGraw-Hill, 1974, pp. 198-212.

Moss, C.B. and J.S. Shonkwiler. "Estimating Yield Distributions with a Stochastic Trend and Non-Normal Errors." *American Journal of Agricultural Economics*. 75(November 1993):1056-1062.

Parry, M., J.C. Zhang, J. Jager (ed.), and H.L. Ferguson (ed.). "The Potential Effect of Climate Change on Agriculture." *Climate Change: Science, Impacts and Policy*. Proceedings of the Second World Climate Conference. Cambridge: Cambridge University Press, UK, 1991, pp. 279-289.

Ramirez, O.A., C.B. Moss, and W.G. Boggess. "Estimation and Use of the Inverse Hyperbolic Sine Transformation to Model Non-Normal Correlated Random Variables." *Journal of Applied Statistics*. 21(December 1994):289-304.

Ramírez, O.A. "Linear Regression with Nonnormal Error Terms Using the Inverse Hyperbolic Sine Transformation". Economics of Production and Conservation Area, CATIE, Costa Rica, 1996.

Reynolds, A. and J.S. Shonkwiler. "Testing and Correcting for Distributional Misspecifications in the Tobit Model: An Application of the Information Matrix Test". *Empirical Economics.* 16(September 1991):313-324.

Richardson, J.W. and G.D. Condra. "A General Procedure for Correlating Events in Simulation Models." *Department of Agricultural Economics*, Texas A&M University, 1978.

Rosenzweig, C., R.M. Adams, J.M. Reilly, P.R. Crosson, and J.E. Schefter. "Climate Change, American Agriculture, and Natural Resources." *American Journal of Agricultural Economics.* 71(December 1989):1265-1271; 1272-1279; 1280-1282; 1283-1285.

Taylor, C.R. "Two Practical Procedures for Estimating Multivariate Nonnormal Probability Density Functions." *American Journal of Agricultural Economics.* 72(February 1990):210-217.

Toure, A., D.J. Major, and C.W. Lindwall. "Sensitivity of Four Wheat Simulation Models to Climate Change." *Canadian Journal of Plant Science.* 75(January 1995):69-74.

Footnotes

1 $$E[y_{it}] = K_i + x_{it}\beta_i = [e^{.5\Theta_i^2\sigma_i^2}(e^{\Theta_i\mu_i} - e^{-\Theta_i\mu_i})]/2\Theta_i + x_{it}\beta_i$$

2 $$V(y_{it}) = [(e^{2\Theta_i^2\sigma_i^2} - e^{\Theta_i^2\sigma_i^2})(e^{2\Theta\mu} + e^{-2\Theta\mu}) + 2(e^{\Theta_i^2\sigma_i^2} - 1)]/4\Theta_i^2$$

3 Notice, however, that the sign of σ affects neither the sign nor the absolute value of the variance of y_t.

4 The maximum value of L_y approaches the one of a concentrated likelihood function set up under the assumption of normality, as well.

5 In order to detrend the random variables to the arbitrary year of 1950, the formulas $y_{1t}-\beta_{11}T$ and $y_{2t}-\beta_{21}T$ (T=0,...,49) were used. Notice, however, that the value of $cov(y_1,y_2)$ in (10) will be the same, independently of the specific year to which the random variables are detrended.

Table 1. Some possible variance-kurtosis combinations for an IHS random (original specification) variable when $\mu = 0$ (no skewness).

Θ	σ	Θ/σ	Variance	Skewness	Kurtosis
±0.0000	±1.7873	0.0000	3.1945	0.0000	0.0000
±0.7689	±1.1500	0.6686	3.1945	0.0000	14.6905
±1.0000	±1.0000	1.0000	3.1945	0.0000	33.1881
±1.3109	±0.8500	1.5422	3.1945	0.0000	82.2300
±4.0000	±0.2500	16.0000	0.1997	0.0000	33.1881
±1.0000	±1.0000	1.0000	3.1945	0.0000	33.1881
±0.2500	±4.0000	0.0625	51.1124	0.0000	33.1881

Table 2. Control of the variance of IHS random variable (original specification) with total independence of the skewness and kurtosis coefficients.

Θ	μ	σ	Variance	Skewness	Kurtosis
0.5000	±0.5000	1.5000	4.499	±0.978	7.809
1.0000	±0.2500	0.7500	1.125	±0.978	7.809
1.5000	±0.1667	0.5000	0.500	±0.978	7.809
2.0000	±0.1250	0.3750	0.281	±0.978	7.809
2.5000	±0.1000	0.3000	0.180	±0.978	7.809
3.0000	±0.0833	0.2500	0.125	±0.978	7.809

Note: Positive values of μ are associated with positive skewness coefficients and vice versa.

Table 3. Effect of the value of μ on the variance, skewness and kurtosis coefficients of an IHS random variable (original specification).

Θ	μ	σ	Variance	Skewness	Kurtosis
±1.0000	0.0000	±1.0000	3.1945	0.0000	33.1881
±1.0000	±0.2500	±1.0000	3.4926	±2.1077	41.6467
±1.0000	±0.5000	±1.0000	4.4628	±3.6279	60.0362
±1.0000	±0.7500	±1.0000	6.3529	±4.7602	77.6816
±1.0000	0.0000	±1.0000	3.1945	0.0000	33.1881
±1.0000	±0.2500	±0.9750	3.1092	±1.9437	34.6043
±1.0000	±0.5000	±0.9250	3.1323	±2.9588	35.0098
±1.1250	±0.7500	±0.7750	3.1391	±3.4785	34.5114

Note: Positive values of μ are associated with positive skewness coefficients and vice-versa.

Table 4. Stage-one, limited information maximum likelihood parameter estimates of the probability density functions for corn yields (n=40).

	Corn Yields									
	Θ_1	μ_1	β_0	β_1	σ_A	σ_B	σ_C	σ_D	MCLLF	MLRT
Parameter/A	-	-	48.3972	1.8812	11.6281	-	-	-	-118.137	-
T-Values	-	-	12.9705	11.9396	8.9655	-	-	-	-	-
Parameter/B	1.1543	-	29.3396	2.2739	14.6840	-	-	-	-111.007	14.260
T-Values	3.4854	-	4.4596	20.4170	2.5195	-	-	-	-	-
Parameter/C	0.9424	1.2359	39.8477	2.2969	-9.9241	-	-	-	-106.363	9.288
T-Values	3.3955	2.0211	13.6151	22.6160	-2.9099	-	-	-	-	-
Parameter/D	0.6201	11.4867	42.8208	2.2446	-10.8515	-7.6257	-	-	-102.485	7.756
T-Values	2.6963	1.3223	16.1158	18.5079	-4.2437	-1.6731	-	-	-	-
Parameter/E	1.4884	11.9899	33.4792	2.3774	-15.3269	-4.6017	8.6551	-	-98.993	6.984
T-Values	6.8164	3.7915	15.8054	107.4299	-7.4706	-9.4935	16.7785	-	-	-
Parameter/F	0.6770	15.0285	44.4468	2.0378	-6.4509	-7.5870	-14.8038	-	-99.087	6.796
T-Values	3.3472	4.7524	36.9485	19.2396	-4.4087	-3.9580	-6.1653	-	-	-
Parameter/G	1.1302	11.2047	42.1220	2.1081	-7.6609	-4.8317	-10.7722	-4.5791	-97.697	9.576
T-Values	3.6948	3.5432	23.2078	29.5521	-4.4795	-3.9489	-6.4696	-2.2114	-	-
Parameter/H	1.1119	15.3169	42.3710	2.0962	-7.4622	-5.0493	-11.1096	-	-97.702	0.010
T-Values	3.6481	4.8436	24.6365	31.2363	-4.4914	-4.0509	-7.4829	-	-	-

Notes: A is the model specification assuming normality; B introduces kurtosis and C kurtosis and skewness. In A, B and C, $\sigma=\sigma_A$ from 1950 to 1989; in D $\sigma=\sigma_A$ from 1950 to 1969 and $\sigma=\sigma_A+\sigma_B$ from 1970 to 1989; in E $\sigma=\sigma_A$ from 1950 to 1969, $\sigma=\sigma_A+\sigma_B$ from 1970 to 1979 and $\sigma=\sigma_A+\sigma_B+\sigma_C$ from 1980 to 1989; in F $\sigma=\sigma_A$ from 1950 to 1959, $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969 and $\sigma=\sigma_A+\sigma_C$ from 1970 to 1989; in G $\sigma=\sigma_A$ from 1950 to 1959, $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969, $\sigma=\sigma_A+\sigma_C$ from 1970 to 1979 and $\sigma=\sigma_A+\sigma_D$ from 1980 to 1989; in H $\sigma=\sigma_A$ from 1950 to 1959, $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969 and from 1980 to 1989, and $\sigma=\sigma_A+\sigma_C$ from 1970 to 1979. MCLLF refers to the maximum value attained by the concentrated log-likelihood function, and MLRT refers to the corresponding likelihood ratio test statistic.

Table 5. Stage-one, limited information maximum likelihood parameter estimates of the probability density functions for soya-beans yields (n=40).

Soya-Beans Yields										
	Θ_2	μ_2	β_0	β_1	σ_A	σ_B	σ_C	σ_D	MCLLF	MLRT
Parameter/A	-	-	21.4083	0.3897	2.6137	-	-	-	-58.431	-
T-Values	-	-	25.2296	10.8703	8.6594	-	-	-	-	-
Parameter/B	1.0886	-	17.6697	0.4477	3.3989	-	-	-	-54.741	7.380
T-Values	3.1935	-	12.8850	15.9706	2.6830	-	-	-	-	-
Parameter/C	0.7489	8.9701	19.6966	0.4714	-3.5881	-	-	-	-48.322	12.838
T-Values	3.4485	28.1555	32.8760	31.9518	-6.0680	-	-	-	-	-
Parameter/D	0.6851	10.6400	20.4619	0.4407	-3.0167	-1.1820	-	-	-47.207	2.230
T-Values	3.3928	72.5842	33.4188	19.7591	-4.8265	-1.4992	-	-	-	-
Parameter/E	0.6767	10.6401	20.5398	0.4382	-2.9784	-0.9788	-0.5589	-	-46.9788	0.456
T-Values	3.0961	72.9511	33.0938	18.6404	-4.5950	-1.3023	-0.5955	-	-	-
Parameter/F	0.6905	10.6401	20.4362	0.4419	-3.0326	0.0336	-1.1305	-	-47.207	0.000
T-Values	1.2917	48.7915	7.0043	6.6756	-1.9856	0.0391	-0.5775	-	-	-

Notes: A is the model specification assuming normality; B introduces kurtosis and C kurtosis and skewness. In A, B and C, $\sigma=\sigma_A$ from 1950 to 1989; in D $\sigma=\sigma_A$ from 1950 to 1969 and $\sigma=\sigma_A+\sigma_B$ from 1970 to 1989; in E $\sigma=\sigma_A$ from 1950 to 1969, $\sigma=\sigma_A+\sigma_B$ from 1970 to 1979 and $\sigma=\sigma_A+\sigma_B+\sigma_C$ from 1980 to 1989; in F $\sigma=\sigma_A$ from 1950 to 1959, $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969 and $\sigma=\sigma_A+\sigma_C$ from 1970 to 1989. MCLLF refers to the maximum value attained by the concentrated log-likelihood function, and MLRT refers to the corresponding likelihood ratio test statistic.

Table 6. Stage-one, limited information maximum likelihood parameter estimates of the probability density functions for wheat yields (n=40).

Wheat Yields										
	Θ_3	μ_3	β_0	β_1	σ_A	σ_B	σ_C	σ_D	MCLLF	MLRT
Parameter/A	-	-	22.6805	0.6831	3.7621	-	-	-	-72.999	-
T-Values	-	-	18.7079	13.2307	8.9527	-	-	-	-	-
Parameter/B	0.0000	0.0000	22.6805	0.6831	3.7621	-	-	-	-72.999	0
T-Values	-	-	18.7079	13.2307	8.9527	-	-	-	-	-
Parameter/C	-	-	22.6891	0.6884	3.2116	1.0342	-	-	-72.254	1.490
T-Values	-	-	20.9473	12.5420	6.2743	1.2041	-	-	-	-
Parameter/D	-	-	22.6634	0.6913	3.2067	1.4623	-0.8777	-	-72.042	0.424
T-Values	-	-	20.8198	14.0799	6.2098	1.2589	-0.6288	-	-	-
Parameter/E	-	-	23.4658	0.6669	3.9644	-1.6638	0.3112	-	-71.163	2.182
T-Values	-	-	19.0793	11.5554	4.1452	-1.4279	0.2612	-	-	-
Parameter/F	-	-	23.5696	0.6623	4.1820	-1.8947	-	-	-71.197	3.604
T-Values	-	-	19.3110	12.0492	7.6093	-2.3026	-	-	-	-

Notes: A is the model specification assuming normality; B attempts to introduce kurtosis, but such model specification degenerates back to normality in this case. In A and B, $\sigma=\sigma_A$ from 1950 to 1989; in C $\sigma=\sigma_A$ from 1950 to 1969 and $\sigma=\sigma_A+\sigma_B$ from 1970 to 1989; in D $\sigma=\sigma_A$ from 1950 to 1969, $\sigma=\sigma_A+\sigma_B$ from 1970 to 1979 and $\sigma=\sigma_A+\sigma_B+\sigma_C$ from 1980 to 1989; in E $\sigma=\sigma_A$ from 1950 to 1959, $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969 and $\sigma=\sigma_A+\sigma_C$ from 1970 to 1989; in F $\sigma=\sigma_A$ from 1950 to 1959 and from 1970 to 1989, and $\sigma=\sigma_A+\sigma_B$ from 1960 to 1969. MCLLF refers to the maximum value attained by the concentrated log-likelihood function, and MLRT refers to the corresponding likelihood ratio test statistic. In F the MLRT is calculated with respect to the MCLLF in C.

Table 7. Stage two, full information maximum likelihood parameter estimates and associated statistics. Corn soya beans and wheat yields (n=120).

FULL INFORMATION - ALL COVARIANCES				RESTRICTED FULL INFORMATION - $\sigma_{13}=\sigma_{23}=0$			
Parameter	Estimate	Std. Err.	T-Value	Parameter	Estimate	Std. Err.	T-Value
CORN							
Θ_1	0.82961	0.31006	2.67564	Θ_1	0.82788	0.30991	2.67133
μ_1	15.32000	0.06942	220.69863	μ_1	15.31690	0.06918	221.41111
β_0	43.13123	1.86214	23.16213	β_0	43.08897	1.89386	22.75199
β_1	2.09278	0.07759	26.97211	β_1	2.09456	0.07946	26.36063
σ_A	-7.40151	1.83675	-4.02968	σ_A	-7.44086	1.86552	-3.98863
σ_B	-6.19696	3.05608	-2.02775	σ_B	-6.16777	3.13224	-1.96913
σ_C	-12.93097	3.93522	-3.28596	σ_C	-12.90267	3.91644	-3.29449
SOYA BEANS							
Θ_2	0.49481	0.11042	4.48126	Θ_2	0.49379	0.17679	2.79309
μ_2	0.64517	0.57388	18.54950	μ_2	10.64357	0.13555	78.52158
β_0	20.65334	0.57772	35.74963	β_0	20.63419	0.61352	33.63269
β_1	0.42569	0.02299	18.51707	β_1	0.42630	0.02653	16.06580
σ_A	-3.94085	0.83655	-4.71087	σ_A	-3.93928	1.18459	-3.32543
σ_B	-1.67136	0.80896	-2.06606	σ_B	-1.70250	1.12396	-1.51474
WHEAT							
β_0	23.49857	1.22019	19.25820	β_0	23.56959	0.35598	66.21072
β_1	0.66497	0.05518	12.05100	β_1	0.66230	0.03246	20.40104
σ_A	4.17323	0.55217	7.55786	σ_A	4.18196	0.50566	8.27037
σ_B	-1.87030	0.82184	-2.27575	σ_B	-1.89466	0.66206	-2.86178
COVARIANCES							
σ_{12}	0.63189	0.10621	5.94969	σ_{12}	0.63124	0.10972	5.75293
σ_{13}	-0.06059	0.16608	-0.36484	σ_{13}	--	--	--
σ_{23}	-0.10184	0.15969	-0.63772	σ_{23}	--	--	--
MAXIMUM VALUE OF LIKELIHOOD FUNTION = -207.477				MAXIMUM VALUE OF LIKELIHOOD FUNTION = -207.678			

Figure 1

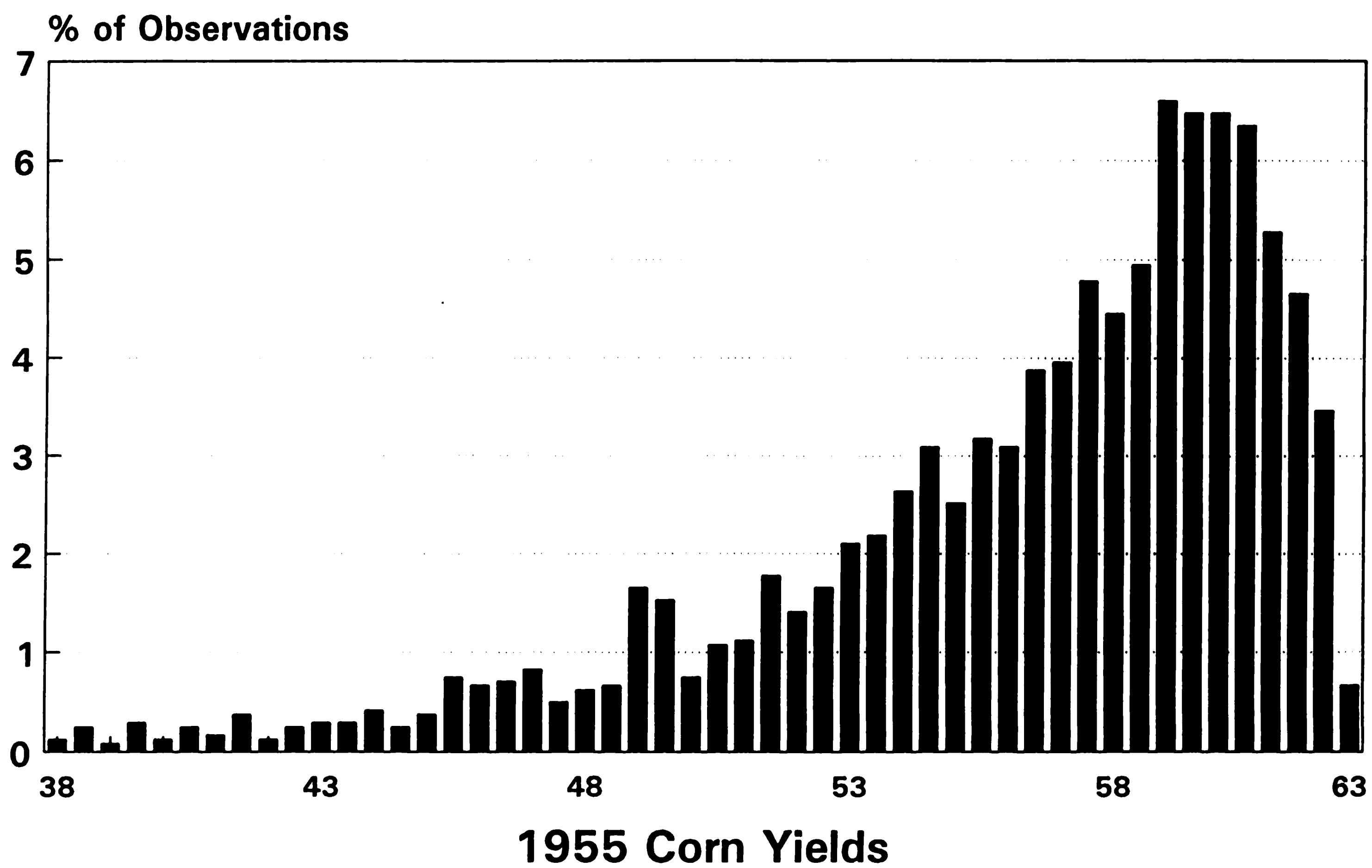
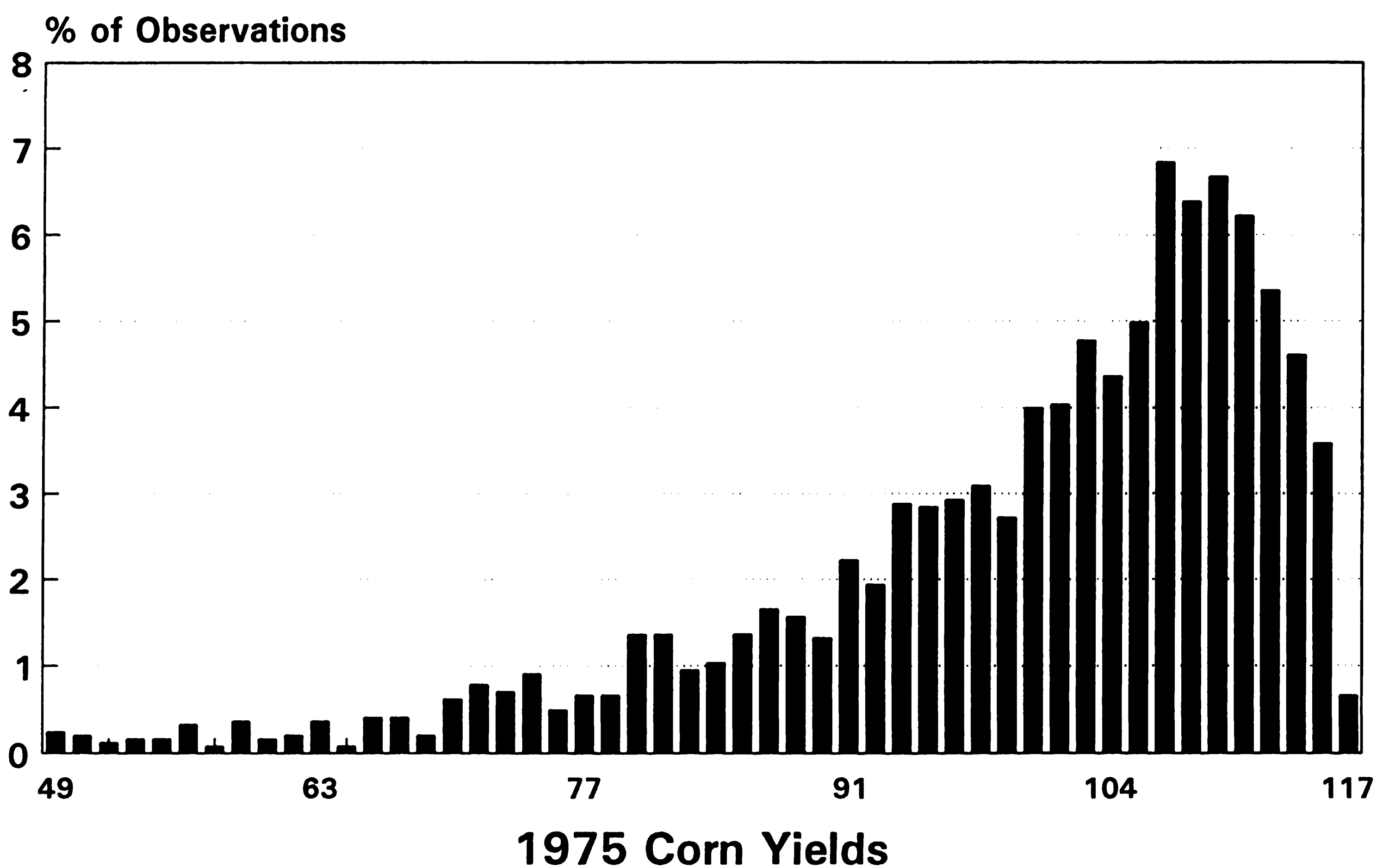


Figure 2



% of Observations
8
7
6
5
4
3
2
1
0
49
63
77
91
104
117
1975 Corn Yields

Figure 3

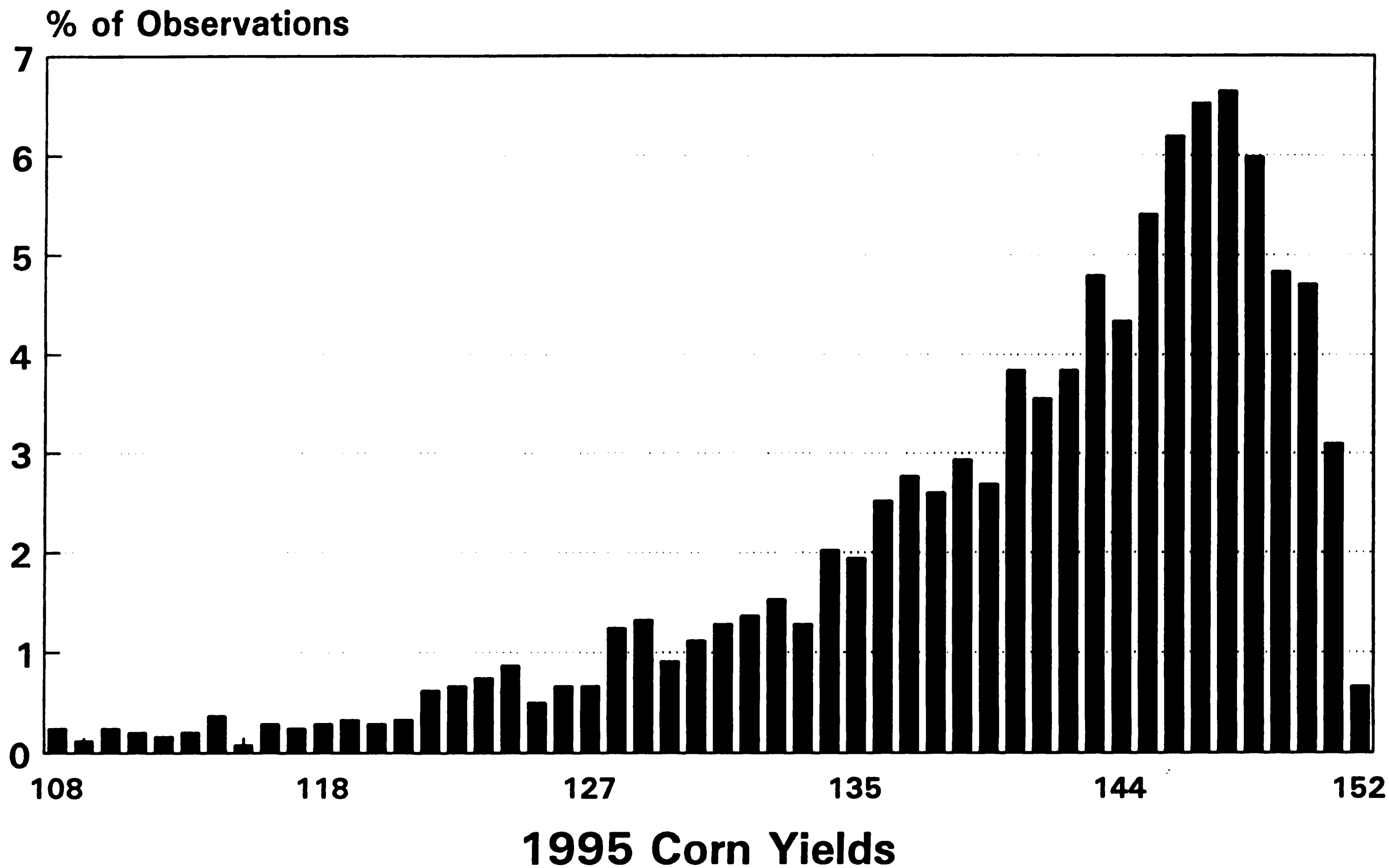

% of Observations
7
6
5
4
3
2
1
0
108
118
127
135
144
152
1995 Corn Yields

Figure 4

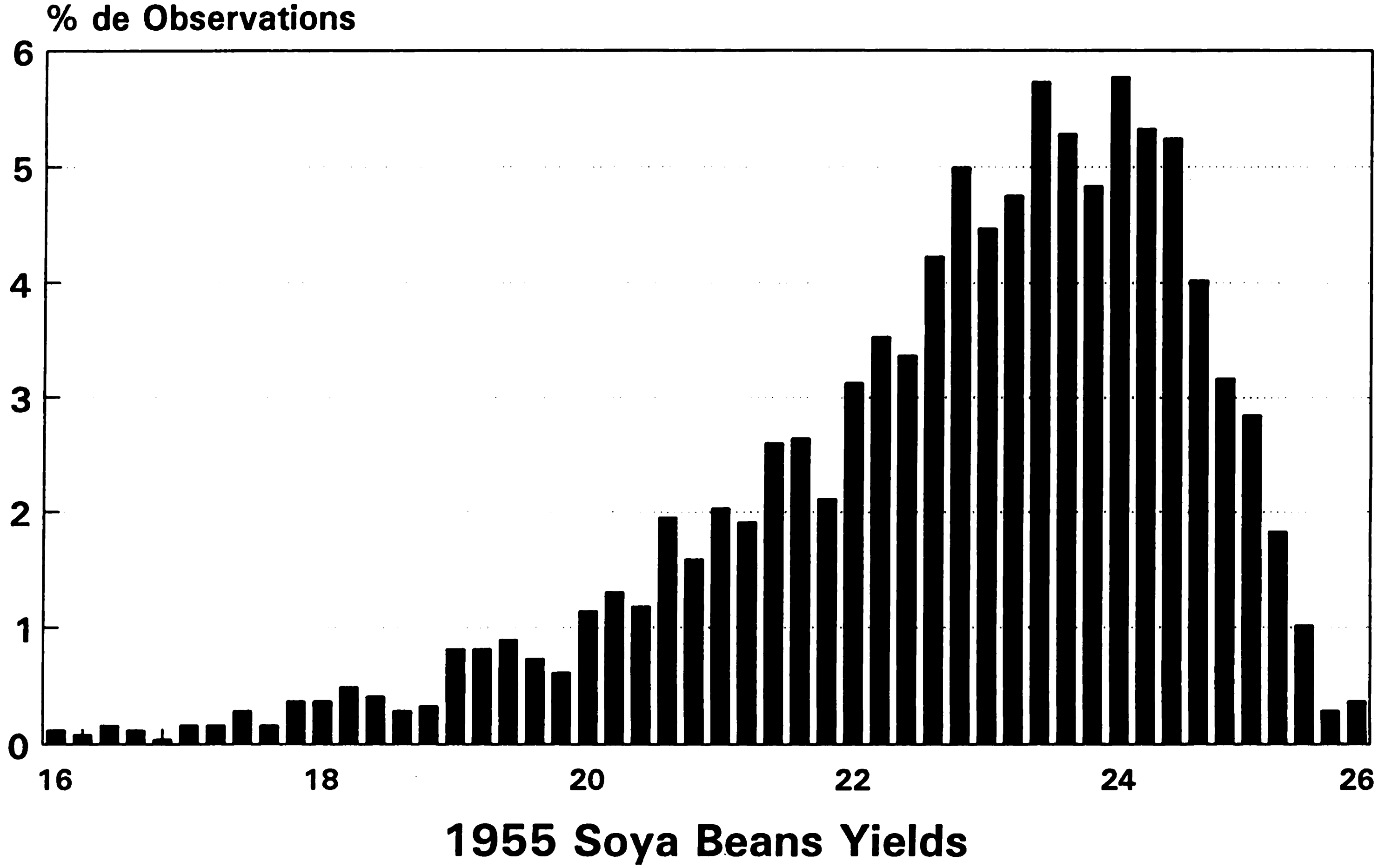

% de Observations
6
5
4
3
2
1
0
16
18
20
22
24
26
1955 Soya Beans Yields

Figure 5

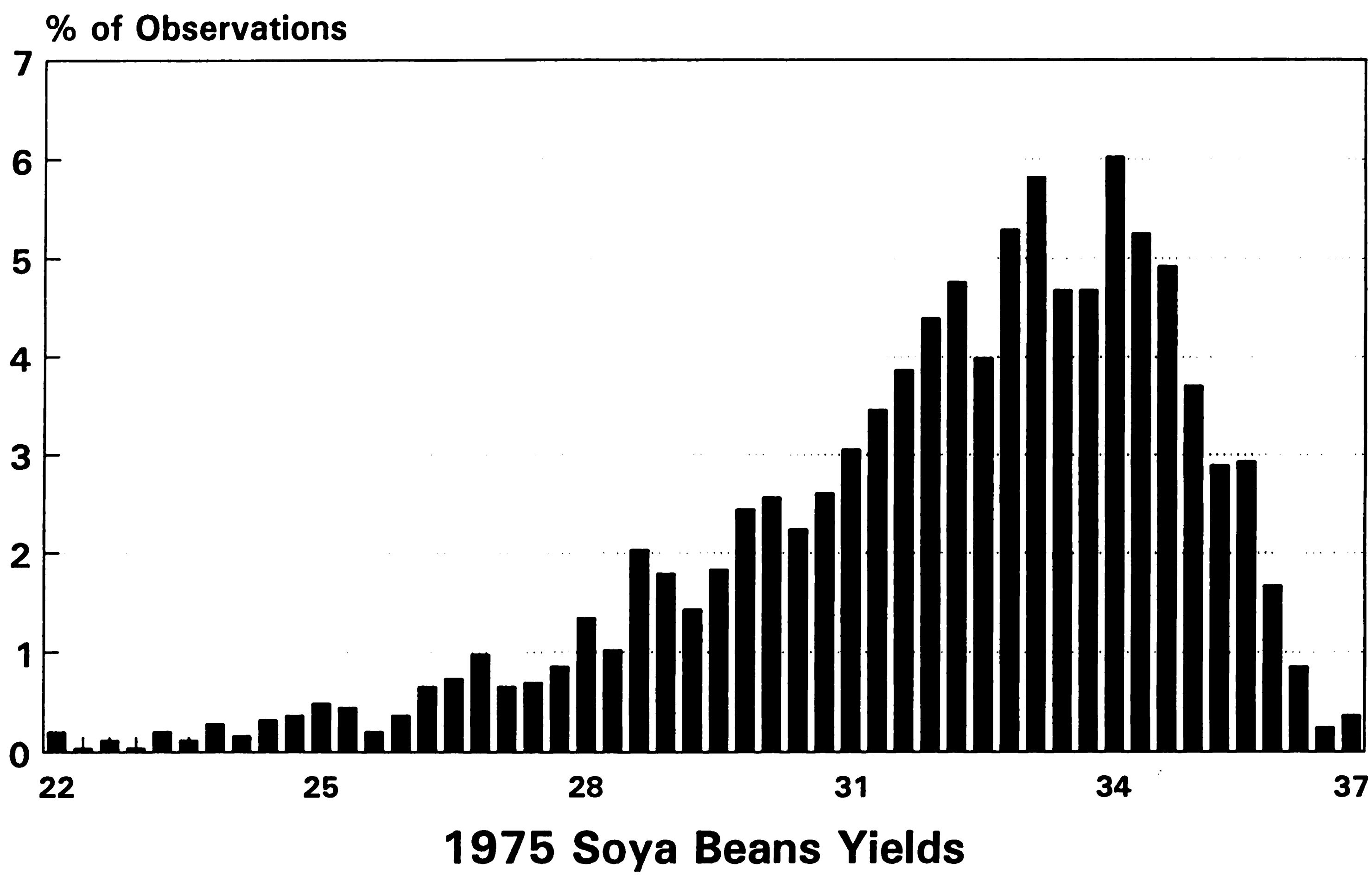
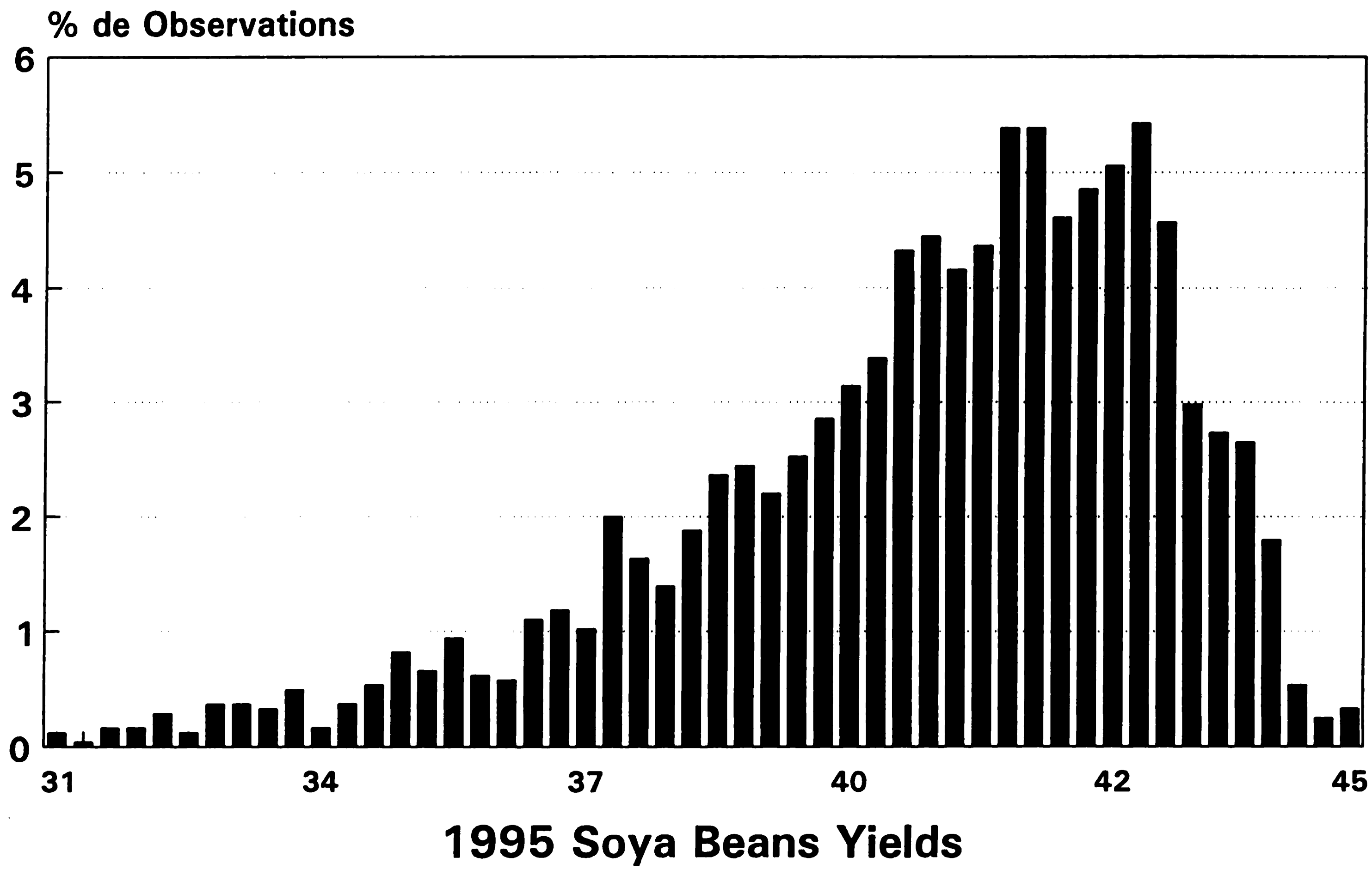

% of Observations
7
6
5
4
3
2
1
0
22
25
28
31
34
37
1975 Soya Beans Yields

Figure 6



% de Observations
6
5
4
3
2
1
0
31
34
37
40
42
45
1995 Soya Beans Yields

APPENDIX

National Agricultural Statistical Service crop yields (bushels) for the Corn Belt.

Year	Corn	Soya Beans	Wheat
50	48.98	23.07	20.32
51	47.53	22.59	17.57
52	55.67	22.97	23.49
53	51.14	19.44	27.50
54	51.12	22.09	29.07
55	51.94	21.42	30.67
56	59.01	24.29	31.03
57	59.22	24.86	22.84
58	64.48	26.77	30.65
59	63.54	25.29	25.25
60	64.56	25.16	31.17
61	74.87	27.63	32.85
62	78.20	26.81	31.53
63	80.91	27.85	37.84
64	73.72	24.90	33.46
65	86.23	27.26	32.07
66	82.50	27.32	39.35
67	90.47	26.61	35.49
68	89.59	31.00	35.71
69	96.16	31.50	36.77
70	77.54	30.22	36.59
71	100.59	31.61	44.03
72	107.73	31.97	44.17
73	100.82	30.70	31.41
74	76.86	25.20	34.14
75	97.54	33.12	39.38
76	98.09	30.36	36.91
77	95.76	35.86	43.09
78	109.78	33.67	37.48
79	122.05	36.41	45.25
80	99.49	33.88	46.58
81	120.60	35.76	45.00
82	122.77	36.13	39.83
83	79.12	29.81	44.77
84	112.38	30.64	43.18
85	127.28	39.67	49.59
86	130.74	38.77	42.18
87	129.13	38.64	55.18
88	79.92	28.13	50.66
89	120.44	36.54	53.34